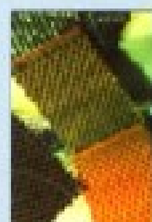
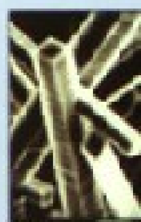
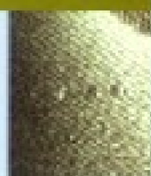
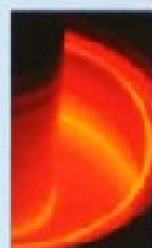
中国材料工程大典

中国机械工程学会　中国材料研究学会


中国材料工程大典编委会

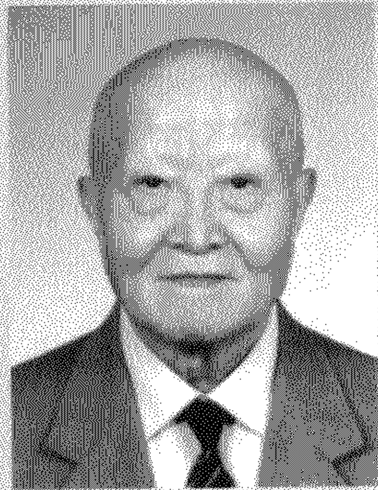
第1卷　材料工程基础


师昌绪　钟群鹏　李成功　主编


化学工业出版社

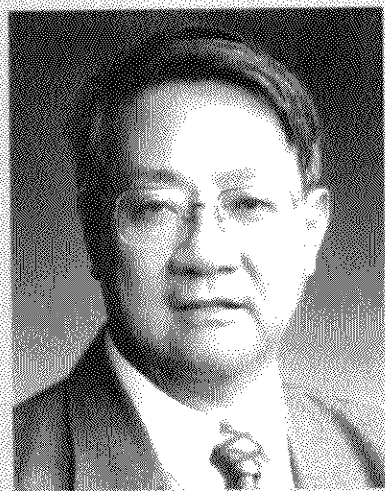
CHINA MATERIALS ENGINEERING CANON



师昌绪

院士

金属学及材料科学家
中国科学院院士
中国工程院院士
国家自然科学基金委员会特邀顾问



钟群鹏

院士

北京航空航天大学教授
中国机械工程学会副理事长
中国工程院院士



李成功

教授

北京航空材料研究院原总工程师
国家973计划材料领域专家咨询组副组长

39位两院院士和1200余位专家教授的知识与经验，编著成7000万字的科技工具书，准能助您获益！

CHINA MATERIALS ENGINEERING CANON

支持单位：

中华人民共和国科学技术部
国防科学技术工业委员会
国家自然科学基金委员会
中国科学技术协会
中国科学院
中国工程院



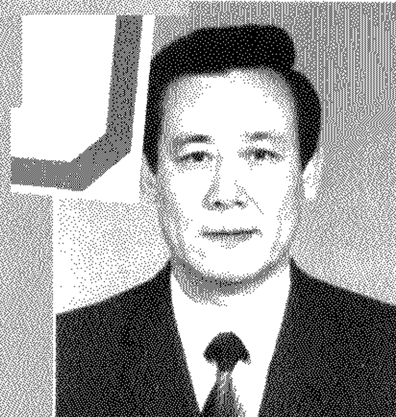
中国材料工程大典编委会



主任：路甬祥

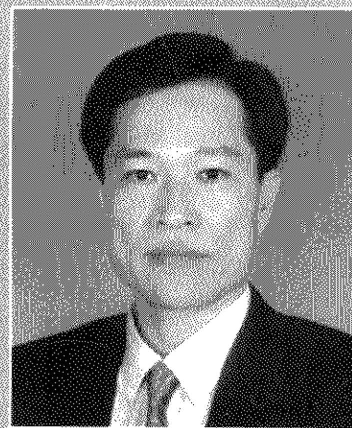


常务副主任：李成功



总策划：宋天虎





总策划：黄远东



总编辑：李骏带





中国材料工程大典编委会

杨　合（西北工业大学教授）
杨　武（上海材料研究所教授）
杨乃宾（北京航空航天大学教授）
杨才福（钢铁研究总院教授）
杨鸣波（四川大学教授）
杨忠民（钢铁研究总院教授）
杨晓华（福州大学教授）
杨海波（北京科技大学教授）
杨焕文（中国有色金属学会副秘书长、教授）
杨德仁（浙江大学教授）
李　强（福州大学教授）
李　晋（上海材料研究所教授）
李　楠（武汉科技大学教授）
李长久（西安交通大学教授）
李龙土（清华大学教授、院士）
李成功（中国材料研究学会荣誉理事、教授）
李光福（上海材料研究所教授）
李志刚（华中科技大学教授）
李明哲（吉林大学教授）
李明辉（上海交通大学教授）
李学勇（科学技术部副部长）
李虹霞（洛阳耐火材料研究院教授）
李恒德（清华大学教授、院士）
李贺军（西北工业大学教授）
李海军（宁夏东方钽业股份有限公司高级工程师）
李骏带（中国材料工程大典编委会高级工程师）
李鹤林（石油天燃气公司管材研究所教授、院士）
严东生（中科院上海硅酸盐研究所教授、院士）
连克仁（苏州特种加工研究所教授）
肖亚庆（中国铝业公司总经理、教授）
吴　行（装甲兵工程学院教授）
吴　昆（哈尔滨工业大学教授）
吴　诚（上海材料研究所教授）
吴永声（四川大学教授）
吴伟仁（国防科工委科技与质量司司长、研究员）
吴性良（复旦大学教授）
吴科如（同济大学教授）
吴恩熙（中南大学教授）
吴谊群（中科院上海光学机械研究所研究员）
吴智华（四川大学教授）
吴德馨（中科院微电子所研究员、院士）
何光远（原机械工业部部长、教授）
何季麟（宁夏东方有色金属集团公司总裁、院士）
佟晓辉（中国热处理行业协会研究员）
邱　勇（清华大学教授）
邱冠周（中南大学副校长、教授）
邱德仁（复旦大学教授）
余金中（中科院半导体研究所研究员）
邹广田（吉林大学教授、院士）
汪明朴（中南大学教授）
沈　真（中航一集团623所研究员）
沈万慈（清华大学教授）
沈德忠（清华大学教授、院士）
宋天虎（中国机械工程学会秘书长、教授）
张　力（国防科工委经济与协调司副司长、研究员）
张　扬（四川大学教授）
张　华（贵州安大航空锻造公司副总经理）
张　杰（北京科技大学教授）
张　金（中国锻压协会秘书长、教授）
张　峥（北京航空航天大学教授）
张子龙（北京航空材料研究院高级工程师）
张用宾（中国建筑材料科学研究院教授）
张立同（西北工业大学教授、院士）
张永俐（昆明贵金属研究所教授）
张吉龙（中国铝业公司教授）
张旭初（中国材料工程大典编委会教授）
张佐光（北京航空航天大学教授）
张晋远（钢铁研究总院教授）
张康侯（昆明贵金属研究所教授）
张道中（中科院物理研究所教授）
张新民（中南大学教授）
陆燕荪（原机械工业部副部长）
陈　琦（沈阳铸造研究所教授）
陈文哲（福州大学教授）
陈世朴（上海交通大学教授）
陈立泉（中科院物理研究所教授、院士）
陈运远（上海材料研究所教授）
陈志良（化学工业出版社编审）
陈国钧（钢铁研究总院教授）
陈治明（西安理工大学校长、教授）
陈南宁（北京钢铁设计研究总院教授）
陈祝年（山东大学教授）
陈晓慈（中国第二重型机械集团公司副总工程师）
陈涌海（中科院半导体研究所研究员）
陈祥宝（北京航空材料研究院研究员）
陈超志（中国机械工程学会高级工程师）
林慧国（钢铁研究总院教授）
欧阳世翕（中国建筑材料科学研究院教授）
卓尚军（中科院上海硅酸盐研究所研究员）
易建宏（中南大学教授）
罗祥林（四川大学教授）
罗豪甦（中科院上海硅酸盐研究所教授）
果世驹（北京科技大学教授）
周　廉（西北有色金属研究院教授、院士）
周伟斌（化学工业出版社副社长、编审）
周国庆（化学工业出版社副总编辑、编审）
郑有炓（南京大学教授、院士）
柳玉起（华中科技大学教授）

鸣　　谢

在编写过程中，得到以下部门和单位的支持和协作，使《中国材料工程大典》得以顺利编撰完成。在此，中国材料工程大典编委会代表全体作者表示衷心感谢！

支持部门：中华人民共和国科学技术部
国防科学技术工业委员会
国家自然科学基金委员会
中国科学技术协会
中国科学院
中国工程院

协作单位

钢铁研究总院
北京有色金属研究总院
北京航空材料研究院
中国建筑材料科学研究院
中国科学院金属研究所
中国科学院上海硅酸盐研究所
上海宝钢集团公司
中国石油化工集团公司
中国铝业公司
清华大学
中南大学
太原钢铁集团公司
西北有色金属研究院
宁夏东方有色金属集团公司
华中科技大学
中国第二重型机械集团公司
北京科技大学
北京航空航天大学
中国航天集团第703研究所
中国特种设备检测研究中心
哈尔滨工业大学
贵州安大航空锻造公司
东北大学
西安重型机械研究所
中国科学院半导体研究所
四川大学
北京航空制造工程研究所
中国科学院物理研究所
西北工业大学
北京矿冶研究总院
沈阳铸造研究所
江苏法尔胜公司

篇　目

第1卷　材料工程基础

主编：师昌绪院士　钟群鹏院士　李成功教授

第1篇	材料科学与工程概论	主编：师昌绪院士　李成功教授　刘治国教授
第2篇	材料成形基础理论	主编：董湘怀教授
第3篇	材料成形数值模拟	主编：柳玉起教授
第4篇	材料成形优化设计方法	主编：赵国群教授
第5篇	材料失效分析	主编：钟群鹏院士　李鹤林院士　张　峥教授
第6篇	材料强度设计	主编：谢里阳教授　王永岩教授

第2、3卷　钢铁材料工程（上、下）

主编：干　勇院士　田志凌教授　董　瀚教授　冯　涤教授　王新林教授

第1篇	概论	主编：干　勇院士
第2篇	钢铁牌号表示方法	主编：林慧国教授
第3篇	铁	主编：祖荣祥教授
第4篇	铸铁与铸钢	主编：陈　琦教授　彭兆弟教授
第5篇	非合金钢	主编：杨忠民教授
第6篇	低合金钢	主编：董　瀚教授　雍歧龙教授　刘清友教授　杨才福教授
第7篇	超细晶钢	主编：刘正才教授
第8篇	镍基和铁镍基耐蚀合金	主编：康喜范教授
第9篇	电热合金	主编：唐昌世教授
第10篇	高温合金	主编：冯　涤教授
第11篇	金属功能材料	主编：王新林教授　陈国钧教授
第12篇	钢铁焊接材料	主编：田志凌教授
第13篇	合金钢	主编：董　瀚教授

第4、5卷　有色金属材料工程（上、下）

主编：黄伯云院士　李成功教授　石力开教授　邱冠周教授　左铁镛院士

第1篇	概论	主编：黄伯云院士　邱冠周教授
第2篇	铝及铝合金	主编：田荣璋教授　肖亚庆教授
第3篇	镁及镁合金	主编：黎文献教授
第4篇	铜及铜合金	主编：汪明朴教授　尹志民教授
第5篇	镍、钴及其合金	主编：唐仁政教授
第6篇	锌、铅、锡及其合金	主编：田荣璋教授
第7篇	钛及钛合金	主编：李成功教授　马济民教授　邓　炬教授
第8篇	钨、钼及其合金	主编：王德志教授　潘叶金教授
第9篇	硬质合金	主编：吴恩熙教授
第10篇	钽、铌及其合金材料	主编：何季麟院士
第11篇	铍、锆、铪及其合金材料	主编：刘建章教授　聂大钧教授
第12篇	贵金属及其合金材料	主编：孙加林教授　张康侯教授　宁远涛教授　张永俐教授
第13篇	有色金属层状复合材料	主编：张新明教授　谢建新教授
第14篇	有色金属新材料	主编：石力开教授　左铁镛院士

第6、7卷　高分子材料工程（上、下）

主编：杨鸣波教授　唐志玉教授

第1篇	概论	主编：杨鸣波教授　唐志玉教授
第2篇	塑料工程	主编：吴智华教授
第3篇	有机纤维	主编：叶光斗教授　徐建军教授
第4篇	橡胶工程	主编：谢邦互教授

第 5 篇	高分子胶粘剂	主编：朱如瑾教授		
第 6 篇	功能高分子	主编：罗祥林教授		
第 7 篇	皮革材料	主编：张　扬教授		

第 8、9 卷　无机非金属材料工程（上、下）

主编：江东亮院士　李龙土院士　欧阳世翕教授　施剑林教授

第 1 篇	概论	主编：江东亮院士		
第 2 篇	结构陶瓷	主编：江东亮院士	黄校先教授	潘振甦教授
第 3 篇	功能陶瓷	主编：李龙土院士	徐廷献教授	
第 4 篇	传统陶瓷	主编：同继锋教授		
第 5 篇	玻璃	主编：马眷荣教授		
第 6 篇	晶体材料	主编：罗豪甦教授	仲维卓教授	
第 7 篇	无机涂层材料	主编：丁传贤院士	赵金榜教授	
第 8 篇	耐火材料	主编：李　楠教授	张用宾教授	李虹霞教授
第 9 篇	碳、石墨材料	主编：李龙土院士	沈万慈教授	
第 10 篇	水泥与混凝土	主编：隋同波教授		
第 11 篇	其他新型无机材料	主编：施剑林教授		

第 10 卷　复合材料工程

主编：益小苏教授　杜善义院士　张立同院士

第 1 篇	复合材料导论	主编：益小苏教授		
第 2 篇	复合材料用增强体材料	主编：冯春祥教授	楚增勇教授	
第 3 篇	聚合物基体材料	主编：陈祥宝教授		
第 4 篇	纺织复合材料	主编：丁　辛教授		
第 5 篇	复合材料界面	主编：黄玉东教授		
第 6 篇	工业聚合物基复合材料与玻璃钢	主编：刘其贤教授	刘占阳教授	高红梅教授
第 7 篇	先进树脂基复合材料	主编：杨乃滨教授		
第 8 篇	热塑性聚合物基复合材料	主编：张　忠教授	傅绍云教授	
第 9 篇	金属基复合材料	主编：耿　林教授	吴　昆教授	
第 10 篇	陶瓷（玻璃）基复合材料	主编：张立同院士	黄　勇教授	
第 11 篇	碳基复合材料	主编：张立同院士	李贺军教授	
第 12 篇	水泥基复合材料	主编：吴科如教授		
第 13 篇	复合材料力学问题与设计	主编：杜善义院士	梁　军教授	
第 14 篇	复合材料结构设计与分析	主编：沈　真教授		
第 15 篇	复合材料性能实验、表征与质量控制	主编：张佐光教授	张子龙教授	
第 16 篇	功能复合材料与新型复合材料	主编：刘献明教授	傅绍云教授	张　忠教授

第 11、12、13 卷　信息功能材料工程（上、中、下）

主编：王占国院士　陈立泉院士　屠海令教授

第 1 篇	概论	主编：王占国院士		
第 2 篇	半导体硅材料	主编：杨德仁教授		
第 3 篇	集成电路制造技术	主编：吴德馨院士	刘　明研究员	
第 4 篇	硅基异质结构材料和器件	主编：余金中研究员		
第 5 篇	化合物半导体材料	主编：屠海令教授	赵有文研究员	
第 6 篇	宽带隙半导体及其应用	主编：郑有炓院士		
第 7 篇	半导体低维结构和量子器件	主编：陈涌海研究员	叶小玲教授	王占国院士
第 8 篇	存储材料	主编：顾冬红研究员	吴谊群研究员	
第 9 篇	显示材料	主编：邱　勇教授	应根裕教授	
第 10 篇	通信光纤材料及其工艺	主编：赵梓森院士		
第 11 篇	全固态激光器及相关材料	主编：许祖彦院士	沈德忠院士	
第 12 篇	稀土磁性材料与自旋电子材料	主编：刘治国教授		
第 13 篇	超导材料	主编：陈立泉院士	蕲常青教授	
第 14 篇	传感器材料	主编：陈治明教授	雷天民教授	
第 15 篇	红外材料	主编：储君浩教授		

第16篇	先进储能材料	主编：陈立泉院士
第17篇	一维纳米材料和纳米结构	主编：张立德教授　解思深院士
第18篇	发光材料	主编：石春山研究员
第19篇	微加工技术	主编：冯　稷教授
第20篇	光子晶体	主编：张道中教授

第14卷　粉末冶金材料工程

主编：韩凤麟教授　马福康教授　曹勇家教授

第1篇	概论	主编：韩凤麟教授
第2篇	金属粉末生产与特性	主编：韩凤麟教授　夏志华教授
第3篇	金属粉末性能测试与相应标准	主编：张晋远教授
第4篇	成形与固结	主编：果世驹教授
第5篇	后续加工与质量控制	主编：贾成厂教授
第6篇	粉末冶金材料	主编：曹勇家教授　马福康教授　易建宏教授
第7篇	粉末冶金材料应用与新发展	主编：王尔德教授　韩凤麟教授

第15卷　材料热处理工程

主编：樊东黎教授　潘健生院士　徐跃明研究员　佟晓辉研究员

第1篇	概论	主编：樊东黎教授
第2篇	材料热处理技术基础	主编：樊东黎教授
第3篇	材料热处理工艺	主编：徐跃明研究员
第4篇	热处理设备	主编：佟晓辉研究员
第5篇	材料热处理	主编：徐跃明研究员
第6篇	热处理CAD/CAM/CAE	主编：潘健生院士
第7篇	热处理清洁生产和安全	主编：樊东黎教授
第8篇	热处理质量控制与无损检测	主编：佟晓辉研究员

第16、17卷　材料表面工程（上、下）

主编：徐滨士院士　刘世参教授

第1篇	概论	主编：徐滨士院士　刘世参教授
第2篇	材料服役中表面的失效行为及防治	主编：涂善东教授
第3篇	表面覆层形成与结合机理	主编：徐滨士院士　朱绍华教授
第4篇	涂装	主编：吴　行教授
第5篇	热喷涂	主编：徐滨士院士　李长久教授
第6篇	堆焊	主编：董祖珏教授
第7篇	电镀与电刷镀	主编：马世宁教授
第8篇	化学镀与转化膜技术	主编：姜晓霞研究员　董首山研究员
第9篇	化学热处理	主编：董汉山教授
第10篇	热浸镀	主编：刘邦津教授
第11篇	气相沉积技术及功能薄膜材料制备	主编：吕反修教授
第12篇	高能束表面处理技术	主编：左铁钏教授
第13篇	纳米表面工程	主编：徐滨士院士
第14篇	封存与包装	主编：梁志杰高工
第15篇	表面工程技术设计	主编：徐滨士院士　朱绍华教授
第16篇	表面工程质量控制与检测	主编：史耀武教授

第18、19卷　材料铸造成形工程（上、下）

主编：柳百成院士　黄天佑教授

第1篇	概论	主编：柳百成院士
第2篇	铸造合金及其熔炼	主编：邢建东教授
第3篇	铸造成形工艺技术基础	主编：郭景杰教授
第4篇	砂型铸造	主编：黄天佑教授
第5篇	特种铸造	主编：姜不居教授

第 6 篇　铸造成形 CAD/CAE　主编：熊守美教授
第 7 篇　铸造生产质量检测及控制　主编：黄天佑教授

第 20、21 卷　材料塑性成形工程（上、下）

主编：胡正寰院士　夏巨谌教授

第 1 篇　概论　主编：夏巨谌教授　张　金教授
第 2 篇　锻造成形　主编：夏巨谌教授　郭会光教授
第 3 篇　板料冲压成形　主编：杨　合教授　华　林教授　刘郁丽教授
第 4 篇　板型管轧制成形　主编：张　杰教授　杨海波教授　施东成教授　陈南宁教授
第 5 篇　零件轧制成形　主编：胡正寰院士　华　林教授
第 6 篇　特种锻造　主编：王高潮教授
第 7 篇　板管特种成形　主编：李明哲教授　蔡中义教授
第 8 篇　型材挤压成形　主编：夏巨谌教授　闫　洪教授
第 9 篇　塑性成形 CAD/CAM　主编：李志刚教授
第 10 篇　塑性成形质量控制与检测　主编：吕　炎教授

第 22、23 卷　材料焊接工程（上、下）

主编：史耀武教授

第 1 篇　概论　主编：史耀武教授
第 2 篇　材料焊接加工技术基础　主编：史耀武教授
第 3 篇　焊接方法与设备　主编：史耀武教授　殷树言教授
第 4 篇　材料焊接　主编：史耀武教授　任家烈教授
第 5 篇　焊接生产过程自动化　主编：蒋力培教授
第 6 篇　焊接结构设计　主编：陈祝年教授
第 7 篇　焊接结构制造　主编：史耀武教授
第 8 篇　焊接生产质量管理与无损检测　主编：解应龙教授
第 9 篇　焊接结构服役与再制造　主编：史耀武教授

第 24、25 卷　材料特种加工成形工程（上、下）

主编：王至尧研究员

第 1 篇　概论　主编：齐从谦教授
第 2 篇　材料电火花成形加工技术　主编：刘晋春教授　白基成教授　郭永丰教授
第 3 篇　材料数控电火花线切割技术　主编：李明辉教授
第 4 篇　材料电化学加工技术　主编：徐家文教授
第 5 篇　材料高能束流加工技术　主编：王亚军教授
第 6 篇　快速原型与快速制造　主编：颜永年教授
第 7 篇　电加工机床质量控制与检测　主编：连克仁教授

第 26 卷　材料表征与检测技术

主编：徐祖耀院士　黄本立院士　鄢国强教授

第 1 篇　概论　主编：徐祖耀院士　黄本立院士　陈文哲教授　鄢国强教授　朱万森教授
第 2 篇　化学成分分析方法　主编：黄本立院士　吴　诚教授　方禹之教授　邱德仁教授　吴性良教授
第 3 篇　常用材料化学成分分析　主编：鄢国强教授　马冲先教授　卓尚军研究员
第 4 篇　材料物理性能测试　主编：陈文哲教授　李　强教授
第 5 篇　材料力学性能测试　主编：陈文哲教授　陈运远教授　杨晓华教授
第 6 篇　材料化学性能测试　主编：杨　武教授　李光福教授
第 7 篇　金相分析　主编：唐汝钧教授　李　晋教授
第 8 篇　无损检测　主编：王务同教授　杨晓华教授
第 9 篇　X 射线衍射分析　主编：漆　玄教授　蒋建中教授
第 10 篇　电子显微分析　主编：陈世朴教授　孙　坚教授
第 11 篇　核技术分析及其他检测与表征技术　主编：陈世朴教授　梁　齐教授

中国材料工程大典

中国机械工程学会 中国材料研究学会



中国材料工程大典编委会

第1卷 材料工程基础

师昌绪 钟群鹏 李成功 主编



化学工业出版社

·北 京·

（京）新登字 039 号

内容简介

中国材料工程大典是中国机械工程学会和中国材料研究学会共同组织全国 39 位院士、百余位各学科带头人、千余位材料工程专家共同执笔编写，全面反映当今国内外材料工程领域发展的最新资料和最新成果，集实用性、先进性和权威性于一体的大型综合性工具书。中国材料工程大典包括材料工程基础、钢铁材料工程、有色金属材料工程、高分子材料工程、无机非金属材料工程、复合材料工程、信息功能材料工程、粉末冶金材料工程、材料热处理工程、材料表面工程、材料铸造成形工程、材料塑性成形工程、材料焊接工程、材料特种加工成形工程、材料表征与检测技术等内容，涵盖了材料工程的各个领域，将最新的实用数据（特别是与国际接轨的标准数据）、图表与先进实用的科研成果系统地集合起来，并附应用实例，充分展示了材料工程各领域的现状和未来。中国材料工程大典不仅可以满足现代企业正确选材，合理用材，应用先进的材料成形加工技术，提高产品质量和性能，降低产品成本，增强产品市场竞争力的需要，而且对推动中国材料科学与材料成形加工技术的不断创新，促进制造业的发展，提高我国制造业的竞争能力，具有重要的现实意义。

本书为第 1 卷，材料工程基础。主要内容包括材料科学与工程概论、材料成形基础理论、材料成形数值模拟、材料成形优化设计方法、材料失效分析、材料强度设计等。

本书主要供具有大专以上文化水平，从事材料工程研究的工程技术人员在综合研究和处理材料工程的各类技术问题时使用，起备查、提示和启发的作用，也可供研究人员、理工院校的有关师生参考。

图书在版编目（CIP）数据

中国材料工程大典．第 1 卷，材料工程基础/师昌绪，钟群鹏，李成功主编．—北京：化学工业出版社，2005.8
ISBN 7-5025-7303-8

Ⅰ．中…　Ⅱ．①师…②钟…③李…　Ⅲ．材料科学　Ⅳ．TB3

中国版本图书馆 CIP 数据核字（2005）第 087226 号

中国材料工程大典
第 1 卷
材料工程基础
中国机械工程学会
中国材料研究学会
中国材料工程大典编委会
师昌绪　钟群鹏　李成功　主编
责任编辑：周国庆　陈志良　李骏带
责任校对：陶燕华
封面设计：雷嘉琦
*
化学工业出版社出版发行
（北京市朝阳区惠新里 3 号　邮政编码：100029）
购书咨询：（010）64982530
（010）64918013
购书传真：（010）64982630
http://www.cip.com.cn
*
新华书店北京发行所经销
北京蓝海印刷有限公司印装
开本 880mm×1230mm 1/16　印张 59 1/4　字数 2770 千字
2006 年 1 月第 1 版　2006 年 1 月北京第 1 次印刷
ISBN 7-5025-7303-8
定价：170.00 元

序

材料是当代社会经济发展的物质基础，也是制造业发展的基础和重要保障。进入21世纪以来，随着经济全球化的发展和中国的崛起，现代制造业的重心正不断向中国转移。据统计，今天中国制造业直接创造国民生产总值的1/3以上，约占全国工业生产的4/5，为国家财政提供1/3以上的收入，占出口总额的90%。但是与发达国家相比，我国制造业的水平不高、自主创新能力不足、高端市场竞争力还不强。我国虽然已是世界制造业大国，但还不是世界制造业强国。在有关因素中，材料工程基础薄弱是制约我国制造业发展的关键因素。广义的材料工程包括材料制备、测试和加工成形过程。为了提高我国制造业的水平和竞争力，突破材料工程这个薄弱环节，中国机械工程学会和中国材料研究学会牵头，会同中国金属学会、中国化工学会、中国硅酸盐学会、中国有色金属学会、中国复合材料学会共同组织编撰《中国材料工程大典》（简称《材料大典》），其目的是力图为我国制造业提供一部集科学性、先进性和实用性于一体的综合性专业工具书。以满足广大科技工作者的迫切需求，为科技自主创新和我国制造业的崛起加强技术基础。

经过5年多的艰苦努力，《材料大典》终将出版了。这部共26卷约7000万字的巨著，是39位两院院士和1200余位参编专家教授们辛勤劳动的智慧结晶。有的作者为此牺牲了健康，如一位退休了的总工程师，为了把他多年的研究成果和实践经验写成书稿，由于长时间写作，导致眼睛视网膜脱落……。这种敬业精神与坚强毅力是值得我们学习铭记的。借此机会，我们要感谢中国金属学会、中国化工学会、中国硅酸盐学会、中国有色金属学会、中国复合材料学会的支持。这些学会的众多专家教授积极参与了《材料大典》编写工作，与中国机械工程学会和中国材料研究学会的专家教授一起完成这项艰巨任务，从而使《材料大典》在完整性与先进性、科学性与实用性的结合上得到了加强；我们要感谢科学技术部、国防科学技术工业委员会、国家自然科学基金委员会、中国科学技术协会、中国科学院、中国工程院，以及各协作单位对编写工作的大力支持和积极帮助；我们也要感谢师昌绪院士等顾问的殷切指导，他们在编委会的两次工作会议上提出了许多重要的意见和建议，平时也给予了经常关心和指导，使我们少走了许多弯路；我们还要对关心和支持《材料大典》编写工作的科研院所、院校、企业以及有关人员表示感谢。没有大家的支持与协同，就不可能有《材料大典》的成功编写和顺利出版。

《材料大典》既总结了10多年来在材料工程方面的最新数据、图表及科研成果，还汇集了国内外在材料工程方面的成熟经验和先进理念，它体现了科学性、先进性和实用性的结合。可供具有大专以上文化水平的有关工程技术人员查阅使用，也可供理工院校的师生参考。

编撰《材料大典》涉及范围广，难度大，书中不可避免地会存在一些缺点和不足之处，恳请各位读者指正。

中国机械工程学会理事长
中国材料工程大典编委会主任

2005年9月23日

前　　言

《材料工程基础》是《中国材料工程大典》26卷中的首卷，是全书其他各卷的共性基础要览，是涉及各类材料及其产品设计选材、制造加工成形以及失效分析等方面的基础。其主要内容涵盖了材料科学与工程概论、材料成形的基础理论、数值模拟和优化设计方法、材料的强度设计以及材料的失效分析等方面。

参加本卷编写工作的有国家自然科学基金委员会、北京航空材料研究院、北京航空航天大学、上海交通大学、中国材料研究学会、南京大学、华中科技大学、山东大学、东北大学、石油管材研究所等单位的40余位专家教授，并由我们三人担任本卷的主编。本卷共设6篇42章，共约270万字。各篇的主编如下：

第1篇	材料科学与工程概论	师昌绪　李成功　刘治国
第2篇	材料成形基础理论	董湘怀
第3篇	材料成形数值模拟	柳玉起
第4篇	材料成形优化设计方法	赵国群
第5篇	材料失效分析	钟群鹏　李鹤林　张　峥
第6篇	材料强度设计	谢里阳　王永岩

本书的编写工作遵循“科学性、先进性和实用性”的原则，以材料工程基础方面比较成熟的理论、方法和数据为主，同时参考了国内外材料工程基础方面的新进展，反映当代材料工程基础的先进水平。本书主要供具有大专以上文化水平的有关工程技术人员阅读使用，也可供理工科院校的有关师生参考。

感谢中国机械工程学会、中国材料研究学会、中国材料工程大典编委会、中国金属学会、中国化工学会、中国硅酸盐学会、中国有色金属学会、中国复合材料学会对本卷编写工作的指导与支持。由于编写时间紧迫和编者水平所限，对书中的不当之处，恳请读者指正并提出宝贵意见。

师昌绪
钟群鹏
李成功

2005年9月12日

目　录

第 1 篇

材料科学与工程概论

主 编 师昌绪 李成功 刘治国
编 写 师昌绪 李成功 刘治国
审 稿 中国材料工程大典编委会

中国机械工程学会
中国材料研究学会
中国材料工程大典编委会

第1章　材料科学与工程的提出与内涵

材料是人类用于制造物品、器件、构件、机器或其他产品的那些物质。材料是物质，但不是所有物质都可以称为材料，如燃料和化学原料、工业化学品、食物和药物，一般都不算是材料。但是这个定义并不那么严格，如炸药、固体火箭推进剂，一般称之为“含能材料”，因为它属于火炮或火箭的组成部分。

材料是人类赖以生存和发展的物质基础。20世纪70年代人们把信息、材料和能源誉为当代文明的三大支柱。80年代以高技术群为代表的新技术革命，又把新材料、信息技术和生物技术并列为新技术革命的重要标志。这主要是材料与国民经济建设、国防建设和人民生活密切相关。

材料除了具有重要性和普遍性以外，还具有多样性。由于其多种多样，分类方法也就没有一个统一标准。从物理化学属性来分，可分为金属材料、无机非金属材料、有机高分子材料和不同类型材料组成的复合材料。从用途来分，又分为电子材料、航天航空材料、核材料、建筑材料、能源材料、生物材料等。更常见的两种分类方法则是结构材料与功能材料，传统材料（基础材料）与新型材料。结构材料是以力学性能为基础，以制造受力构件所用材料，当然，结构材料对物理或化学性能也有一定要求，如光泽、热导率、抗辐照、抗腐蚀、抗氧化等。功能材料则主要是利用物质的独特物理、化学性质或生物功能等而形成的一类材料。一种材料往往既是结构材料又是功能材料，如铁、铜、铝等。传统材料是指那些已经成熟且在工业中已批量生产并大量应用的材料，如钢铁、水泥、塑料等。这类材料由于其量大、产值高、涉及面广泛，又是很多支柱产业的基础，所以又称为基础材料。新型材料（先进材料）是指那些正在发展，且具有优异性能和应用前景的一类材料。新型材料与传统材料之间并没有明显的界限，传统材料通过采用新技术，提高技术含量，提高性能，大幅度增加附加值而成为新型材料；新型材料在经过长期生产与应用之后也成为传统材料。传统材料是发展新型材料和高技术的基础，而新型材料又往往能推动传统材料的进一步发展。

为了规范新型材料的涵义，2004年曾召开小型座谈会，认为具备以下条件之一的材料可称之为新型材料：

1）新出现或正在发展中的具有传统材料所不具备的优异性能的材料；

2）高技术发展需要，具有特殊性能的材料，例如电视机的阴极罩，材料是低碳钢，由于对材料质量和性能有特殊要求，必须经过特殊加工过程才能达到，从而普通碳钢成为新型材料；

3）由于采用新技术（工艺、装备）明显提高了性能，或出现新的功能的材料，如超级钢属于前一种类型；纳米氧化锌则属于后者。

1　材料是人类社会进步的里程碑

人类发展的历史证明，材料是社会进步的物质基础，是人类进步程度的主要标志，所以人类社会的进步以材料作为里程碑。纵观人类发现材料和利用材料的历史，每一种重要材料的发现和广泛利用，都会把人类支配和改造自然的能力提高到一个新水平，给社会生产力和人类生活水平带来巨大的变化，把人类的物质文明和精神文明向前推进一步。

早在一百万年以前，人类开始用石头做工具，使人类进入旧石器时代。大约一万年以前，人类知道对石头进行加工，使之成为精致的器皿或工具，从而使人类进入新石器时代。在新石器时代，人类开始用皮毛遮身。8000年前中国就开始用蚕丝做衣服，4500年前印度人开始种植棉花，这些都标志着人类使用材料促进文明进步。在新石器时代。人类已知道使用自然铜和天然金，但毕竟数量太少，分散细小，没有对人类社会产生重要影响。

大约在8000～9000年前，人类还处于新石器时代，已发明了用黏土成型，再火烧固化而成为陶器。陶器的出现，不但用于器皿，而且成为装饰品，是对精神文明的一大促进，历史上虽无陶器时代的名称，但其对人类文明的贡献是不可估量的。在烧制陶器过程中，偶然发现金属铜和锡，当然那时还不明白这是铜、锡的氧化物在高温下被碳还原的产物，进而又生产出色泽鲜艳、又能浇铸成型的青铜，从而使人类进入青铜时代。这是人类较大量利用金属的开始，也是人类文明发展的重要里程碑。世界各地开始青铜时代的时间各不相同，希腊在公元前3000年前，埃及在公元前2500年前，巴比伦在公元前19世纪中叶，印度大约在公元前3000年已广泛使用青铜器。中国的青铜器在公元前2700年已经使用了，至今约5000年的历史，到商周（公元前17世纪到公元前3世纪）进入了鼎盛时期，如河南安阳出土的重达875 kg的鼎、湖北隋县的编钟、西安青铜车马都充分反映了当时中国冶金技术水平和制造工艺的高超。

公元前13～14世纪前，人类已开始用铁，3000年前铁工具比青铜工具更为普遍，人类开始进入了铁器时代。中国最早出土的人工冶铁制品约在公元前9世纪。到春秋（公元前700～476年）末期，生铁技术有较大突破，遥遥领先于世界其他地区，如用生铁退火而制成韧性铸铁及生铁炼钢技术的发明，促进了中国生产力的大发展，对战国和秦汉农业、水利和军事的发展起到很大作用。早在公元2世纪中国的钢和丝绸已驰名罗马帝国。生铁技术在公元前5世纪即春秋末叶已经在黄河长江流域传播。这些技术于公元6～7世纪传入朝鲜半岛、日本和北欧，推动了整个世界文明的进步。

随着世界文明的进步，18世纪发明了蒸汽机，19世纪发明了电动机，对金属材料提出了更高要求，同时对钢铁冶金技术产生了更大的推动作用。1854年和1864年先后发明了转炉和平炉炼钢，使世界钢产量有一个飞速发展。如1850年世界钢产量为6万吨，1890年达2 800万吨，大大促进了机械制造、铁道交通及纺织工业的发展。随之电炉冶炼开始，不同类型的特殊钢相继问世，如1887年高锰钢、1900年18-4-1（W18Cr4V）高速钢、1903年硅钢及1910年奥氏体镍铬（Cr18Ni8）不锈钢，把人类带进了现代物质文明。在此前后，铜、铝也得到大量应用，而后镁、钛和很多稀有金属都相继出现，从而金属材料在整个20世纪占据了结构材料的主导地位。

随着有机化学的发展，19世纪末叶西方科学家仿制中国丝绸发明了人造丝，这是人类改造自然材料的又一里程碑。20世纪初，人工合成有机高分子材料相继问世，如1909年的酚醛树脂（电木），1920年的聚苯乙烯，1931年的聚氯乙烯及1941年的尼龙等，以其性能优异、资源丰富、建设投资少、收效快而得到迅速发展。目前世界三大有机合成材料（树脂、纤维和橡胶）年产量愈亿吨。而且有机材料的性能不断提高，附加值大幅度增加，特别是特种聚合物正向功能材料各个领域进军，显示其巨大的潜力。

陶瓷本来用作建筑材料、容器或装饰品等。由于其资源

丰富、密度小、高模量、高硬度、耐腐蚀、膨胀系数小、耐高温、耐磨等特点，到了20世纪中叶，通过合成及其他制备方法，做出各种类型的先进陶瓷（如 Si_3N_4、SiC、ZrO_2 等），成为近几十年来材料中非常活跃的研究领域，有人甚至认为"新陶器时代"即将到来。但由于其脆性问题难以解决，且价格过高，作为结构材料没有得到如钢铁或高分子材料一样的广泛应用。

复合材料是20世纪后期发展的另一类材料。众所周知，天然材料很多是复合材料，如木材、皮革、竹子等。事实上，人类很早就制造复合材料，如泥巴中混入碎麻或麦秆用以建造房屋，钢筋水泥是脆性材料和韧性材料的复合。近几十年来，利用树脂的易成型和金属韧性好，无机非金属材料的高模量、高强度、耐高温，做成了树脂基复合材料或金属基复合材料，前者已得到广泛应用。后者以其制作困难、价格高而受到一定限制。为了改善陶瓷的性能，也制成陶瓷基复合材料。碳是使用温度最高的材料（可达2 500℃），为了克服热震性能差，并提高其力学性能而制出的碳-碳复合材料已广泛用于军工，并扩展到民用。

图1.1-1以结构材料为主线，概述了材料的发展历史。可以看出，自19世纪中叶现代炼钢技术出现以后，金属材料的重要性急剧增加，一直到20世纪中叶，人工合成有机材料、陶瓷材料及先进复合材料迅速发展，金属材料的重要性逐渐下降，但一直到21世纪上半叶，金属材料仍将占重要位置。

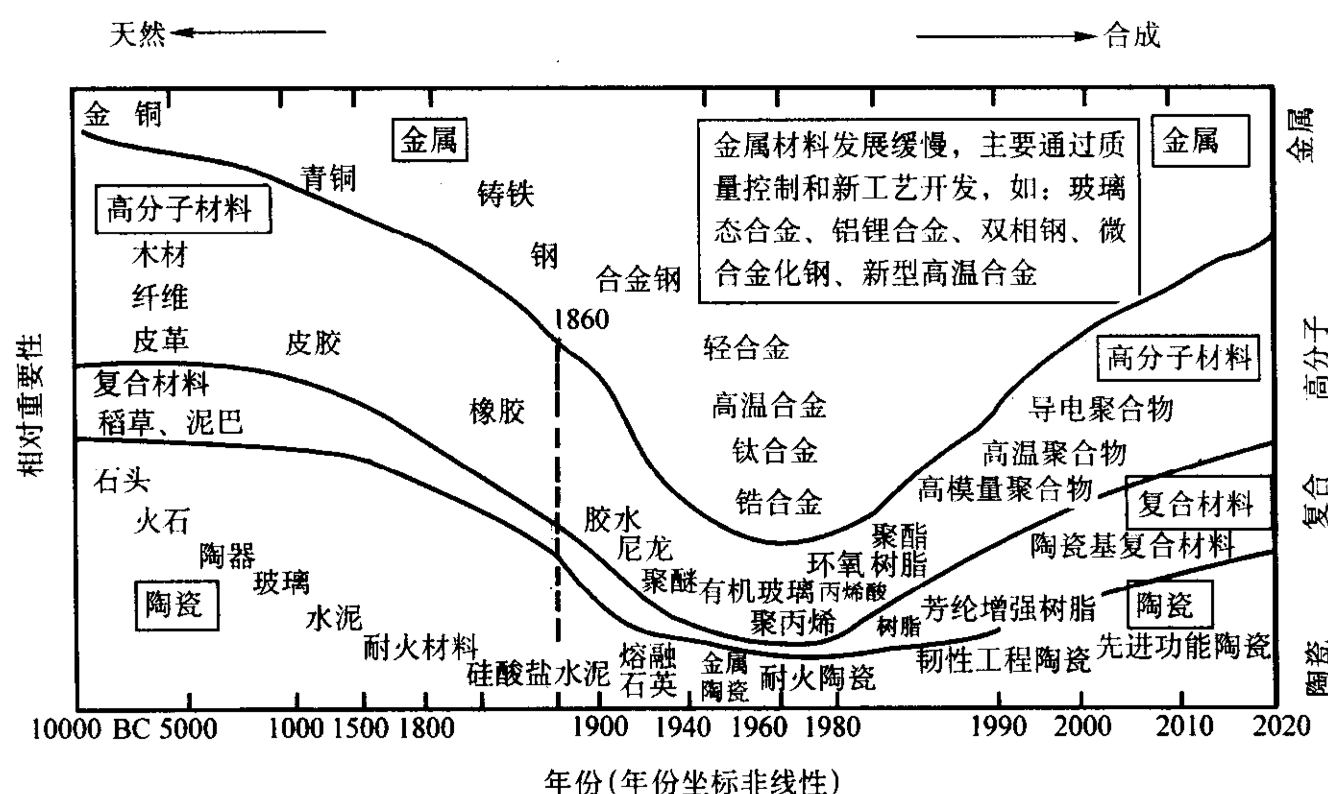


图1.1-1 各类结构材料相对重要性随年代变化

功能材料自古就受到重视，早在战国（公元前3世纪）已利用天然磁铁矿来制造司南，到宋代用钢针磁化制出了罗盘，为航海的发展提供了关键技术。功能材料是信息技术及自动化的基础，特别是半导体材料出现以后，加速了现代文明的发展。1947年发明了第一只具有放大作用的晶体管，10余年后又研制成功集成电路，使以硅材料为主体的计算机的功能不断提高，体积不断缩小，价格不断下降，加之高性能的磁性材料不断涌现，激光材料与光导纤维的问世，使人类社会进入了信息时代。因为硅是微电子技术的关键材料，所以有人称之为硅器件为代表的"硅材料时代"，再一次说明材料对人类文明起了关键的作用。

2 材料科学的形成与内涵

"材料"是早已存在的名词，但"材料科学"的提出只是20世纪60年代初的事。1957年前苏联人造卫星首先上天，美国朝野上下为之震惊，认为自己落后的主要原因之一是先进材料落后，于是在一些大学相继成立了十余个材料科学研究中心。采用先进的科学理论与实验方法对材料进行深入的研究，取得重要成果。从此，"材料科学"这个名词便开始流行。

"材料科学"的形成实际是科学技术发展的结果。

首先，固体物理、无机化学、有机化学、物理化学等学科的发展，对物质结构和物性的深入研究，推动了对材料本质的了解；同时，冶金学、金属学、陶瓷学、高分子科学等的发展也使对材料本身的研究大大加强，从而对材料的制备、结构与性能，以及它们之间的相互关系的研究也愈来愈深入，为材料科学的形成打下了比较坚实的基础。

其次，在材料科学这个名词出现以前，金属材料、高分子材料与陶瓷材料都已自成体系，目前复合材料也获得广泛应用，其研究也逐步深入。但它们之间存在着颇多相似之处，对不同类型材料的研究可以相互借鉴，从而促进学科的发展。如马氏体相变本来是金属学家提出来的，而且广泛地被用来作为钢热处理的理论基础，但在氧化锆陶瓷中也发现了马氏体相变现象，并用来作为陶瓷增韧的一种有效手段。又如材料制备方法中的溶胶-凝胶法，是利用金属有机化合物的分解而得到纳米级高纯氧化物粒子，成为改进陶瓷性能的有效途径。虽然不同类型的材料各有其专用测试设备与生产装置，但各类材料的研究检测设备与生产手段有颇多共同之处，例如显微镜、电子显微镜、表面测试及物性与力学性能测试设备等。在材料生产中，许多加工装置的原理也有颇多相通之处，可以相互借鉴，从而加速材料的发展。

最后，许多不同类型的材料可以相互替代和补充，能更充分发挥各种材料的优越性，达到物尽其用的目的。但长期以来，金属、高分子及无机非金属材料自成体系，缺乏沟通。由于互不了解，不利于发展创新，对复合材料的发展也极不利。

尽管从材料发展需要和共性来看，有必要形成一门材料科学，但是由于各类材料的学科基础不同，还存在不小的分歧，特别是无机材料与有机材料之间分歧较大，但由于软物质科学的崛起，还可以软硬兼顾，互相借鉴。

材料科学所包括的内容往往被理解为研究材料的组织、结构与性质的关系，探索自然规律，这属于基础研究。实际上，材料是面向实际、为经济建设服务的，是一门应用科学，研究与发展材料的目的在于应用，而材料又必须通过合理的工艺流程才能制备出具有实用价值的材料来，通过批量生产才能成为工程材料。所以，在"材料科学"这个名词出

现后不久，就提出了“材料科学与工程”。工程是指研究材料在制备过程中的工艺和工程技术问题。许多大学的冶金系、材料系也就此改变了名称，多数改为“材料科学与工程系”，偏重基础方面的就称“材料科学系”，偏重工艺方面的称“材料工程系”，也有不肯放弃“冶金”而称为“冶金与材料科学系”的，如英国的剑桥大学。同时，有关材料科学或材料科学与工程方面的杂志和书籍也应运而生。第一部《材料科学与工程百科全书》由美国麻省理工学院的科学家主编，由英国 Pergamon 自 1986 年陆续出版。它对材料科学与工程下的定义为：材料科学与工程就是研究有关材料组成、结构、制备工艺流程与材料性能和用途的关系的知识的产生及其作用。换言之，材料科学与工程是研究材料组成、结构、生产过程、材料性能与使用效能以及它们之间的关系。因而把组成与结构（composition-structure）、合成与生产过程（synthesis-processing）、性质（properties）及使用效能（performance）称之为材料科学与工程的四个基本要素（basic elements）。把四要素连结在一起，便形成一个四面体（tetrahedron），如图 1.1-2a。

考虑在四要素中的组成/结构并非同义词，即相同成分或组成通过不同的合成或加工方法，可以得出不同结构，从而材料的性质或使用效能都不会相同。因此，本人提出一个五个基本要素的模型，即成分（composition）、合成/加工（synthesis/processing）、结构（structure）、性质（properties）和使用效能（performance）。如果把它们连接起来，则形成一个六面体（hexahedron），如图 1.1-2b。

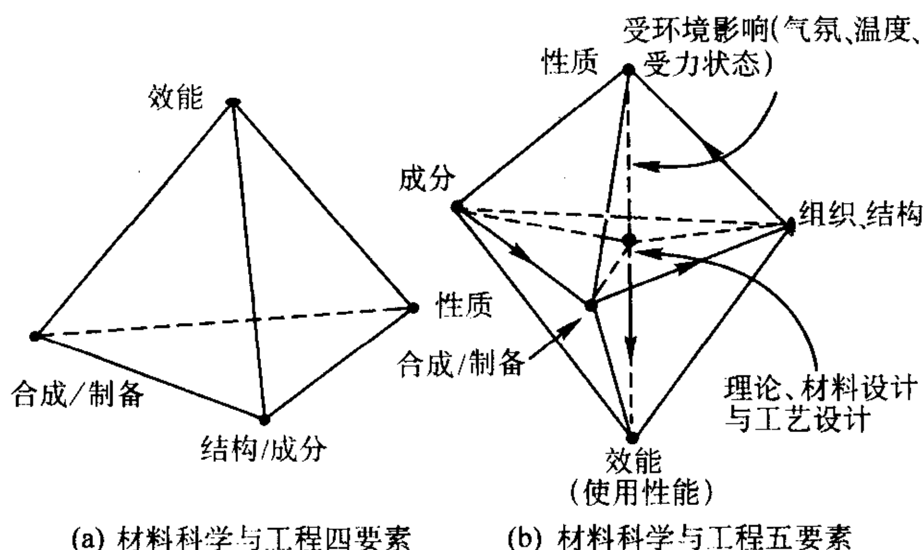


图 1.1-2　材料科学与工程要素图

材料科学与工程五要素模型的特点主要有两个：一是性质与使用效能有一个特殊的联系，材料的使用效能便是材料性质在使用条件下的表现。环境对材料性能的影响很大，如受力状态、气氛、介质与温度等。有些材料在一般环境下的性能良好，而在腐蚀介质下性能却下降显著；有的材料在光滑样品时表现很好，而在有缺口的情况下性能大为下降，特别有些高强度材料表现尤为突出，只要有一个划痕，就会造成灾害性破坏。因此，环境因素的引入对工程材料来说十分重要。二是材料理论和材料设计或工艺设计有了一个适当位置，它处在六面体的中心。因为这五个要素中的每一个要素，或几个相关要素都有其理论，根据理论建立模型（modeling），通过模型可以进行材料设计或工艺设计，以达到提高性能及使用效能、节约资源、减少污染或降低成本的最佳状态。这是材料科学与工程最终努力的目标。应该说明，目前国际流行的仍是四要素模型，五要素模型在国际上也有人引用。

根据以上所述，材料科学有三个重要属性：一是多学科交叉，它是物理学、化学、冶金学、金属学、陶瓷学、高分子化学及计算科学相互融合与交叉的结果，如生物医用材料要涉及医学、生物学及现代分子生物学等学科；二是一种与实际使用结合非常密切的科学，发展材料科学的目的在于开发新材料，提高材料的性能和质量，合理使用材料，同时降低材料成本和减少污染等；三是材料科学是一个正在发展中的科学，不像物理学、化学已经有一个很成熟的体系，材料科学将随各有关学科的发展而得到充实和完善。

3　先进材料是社会现代化的先导

材料既是人类社会进步的里程碑，又是社会现代化的物质基础与先导，特别是先进材料的研究、开发与应用反映着一个国家科学技术与工业水平。现仅就现代科学技术的发展与先进材料的关系举几个典型事例来说明。

3.1　电子技术的发展

从电子技术的发展可以看出材料所起的作用。1906 年发明了电子管，从而出现了无线电技术、电视机、电子计算机。1948 年发明了半导体晶体管，致使电子设备的小型化、轻量化、节能化及成本的降低、可靠性的提高与寿命的延长。1958 年出现了集成电路，使计算机及各种电工设备发生再一次飞跃。如以 1946 年电子管计算机与 1976 年微机的一些指标来对比，由于集成电路的采用，使计算机体积缩小了 30 万倍，功耗降低了 5 万多倍，质量降低了 6 万倍，平均故障率也大为减少，而且价格大幅度下降。这样为计算机的普及创造了条件。

随着芯片集成度的不断提高，单元体积和价格不断下降，图 1.1-3 说明芯片发展历程及与硅晶片尺寸的关系。可以看出，40 年间（1958～1998 年），芯片集成度提高了一百万倍，每单元价格下降到一百万分之一，这是因为单元价格与特征尺寸（feature length）的平方成正比，与晶片直径的平方成反比。特征尺寸的缩小，一方面与制作技术有关，另一方面与相关材料的不断改进有直接关系。为了适应集成度的不断提高，特征尺寸要不断缩小，晶片尺寸要不断增加，对硅单晶的质量要求也要不断提高。由于芯片密度的大幅度提高，用以制造计算机的性能大为提高，如 20 世纪末计算机的计算速度已逾每秒万亿次。信息存储是现代化的另一重要标志，其要求是容量大，密度高，易于快速随机存取，可擦除和反复使用，这就要求材料不断改进才能满足，迄今已出现磁存储、半导体存储和光存储等，如一张光盘可以存储 10 万幅图像或 50 万页文字信息量。计算机是工业自动化的关键，计算机控制的精度决定于传感器的敏感程度。因此，高精度、高灵敏度、性能稳定的各种类型的敏感材料便成为关键。

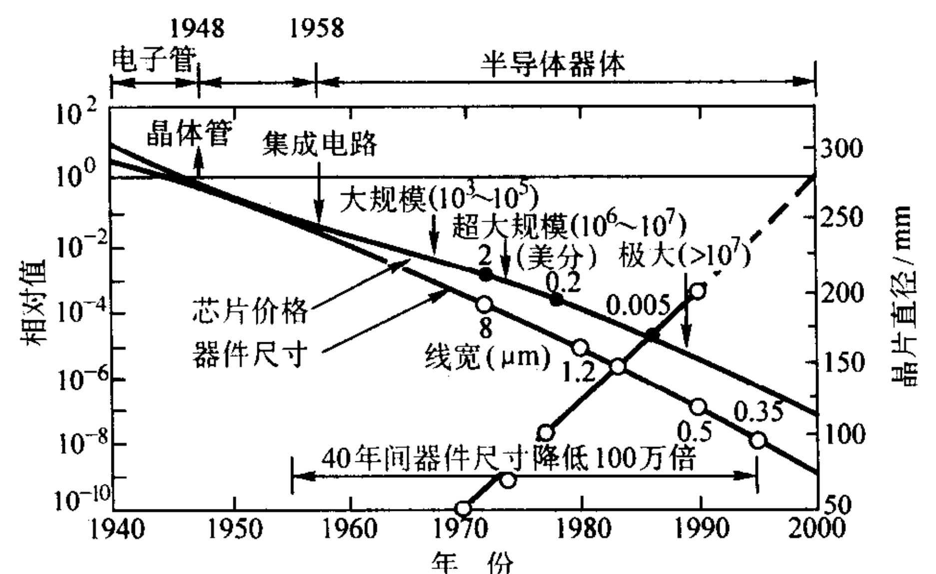


图 1.1-3　集成电路发展历程

3.2　光纤通信的诞生

1966 年高琨提出当光纤传输的损耗小于 20 dB/km 时可

实现光通信。1970年采用石英掺杂氧化锗等，达到了这一指标，而后光纤通信比集成电路发展更快，代替了同轴电缆。因为其信息容量比同轴电缆大几个量级，而且质量轻，消耗原材料少，保密性强，又抗电磁干扰，所以几十年来，从连接各大陆的海底电缆到很多家庭，都遍布了光缆，成为现代信息高速公路与信息网络化的重要组成部分。光纤通信除了光导纤维以外，还要有激光器和探测器，这些都是化合物半导体。表1.1-1为光纤通信系统的发展历程，除第四代尚在研制阶段外，其余都达到应用状态。

表1.1-1 光纤通信系统的发展历程

技术指标	第一代（短波长）	第二代（长波长）	第三代（长波长）	第四代（超长波长）
波长/μm	0.83	1.3	1.55	2~5
光纤	熔石英	熔石英	熔石英（+铒）	氟化物玻璃
激光器	GaAlAs	InGaAsP	InGaAsP	GaInAsSb
探测器	Si	Ge，InGaAsP	InGaAsP	InAsGaSb
损耗/$dB \cdot km^{-1}$	2~3	0.5~1	<0.1	0.1~0.001
传输距离/km	约10	约30	约400	>500

除光纤通信外，通信设备小型化是“漫游电脑”时代的必要条件，因此，促进了低电压、低功耗、高速集成电路的发展。

另外，计算机系统的存储和显示莫不与材料科学的进步相关。

3.3 航空航天技术的进步

21世纪全球经济的一体化，一靠信息网络的大发展，形成全球性的电子商业；二是运输工具的高效、远程、大容量。前者是电子及光电技术发展的结果，后者是航空技术与水上、地面运输工具不断改进的结果，这些都与材料有着密切联系。以飞机而论，由于战争的需要，20世纪40年代出现了喷气技术，这种技术的实现是以高温材料及高性能结构材料为依托的。高温合金的发展，使涡轮前温度不断提高；高比强度、高比刚度结构材料的不断进步，使今天大型客机的安全性及有效载荷大为提高，续航时间不断延长，飞机发动机寿命显著延长，油耗不断下降。有人对歼击机的发展有一个估计，飞机性能的改善有2/3依靠于材料，航空发动机性能的提高在很大程度上也有赖于材料的改进。由此说明材料对飞机的重要性。地面运输工具如汽车和火车，水上运输工具如轮船的发展莫不与材料有密切关系。

航天与卫星是科学技术现代化的重要标志之一。对材料的要求除了高比强度和高比刚度之外，还需要耐超高温、抗辐射、耐粒子云以及原子氧侵蚀等的材料，对比强度和比刚度的要求更为苛刻，因为航天飞行器每减重1 kg，可使运载火箭减轻几百千克。这就说明材料在航天技术的重要性。

从以上几个方面说明，材料，特别是先进材料与现代文明的关系十分密切，而且先进材料是先进技术的先导。没有半导体和其他功能材料，就不会有今天的信息社会，没有高温和超高温材料及高比强度、高比刚度材料，就不会有今天的航空航天技术，从而全球经济的一体化、人类社会的频繁往来就会发生很大困难。

4 传统材料（基础材料）在国民经济中的地位与可持续发展

谈到传统材料往往和“夕阳工业”联系在一起。实际上传统材料与国民经济支柱产业是密不可分的。钢铁曾经是衡量一个国家实力的重要标志，今天在一些工业发达国家也仍把钢铁视为支柱产业，因为钢有不可代替的优良性能，其价格又十分低廉，表1.1-2为一些结构材料的价格与其比强度的关系。钢的价格与比强度的比值仅高于木材，而比任何其他材料都低。同时根据多年统计结果，很多机械产品和材料的价格都在上涨，而钢的价格由于其工艺的不断改进、降低能耗和提高劳动生产率而长期未上涨。建筑行业在任何一个国家都被视为支柱产业，建筑材料是建筑行业的重要组成部分。它涉及人民生活的提高，量大面广。水泥、玻璃、木材以及砖、瓦、石材等属于传统产业。汽车工业在很多国家也被视为支柱产业，但是所用材料中钢、铁、铝、塑料及玻璃等占了汽车用结构材料的90%~95%。人工合成纤维、树脂、塑料、橡胶在国民经济中占有非常重要位置，而且逐年增加，这都属于传统产业。此外，机械制造、造船、机车等莫不是以钢铁及其他传统材料为基础的，所以说传统材料是国民经济的基础，不能稍有忽视。

表1.1-2 一些结构材料的价格与其比强度的关系

材料	钢	铝	钛	水泥	Al_2O_3	碳纤维	木材	聚丙烯
比强度	5.2	11.1	13.3	0.8	97.4	160.9	12.5	3.0
价格/比强度（相对值）	1	3.0	16.7	6.4	4.0	5.2	0.8	3.8

另外，传统材料有以下几个特点。

1）传统材料量大面广。传统材料只要有一点改进，收益就十分可观。据估计，美国道路与桥梁的使用寿命增加1%，其收益可达300亿美元。世界易拉罐每年消耗量以600亿元计，其造价每个省1分，则有6亿元的收益。所以对传统材料性能的改进和生产成本的下降必须给以足够的重视。

2）传统材料是矿产资源消耗大户。根据专家估计，全世界铁矿资源最多能维持200~300年，铝矿近百年，而铜、铅、锌矿只能够几十年。因此，除了提高矿产的收得率外，要开拓新的矿源，探索某些关键材料的代用品。如铝和镁有大致相似的性质，而镁在地壳表层储量仅次于铁、铝，居第三位，但在海水中储量巨大（2.1×10^{15} t），可以说取之不尽，用之不竭，所以开展以镁代铝的研究应给予重视。

3）传统材料在制备过程中污染严重。根据美国统计，含毒污水的排放，近80%来自化工、钢铁与有色金属冶炼，因而有机材料和金属是重要污染源。水泥生产除了高能耗外，也是一个重要污染源。1997年我国生产水泥5亿吨，占世界产量近40%，产生CO_2近4亿吨，SO_2达49万吨，NO_x为100万吨，粉尘1 100多万吨。这些废气加剧了温室效应，促进了酸雨。而炼1吨钢所产生的CO_2至少2倍于水泥。所以从大气污染看，材料的生产对环境的污染是十分严重的。

资源的大量消耗和环境污染都不利于材料的可持续发展。为此，必须注意以下几点。

1）开发节约资源、低污染的生产流程。当前正在开发一种所谓“零排放”流程，即所输入的原料及能源全部生成产品而没有废物排出，这当然是一种理想状态，但是至少应尽量开发充分利用资源而少污染的生产流程。

2）发展环境友好材料，或称环境材料（ecomaterials）。就是指与环境相适应的材料，即节约资源和能源、少污染、易回收或可降解的材料。

3）开发高性能、长寿命材料是节约资源、减少污染最有效的途径。因此，应把提高传统材料性能作为主要奋斗目标，如钢的强度大幅度提高；开发既抗腐蚀又可大幅度提高强度的水泥，以减少水泥用量，从而减少污染和节约资源和能源。

4）用新技术改造传统材料生产流程。一方面提高劳动生产率，改善产品质量，降低成本，另一方面使传统材料升级换代，扩大材料的用途，以增加竞争能力。

总之，传统材料仍有很多值得探索的科学技术问题。即使在工业发达国家，也都花很大力气在这方面开展研究与开发工作，否则便无竞争能力。根据我国1998年在《国家重点基础研究发展规划》中的新一代钢铁材料重大基础研究的结果表明，利用钢的晶粒细化、成分的纯净化和均匀化，在不增加任何元素的情况下，可使碳钢的强度从200 MPa提高到400 MPa，使低合金钢的强度从400 MPa提高到800 MPa，而且生产实践证明，在不明显改变现有生产条件下就可实现工业生产。

近年来提出生态设计（ecodesign）的概念，全面考虑材料对环境的影响，并实现有效利用资源，提高性能、减少污染，延长使用寿命，以期资源的使用效率达到目前的10倍（factor10）乃至20倍以上，这是人类的奋斗目标。

5 材料科学技术发展的重点

材料科学技术为开发先进材料、改进现有材料和合理使用材料服务。根据国际动向并结合国内具体情况，材料科学技术发展的重点有以下几个方面。

5.1 材料制备工艺与技术的开发

任何一种新材料从发现到应用于实际，必须经过适宜的制备工艺才能成为工程材料。高温超导自1986年发现以后到20世纪末，已有15年的历史，但仍不能普遍应用于电力工业，主要是因为没有找到价廉而稳定的生产线材的工艺。C_{60}也是如此，尽管在发现之初认为它的用途十分广泛，但到20世纪末仍处于科研阶段。传统材料也需要不断改进生产工艺或流程，以提高产品质量、降低成本和减少污染，从而提高竞争能力。

分子束外延技术的出现，可以控制薄膜的生长精确到几个原子的厚度，从而实现了“原子工程”或“能带工程”，为原子、分子设计提供了有效手段。快冷技术（即每秒冷却速度达10^4 ~ 10^8 K）的采用，为金属材料的发展开辟了一条新途径。首先是金属玻璃的形成，提高了金属强度、耐磨耐蚀性能和磁学性能。其次通过快冷可得到超细晶粒，成为改进性能的有效方法。最后是通过快冷发现了准晶，由此改变了晶体学的传统观念。所以材料制备方法的研究与开发成为材料科学技术的重点。

材料制备工艺的重点：一是工艺流程的智能化；二是实现原子或分子加工，使材料或器件依照人们的意志达到微型化、多功能化和智能化。

材料的现代制备工艺或技术往往与极端条件密切相联，如利用空间失重条件进行晶体生长，可得出无偏析或低偏析的材料；强磁场、强冲击波、超高压、超高真空以及强制快冷等都可能成为改进材料性能的有效手段。

5.2 材料的应用研究与开发

材料的广泛应用是材料科学技术发展的主要动力，实验研究出来的具有优异性能的材料不等于具有实用价值，必须通过大量应用研究，才能发挥其应有的作用。这里对材料的应用问题稍加讨论，如此对材料有一个比较全面的了解。

材料的应用要考虑以下几个因素：一是材料的使用性能（performance）；二是使用寿命（durability）及可靠性（reliability）；三是环境适应性（environmental compliance），包括生产过程与使用期间；四是价格（cost）。当然，不同材料及使用的对象不同，考虑的重点就不一样，有些量大面广的材料，价格低廉是主要的，因而生产要低成本，检验不十分复杂，如建材与包装材料；相反，有些关键技术所用关键材料，如航空航天及医用生物材料，一旦发生意外，则损失严重，因而必须高质量、安全可靠，加强检验，否则后果不堪设想，所以有时检验费用比材料本身花费还高。以航空发动机所用高温合金为例，作为涡轮叶片及涡轮盘材料，一旦在飞行过程中出现断裂，很可能造成机毁人亡，因而在要求长寿命（几万小时）的同时，对可靠性的要求特别严格。为了保证材料的质量，采用三次熔炼，即真空感应炉熔炼，以保证严格控制成分（去气、去有害杂质）；再用电渣重熔，以去除非金属夹杂物；最后真空自耗电弧重熔，可以得到无宏观缺陷的合金锭，如此保证材料质量的均一性和完整性，再经锻造，或重熔铸造加工成零件，最后经过高灵敏度的检验合格后再装机使用。对医用生物材料来说，质量保证更为严格，因为一旦因质量事故而产生不良后果，则后患无穷。

材料是否有竞争能力，除了质量和质量稳定性以外，还有材料的生产成本和价格。价格的影响因素很多，其中产量是决定因素之一。因为高产量不但可以实现自动化，而且质量稳定，成品率大幅度提高，成本明显下降。因此扩大应用范围是促进生产量提高的必由之路。

材料应用研究又是机械部件与电子元件失效分析的基础。失效分析的准确性与时效性代表一个国家的科学技术水平，因而通过材料的应用研究来培养一批既有实践经验又有理论基础的工程技术人才是一个重要途径。通过材料应用研究也可以发现材料中规律性的东西，进而指导材料的改进和发展。

5.3 开发先进材料，发展高技术产业

当今是高技术主宰着社会，一方面高技术促进社会的发展，又是国防安全的保证，另一方面高技术是传统产业改造和发展支柱产业不可或缺的组成部分，而先进材料又是高技术的先导和基础，所以开发先进材料必须受到高度重视。根据目前情况，先进材料的重点有以下几个方面。

1）信息功能材料将得到更高的重视 人类已进入信息时代，为了实现装置的小型化、低功耗、多功能化和智能化，信息功能材料将受到更高的重视。所谓信息功能材料，就是指信息的产生、获取、存储、传输、转换、处理、显示所需材料。主要用于计算机、通讯和控制，成为3C（computer，communication及control）所需材料。其特点是要求高、发展快、种类繁多，这里不可能一一列举。但总的来说，信息技术（IT，information technology）以集成电路为基础（细胞），计算机为中枢（大脑），网络为媒介（神经）。

现仅以集成电路所需材料为例说明。图1.1-3为集成电路发展的历程，根据Moore's定理，到2010年集成电路的特征尺寸为0.07 μm。由于制作困难，再加发热过高以及量子效应等可能达到极限，这时硅晶片的直径将在450 mm以上，对硅单晶质量要求进一步提高，如表1.1-3。因此硅单晶的

生长一直是研究的主题，因为晶体直径愈大，杂质的偏析愈严重，愈容易出现缺陷。在硅集成电路达到极限以后，下一代集成电路将如何发展，是目前正在探索中的问题：量子点、光集成，还是采用别的途径？

计算机作为控制装置的组成部分，敏感元件传感器是关键，因此敏感材料的灵敏度和稳定性决定计算机可能控制的精度。敏感材料是利用材料对物理、化学和生物的变化而产生不同反应发展出来的，种类繁多，发展很快。表1.1-4列举了几种无机非金属敏感材料及其原理，由此说明对物质的基础研究至关重要。

表1.1-3 1998年公布的硅片表面颗粒度指标

参数	年份					
	1997	1999	2002	2005	2008	2011
IC工艺/nm	250	180	130	100	70	50
颗粒大小/nm	125	90	65	50	35	25
颗粒密度/cm^{-2}	≤0.14	≤0.088	≤0.055	≤0.038	≤0.023	≤0.014

表1.1-4 无机非金属敏感材料

探测性能	原理	材料	探测性能	原理	材料
氧含量	体离子导电	$Zr_{1-x}CaO_{2-y}$	电压	晶界面隧道	$ZnO \cdot Bi_2O_3$
湿度	表面体离子导电	$MgCr_2O_4 \cdot TiO_2$	PTC热敏电阻	晶界面相变	$Ba_{1-x}Ce_xTiO_3$
压力	压电	$PbZr_{1-x}TiO_3$	化学	表面电子导电	$ZrO \cdot CuO$
酸度	表面化学反应	IrO_{2-x}	光学	光电阻	CdS

2) 先进结构材料的研究与开发是永恒的主题　结构材料一般数量大，资源与能源消耗高，污染严重，对可持续发展有决定作用。结构材料性能的提高，无疑会减缓上述压力，因性能的提高，使用寿命的延长，可减少材料的用量。特别对运载机械来说，工作温度愈高，效率也愈高，所以，高温材料的研究与开发是当务之急。高比强度、高比刚度，对提高力学性能十分重要；耐磨、耐蚀、抗疲劳、抗老化是延长使用寿命的关键，因此，在材料的发展中，必须给予高度重视。从材料类型来说，金属材料在21世纪上半叶仍将占主要地位。先进陶瓷材料要进一步提高韧性，降低成本。有机材料以其再生资源的优势和优异性能，而且可实现分子设计而进一步得到发展。先进复合材料类型很多，有些已得到广泛应用，如碳纤维增强树脂基复合材料；有的因价格高，制作难度大而需进一步发展新工艺，如金属基复合材料。有些超高温材料，如碳-碳复合材料，必须解决抗氧化问题才能得到更大范围的应用。

3) 能源材料的开发有广阔的前景　能源的利用是人类进步的标志之一。目前化石燃料一是储量有限，二是污染严重，必须大力开发无污染、可再生能源。太阳能虽然密度低，又受气候影响，但辐射于地表的能量，一万倍于人类开发的能源。因此，开发光-电转换效率高而又价廉长寿命的材料是当务之急。海水中氢的同位素氘可谓取之不尽，用之不竭，认为是人类的最终能源。科学家已采用多种方法（如等离子、激光）点火，实现可控热核聚变，用聚变能发电，预测21世纪内可实现商业化，其中抗辐射、耐高温、耐氦脆材料是关键之一，必须得到解决。除开发新能源以外，节能也十分重要，如超导的利用。低温超导以金属（NbTi合金、Nb_3Sn等）为主，但必须液氦冷却（4 K），成本太高而难以普及。20世纪80年代发现可在液氮温度下出现超导现象的氧化物超导体，称为高温超导，但制成符合工业应用的线材和块体还存在差距，不能普遍应用。利用超导输电，可以减少线路损失；超导储电，效率可提高；超导电机功率大、体积小、损耗低，十分诱人。目前正在探索室温超导，一旦有所突破，前途将更是无量。

4) 有机高分子材料将有更大发展　有机高分子材料以其可再生，资源丰富，又有优异性能，特别可实现分子设计，将会更大发展。高分子聚合物不但是重要的结构材料，而且正在发展成为重要的功能材料，从半导体到超导体都有发展前途。作为电导体，其电导率可与铜相比。高分子光学材料种类繁多，正在研究开发。高分子材料也具有铁磁性，这些都是21世纪要开展研究与开发的重要领域。但是高分子的某些缺点也是必须下大力气才能得到解决的，如稳定性、抗老化性及阻燃性。

5) 生物材料将受到更大的重视　随着生物技术的发展及人类寿命的延长与生活质量的提高，生物医用材料已成为人们非常关注的领域，人的器官更换、药物缓释及组织工程的发展将逐步深入。生物模拟是另一正在兴起的学科，使材料的功能进一步提高，并能达到自恢复、自修复或智能化。生物材料的更长远目标是使生物技术原理用于工业化生产，改变高温、高压及耗能高的生产方式，催化剂已迈出第一步。光合作用使水和CO_2合成碳水化合物，人类正在寻求利用生物技术，通过像催化剂的效应以工厂方式合成粮食，如此除了可以解决世界粮食问题，CO_2过剩问题也可得到解决。当然，这是很长的历史过程。

6) 纳米材料及制备技术的研究与开发迫在眉睫　当物质到纳米尺度时，由于其尺寸效应、晶界效应和量子效应等，材料显示出奇特的物理、化学性能，或其生物功能有明显改变，利用这一效应可大幅度提高结构材料的强度，改善其脆性。对功能材料来说，应用方面更广，高效率的催化作用，以及多变的光学性质、电学性质、磁学性质和热学性质等，将使目前的功能材料性能明显提高。这些改变将影响到工业和国防的各个方面，所以有人认为纳米技术将如信息技术或生物技术一样，将导致下一代工业革命。应该指出，纳米科学技术仍处于基础研究阶段，特别是纳米电子学、纳米医学所需材料，尚处于探索阶段，但是纳米技术用于结构材料的改性及用作某些功能材料已显示出明显的优越性，有些已进入产业化。

5.4 材料设计

材料通过设计而得到所需要的性能是材料科学技术工作者的奋斗目标。材料的研制从以经验为主过渡到以科学设计为主，从而进入“设计材料”（materials by design）时代。材

料设计的概念始于20世纪80年代初，随着有关学科的发展，设计材料愈来愈受到重视。材料设计可以分为几个层次，最基础的是量子设计（quantum design）。这是由于电子运动而引起的多种现象，如光学，这是功能材料的基础。其次为原子设计（atomic design），这是纳米技术的基础。原子排列是决定材料力学性质与化学性质的基础，原子操纵可能是下一代芯片的一种技术。第三个层次是微观设计（micro-design），即微米级结构的设计，金属的相变、晶界的控制都属于这一范畴。第四个层次是宏观设计（macro-design），以毫米到厘米为对象，像金属在凝固过程的结构与偏析便属于此。

材料设计的最终目标是根据最终需求，设计出合理成分，制订最佳生产流程，而后生产出符合要求的材料。应该指出，材料设计十分复杂，如模型的建立往往是基于平衡态，而实际材料多处于非平衡态，如凝固过程的偏析和相变等。材料的力学性质往往对结构十分敏感，因此，结构的任何微小变化，性能都会发生明显变化，何况有些性质，如脆性、裂纹的萌生与扩展等还很不清楚，因此，要想得到确切真实的结果决非易事。另外，相图也是材料设计不可或缺的组成部分。

可以肯定，随着科学技术的进步，各学科的发展与对材料的了解日益深入，计算机的巨型化和智能化，一个更加科学化发展新材料的时期将会到来。要做到这一点，最关键的是建立完善的数据库与知识库，发展符合实际的物理模型，这就要求应用数学家、物理学家、化学家、材料科学家与工程技术等方面人员的密切合作。

5.5 科学仪器与检测装置

科学技术的发展很大程度上依赖于新科学仪器的不断发明和性能的不断提高，以物理学诺贝尔奖金获得者为例，大约有半数得益于新仪器的发明或新测试手段的采用。工业产品质量的改进往往取决于检测装置精度的提高。所以工业发达国家对科学仪器与检测装置的开发与制造都给以足够的重视。以电子光学仪器为例，1863年金属材料研究首先应用光学显微镜。几十年后出现了电子显微镜、扫描电镜（SEM）、高分辨电镜（HREM），其分辨率达0.2 nm，足以观察到原子。后又有扫描透射电镜（STEM），不但可以观察到原子，而且可以分析微小区域的化学组成和结构。20世纪80年代初出现的扫描隧道显微镜（STM），可以在非真空条件下观察到原子，并用来进行原子加工。一种更为引人注意的显微技术称之为扫描探针显微镜（scanning probe microscope），简易价廉，它基于压电原理，可以在不同结构中观察到单个原子或分子。原子力显微镜（AFM）可以观察到金属脆性或韧性断裂过程。红外原子力显微镜（infrared AFM）可以观察高分子聚合物的苯环。这些分析装置与高功能计算机相结合，可以得出结构细节，使结构分析从以经验为主，费时而不准确，过渡到直接、快速地解决材料中的问题。

除显微结构以外，材料的晶体结构分析开始于劳厄（Laue）等，于1912年发现X射线通过晶体产生衍射花样，随后又推广到电子衍射与中子衍射，成为探明材料原子结构的重要手段，而后又出现了基于不同原理的多种谱仪，成为分析表征材料结构与性能的有力装置。材料科学工作者不仅利用现有先进设备研究和开发材料，同时也应致力于发明与创造新的检验与测试装置，因为任何已商业化的设备至少落后了5年，突破性的工作很难在这些已商品化的设备上做出来。

检测是控制工艺流程及产品质量的主要手段，如无损检测装置不但可以检查宏观缺陷，也可监控裂纹的萌生与扩展，提高无损检测装置的精度十分重要，如先进陶瓷的临界裂纹，只有几十微米，没有精密的无损探伤仪，陶瓷的质量控制便谈不上。当前检测装置除了扩大品种，提高灵敏度和稳定性以外，主要发展在线、动态以及适应各种恶劣环境的检测装置。

总之，材料的重要性任何时候都不会下降，但是随着科学技术的进步，材料与器件或构件的结合愈来愈密切，全新的人造材料或人工合成材料愈来愈占更主要的位置，具有创造性的新材料将层出不穷。但是，在很多情况下，新材料的出现往往难以预料，如高温超导材料和C_{60}等。因此，材料科学工作者必须富有探索精神，同时也要为他们建立一个宽松环境的氛围，在打好本学科基础上，重视交叉学科，广泛接触实际，抓住任何一个“灵感”火花，才会做出创新成果。

编写：师昌绪（国家自然科学基金委员会）

第 2 章　材料科学基础

材料科学是研究材料的组织、结构与性质关系的科学，又是一门交叉学科，同时也是材料工程学的基础。其基本原理植根于凝聚态物理学、物理化学与合成化学等学科。作为“中国材料工程大典”系列丛书的先导，本章尝试从不同的侧面非常简略地介绍关于材料的结构与缺陷，结构转变和重要物性的基本概念，以服务于其后各卷的专门论述。本章的体系和内容取舍较多地参考了冯端，师昌绪，刘治国主编的“材料科学导论-融贯的论述”一书，甚至还采用了其中的一些插图，在此向有关作者先行致谢。

1　材料的结构

材料的结构是材料科学与工程的核心问题，也是理解材料的性质和掌握材料制备方法的基础。可以从几个不同的侧面来描述材料的结构，例如从化学的侧面来理解原子的构造和原子之间的键合，从晶体学的侧面来理解对称性在决定材料结构方面所起的关键作用，从晶体缺陷和微结构的侧面来理解实际材料丰富多彩的行为等。

1.1　原子的键合

组成材料的基本单元是原子，原子通过电磁交互作用结合在一起，结合的方式称为化学键。按其作用机制化学键可分为四类。实际材料中的键往往并不是典型的某一类键，而可能同时具有不同类键的特征。

1.1.1　原子结构简介

19 世纪末门捷列夫就已发现元素及其化合物的性质随着元素原子量的增加而呈周期性的变化并提出了著名的元素周期律。这为人们理解原子的结构提供了重要的线索。标明各元素原子及其单质主要特性的元素周期表见图 1.2-1。

20 世纪早期量子力学的建立则奠定了描述原子结构的理论基础。原子具有核式结构，由带正电荷的原子核和在分立的轨道上绕原子核旋转的电子组成。依据量子力学原理，物质同时具有粒子性和波动性，人们传统认为是粒子的电子同时也具有波动性。这种与粒子相联系的波称为德布罗意波，描述具有波动性的粒子的运动方程为薛定谔方程。在这个方程中粒子的各物理量都用相应的算符表示并依条件的不同只能取一系列分立的数值，即“量子化”。氢原子只有一个电子，处于原子核形成的球形库仑场中，可以用薛定谔方程求出其严格解。电子的状态用四个量子数来描述：主量子数 $n=0,1,2,3\cdots\cdots$，描述电子的轨道能量；角量子数 $l=0,1,2\cdots\cdots,n-1$，描述电子的角动量；磁量子数 $m=l,l-1,\cdots\cdots,-l$，描述电子的角动量在 z 轴方向的分量；自旋量子数 $m_s=1/2$ 或 $-1/2$，描述电子的自旋状态。其中前三个量子数完全确定了电子的空间运动。

氢以外的元素的原子有一个以上电子，数目与其原子序数相同，其每个电子不仅受到原子核库仑场的作用，还受到其他电子库仑场的作用，其理论计算成为复杂的量子力学多体问题。这时只能求数值解或用单电子近似求近似解。人们将多电子原子中的电子态仍用前述四个量子数来描述。通常人们将角量子数 $l=0,1,2,3,\cdots\cdots$ 记为 s，p，d，f，……，如 $n=3$，$l=2$ 的能级便记为 3d。主量子数 n 和角量子数 l 基本上确定了电子的能级。n 一定时 l 越大能级越高，l 一定时 n 越大能级越高。如果比较 n 和 l 都不同的能级，则情况比较复杂，原子能级出现交错现象。图 1.2-2 为泡利 (Pauling) 总结理论计算和实验结果提出的多电子原子能级示意图，电子由低能级到高能级一个一个填上去便形成了原子的壳层结构。更详细的研究表明电子的排布要遵循以下三原则：同一电子态只能由一个电子占据（泡利不相容原理）；电子优先占据能量较低的能级（能量最低原理）；未填满的壳层中电子自旋值应尽可能大（洪德规则）。

1.1.2　离子键

离子键的基本特征是成键双方分别以正，负离子的形式通过库仑作用结合在一起。通过离子键结合而成的晶体称为离子晶体，其中正负离子相间排列，最典型的离子晶体为碱金属和卤族元素形成的化合物如氯化钠。由于正负离子都是满壳层结构而具有球对称性，离子键没有明显的方向性。原子失去一个外层电子即成为正离子，这个过程称为电离，所需能量称为电离能。电离能随原子序数作周期性变化。碱金属因其外壳层只有一个电子而电离能最小。原子也可吸收一个外电子并放出能量而变为负离子，所放出的能量称为亲合能，它也随原子序数作周期性变化。卤族元素吸收一个外电子后其外壳层电子可以变满，因而其亲和能最大。以氯化钠分子的形成为例，钠原子失去一个电子变成钠离子同时吸收电离能 5.14 eV，氯原子得到一个电子变成氯离子同时放出亲合能 3.61 eV，钠离子和氯离子通过库仑作用结合为氯化钠分子同时放出内聚能 7.9 eV。在整个反应中放出的总能量为 3.16 eV + 7.9 eV − 5.14 eV = 5.92 eV。该过程的示意见图 1.2-3。离子键是一种强键，每个分子的结合能在 4 ~ 8 eV 之间。当然，孤立的氯化钠分子在晶体中并不存在，其结合能的计算方法可参见固体物理学教课书关于马德隆常数计算的章节。

1.1.3　共价键

共价键的特征是自旋相反的外层电子间的交换作用使成键的两原子之间的区域出现较高的负电荷密度，它与离子的静电吸引力将两原子结合在一起，这也常被形象地称为电子的共有化。共价键的典型例子是氢分子。Heitler 和 London 基于量子力学计算提出了氯分子的共价键理论。图 1.2-4 是他们的结果的示意。该图显示两个电子自旋平行，反平行时氢分子的能量与两个氢原子核之间距离的关系，作为对比，经典力学计算的结果也标在图中。显然，两电子自旋反平行时氢分子能量随两原子接近逐步下降直至一个极小值，这种状态称为成键态，这时两原子间的距离对应于分子中两原子的平衡间距。若两电子自旋平行，氢分子能量则随两原子接近单调上升，此时两原子不成键，也不能结合成分子。

共价键的基本特征是其饱和性和方向性。“饱和性”指一个原子只能形成一定数目的共价键，也就是说只能和一定数目的最近邻原子结合。其原因在于每个原子通过与相邻各原子共享价电子的方式使其自身具有满壳层的稳定结构。Ⅳ族的 C，Si，Ge 具有最典型的共价键，它们的每个原子有四个价电子，通过形成金刚石结构与四个同类原子为邻，每个原子与相邻的四个原子共享价电子使其周围有八个电子形成满壳层的稳定结构。“方向性”指原子只在一些特定的方向上形成共价键。这是因为电子轨道有其方向性，相邻原子只能在这些方向上成键。共价键是一种强键，每个原子的结合能在 1 ~ 5 eV 之间，典型的共价晶体金刚石是自然界最硬的材料。Ⅳ族元素，以共价键结合的Ⅲ − Ⅴ族化合物及Ⅱ − Ⅵ族化合物是主要的半导体材料。不同元素的原子结合时其键合可同时具有共价键和离子键的特征。

元素周期表

图例：30 65.38 1180 692.73 锌 Zn 7.14 [Ar]3d¹⁰4s² Zinc —— 原子序数、原子量、沸点(K)、熔点(K)、符号、300K密度 (g/cm³)、电子组态、英文名

原子序数	符号	中文名	英文名	原子量	沸点(K)	熔点(K)	300K密度(g/cm³)	电子组态	族
1	H	氢	Hydrogen	1.0079	20.268	14.025	0.0899	$1s^1$	IA
2	He	氦	Helium	4.00260	4.215	0.95	0.1787	$1s^2$	0
3	Li	锂	Lithium	6.941	1615	453.7	0.53	$1s^22s^1$	IA
4	Be	铍	Beryllium	9.01218	2745	1560	1.85	$1s^22s^2$	IIA
5	B	硼	Boron	10.81	4275	2300	2.34	$1s^22s^2p^1$	IIIA
6	C	碳	Carbon	12.011	4728	3652	2.62	$1s^22s^2p^2$	IVA
7	N	氮	Nitrogen	14.0067	77.35	63.14	1.251	$1s^22s^2p^3$	VA
8	O	氧	Oxygen	15.9994	90.18	50.35	1.429	$1s^22s^2p^4$	VIA
9	F	氟	Fluorine	18.998403	84.95	53.48	1.696	$1s^22s^2p^5$	VIIA
10	Ne	氖	Neon	20.179	27.096	24.553	0.901	$1s^22s^2p^6$	0
11	Na	钠	Sodium	22.98977	1156	371.0	0.97	$[Ne]3s^1$	IA
12	Mg	镁	Magnesium	24.305	1363	922	1.74	$[Ne]3s^2$	IIA
13	Al	铝	Aluminum	26.98154	2793	933.25	2.70	$[Ne]3s^2p^1$	IIIA
14	Si	硅	Silicon	28.0855	3540	1685	2.33	$[Ne]3s^2p^2$	IVA
15	P	磷	Phosphorus	30.97376	550	317.30	1.82	$[Ne]3s^2p^3$	VA
16	S	硫	Sulfur	32.06	717.75	388.36	2.07	$[Ne]3s^2p^4$	VIA
17	Cl	氯	Chlorine	35.453	239.1	172.16	3.17	$[Ne]3s^2p^5$	VIIA
18	Ar	氩	Argon	39.948	87.30	83.81	1.784	$[Ne]3s^2p^6$	0
19	K	钾	Potassium	39.0983	1032	336.35	0.86	$[Ar]4s^1$	IA
20	Ca	钙	Calcium	40.08	1757	1112	1.55	$[Ar]4s^2$	IIA
21	Sc	钪	Scandium	44.9559	3104	1812	3.0	$[Ar]3d^14s^2$	IIIB
22	Ti	钛	Titanium	47.90	3562	1943	4.50	$[Ar]3d^24s^2$	IVB
23	V	钒	Vanadium	50.9415	3682	2175	5.8	$[Ar]3d^34s^2$	VB
24	Cr	铬	Chromium	51.996	2945	2130	7.19	$[Ar]3d^54s^1$	VIB
25	Mn	锰	Manganese	54.9380	2335	1517	7.43	$[Ar]3d^54s^2$	VIIB
26	Fe	铁	Iron	55.847	3135	1809	7.86	$[Ar]3d^64s^2$	VIII
27	Co	钴	Cobalt	58.9332	3201	1768	8.90	$[Ar]3d^74s^2$	VIII
28	Ni	镍	Nickel	58.70	3187	1726	8.90	$[Ar]3d^84s^2$	VIII
29	Cu	铜	Copper	63.546	2836	1357.6	8.96	$[Ar]3d^{10}4s^1$	IB
30	Zn	锌	Zinc	65.38	1180	692.73	7.14	$[Ar]3d^{10}4s^2$	IIB
31	Ga	镓	Gallium	69.72	2478	302.90	5.91	$[Ar]3d^{10}4s^2p^1$	IIIA
32	Ge	锗	Germanium	72.59	3107	1210.4	5.32	$[Ar]3d^{10}4s^2p^2$	IVA
33	As	砷	Arsenic	74.9216	876	1081	5.72	$[Ar]3d^{10}4s^2p^3$	VA
34	Se	硒	Selenium	78.96	958	494	4.80	$[Ar]3d^{10}4s^2p^4$	VIA
35	Br	溴	Bromine	79.904	332.25	265.90	3.12	$[Ar]3d^{10}4s^2p^5$	VIIA
36	Kr	氪	Krypton	83.80	119.80	115.78	3.74	$[Ar]3d^{10}4s^2p^6$	0
37	Rb	铷	Rubidium	85.4678	961	312.64	1.53	$[Kr]5s^1$	IA
38	Sr	锶	Strontium	87.62	1650	1041	2.6	$[Kr]5s^2$	IIA
39	Y	钇	Yttrium	88.9059	3611	1799	4.5	$[Kr]4d^15s^2$	IIIB
40	Zr	锆	Zirconium	91.22	4682	2125	6.49	$[Kr]4d^25s^2$	IVB
41	Nb	铌	Niobium	92.9064	5017	2740	8.55	$[Kr]4d^45s^1$	VB
42	Mo	钼	Molybdenum	95.94	4912	2890	10.2	$[Kr]4d^55s^1$	VIB
43	Tc	锝	Technetium	(98)	4538	2473	11.5	$[Kr]4d^55s^2$	VIIB
44	Ru	钌	Ruthenium	101.07	4423	2523	12.2	$[Kr]4d^75s^1$	VIII
45	Rh	铑	Rhodium	102.9055	3970	2236	12.4	$[Kr]4d^85s^1$	VIII
46	Pd	钯	Palladium	106.4	3237	1825	12.0	$[Kr]4d^{10}$	VIII
47	Ag	银	Silver	107.868	2436	1234	10.5	$[Kr]4d^{10}5s^1$	IB
48	Cd	镉	Cadmium	112.41	1040	594.18	8.65	$[Kr]4d^{10}5s^2$	IIB
49	In	铟	Indium	114.82	2346	429.76	7.31	$[Kr]4d^{10}5s^2p^1$	IIIA
50	Sn	锡	Tin	118.69	2876	505.06	7.30	$[Kr]4d^{10}5s^2p^2$	IVA
51	Sb	锑	Antimony	121.75	1860	904	6.68	$[Kr]4d^{10}5s^2p^3$	VA
52	Te	碲	Tellurium	127.60	1261	722.65	6.24	$[Kr]4d^{10}5s^2p^4$	VIA
53	I	碘	Iodine	126.9045	458.4	386.7	4.92	$[Kr]4d^{10}5s^2p^5$	VIIA
54	Xe	氙	Xenon	131.30	165.03	161.36	5.89	$[Kr]4d^{10}5s^2p^6$	0
55	Cs	铯	Cesium	132.9054	944	301.55	1.87	$[Xe]6s^1$	IA
56	Ba	钡	Barium	137.33	2171	1002	3.5	$[Xe]6s^2$	IIA
57	La	镧★	Lanthanum	138.9055	3730	1193	6.7	$[Xe]5d^16s^2$	IIIB
72	Hf	铪	Hafnium	178.49	4876	2500	13.1	$[Xe]4f^{14}5d^26s^2$	IVB
73	Ta	钽	Tantalum	180.9479	5731	3287	16.6	$[Xe]4f^{14}5d^36s^2$	VB
74	W	钨	Tungsten	183.85	5828	3680	19.3	$[Xe]4f^{14}5d^46s^2$	VIB
75	Re	铼	Rhenium	186.207	5869	3453	21.0	$[Xe]4f^{14}5d^56s^2$	VIIB
76	Os	锇	Osmium	190.2	5285	3300	22.4	$[Xe]4f^{14}5d^66s^2$	VIII
77	Ir	铱	Iridium	192.22	4701	2716	22.5	$[Xe]4f^{14}5d^76s^2$	VIII
78	Pt	铂	Platinum	195.09	4100	2045	21.4	$[Xe]4f^{14}5d^96s^1$	VIII
79	Au	金	Gold	196.9665	3130	1337.58	19.3	$[Xe]4f^{14}5d^{10}6s^1$	IB
80	Hg	汞	Mercury	200.59	630	234.28	13.53	$[Xe]4f^{14}5d^{10}6s^2$	IIB
81	Tl	铊	Thallium	204.37	1746	577	11.85	$[Xe]4f^{14}5d^{10}6s^2p^1$	IIIA
82	Pb	铅	Lead	207.2	2023	600.6	11.4	$[Xe]4f^{14}5d^{10}6s^2p^2$	IVA
83	Bi	铋	Bismuth	208.9804	1837	544.52	9.8	$[Xe]4f^{14}5d^{10}6s^2p^3$	VA
84	Po	钋	Polonium	(209)	1235	527	9.4	$[Xe]4f^{14}5d^{10}6s^2p^4$	VIA
85	At	砹	Astatine	(210)	610	575	-	$[Xe]4f^{14}5d^{10}6s^2p^5$	VIIA
86	Rn	氡	Radon	(222)	211	202	9.91	$[Xe]4f^{14}5d^{10}6s^2p^6$	0
87	Fr	钫	Francium	(223)	950	300	-	$[Rn]7s^1$	IA
88	Ra	镭	Radium	226.0254	1809	973	5	$[Rn]7s^2$	IIA
89	Ac	锕	Actinium	227.0278	3473	1323	10.07	$[Rn]6d^17s^2$	IIIB
90	Th	钍★★	Thorium	232.0381	5061	2028	11.7	$[Rn]6d^27s^2$	IVB

★

原子序数	符号	中文名	英文名	原子量	沸点(K)	熔点(K)	300K密度(g/cm³)	电子组态
58	Ce	铈	Cerium	140.12	3699	1071	6.78	$[Xe]4f^26s^2$
59	Pr	镨	Praseodymium	140.9077	3785	1204	6.77	$[Xe]4f^36s^2$
60	Nd	钕	Neodymium	144.24	3341	1289	7.00	$[Xe]4f^46s^2$
61	Pm	钷	Promethium	(145)	3785	1204	6.475	$[Xe]4f^56s^2$
62	Sm	钐	Samarium	150.4	2064	1345	7.54	$[Xe]4f^66s^2$
63	Eu	铕	Europium	151.96	1870	1090	5.26	$[Xe]4f^76s^2$
64	Gd	钆	Gadolinium	157.25	3539	1585	7.89	$[Xe]4f^75d^16s^2$
65	Tb	铽	Terbium	158.9254	3496	1630	8.27	$[Xe]4f^96s^2$
66	Dy	镝	Dysprosium	162.50	2835	1682	8.54	$[Xe]4f^{10}6s^2$
67	Ho	钬	Holmium	164.9304	2968	1743	8.80	$[Xe]4f^{11}6s^2$
68	Er	铒	Erbium	167.26	3136	1795	9.05	$[Xe]4f^{12}6s^2$
69	Tm	铥	Thulium	168.9342	2220	1818	9.33	$[Xe]4f^{13}6s^2$
70	Yb	镱	Ytterbium	173.04	1467	1097	6.98	$[Xe]4f^{14}6s^2$
71	Lu	镥	Lutetium	174.967	3668	1936	9.84	$[Xe]4f^{14}5d^16s^2$

★★

原子序数	符号	中文名	英文名	原子量	沸点(K)	熔点(K)	300K密度(g/cm³)	电子组态
91	Pa	镤	Protactinium	231.0359	-	-	15.4	$[Rn]5f^26d^17s^2$
92	U	铀	Uranium	238.029	4407	1405	18.90	$[Rn]5f^36d^17s^2$
93	Np	镎	Neptunium	237.0482	-	910	20.4	$[Rn]5f^46d^17s^2$
94	Pu	钚	Plutonium	(244)	3503	913	19.8	$[Rn]5f^67s^2$
95	Am	镅	Americium	(243)	2880	1268	13.6	$[Rn]5f^77s^2$
96	Cm	锔	Curium	(247)	-	1340	13.511	$[Rn]5f^76d^17s^2$
97	Bk	锫	Berkelium	(247)	-	-	-	$[Rn]5f^97s^2$
98	Cf	锎	Californium	(251)	-	900	-	$[Rn]5f^{10}7s^2$
99	Es	锿	Einsteinium	(252)	-	-	-	$[Rn]5f^{11}7s^2$
100	Fm	镄	Fermium	(257)	-	-	-	$[Rn]5f^{12}7s^2$
101	Md	钔	Mendelevium	(258)	-	-	-	$[Rn]5f^{13}7s^2$
102	No	锘	Nobelium	(259)	-	-	-	$[Rn]5f^{14}7s^2$
103	Lr	铹	Lawrencium	(260)	-	-	-	$[Rn]5f^{14}6d^17s^2$

注：物态与导电态一般指室温常压下的物相
其中砷的熔点为28个大气压下的数据，无沸点，876K为升华温度

图 1.2-1 元素周期表

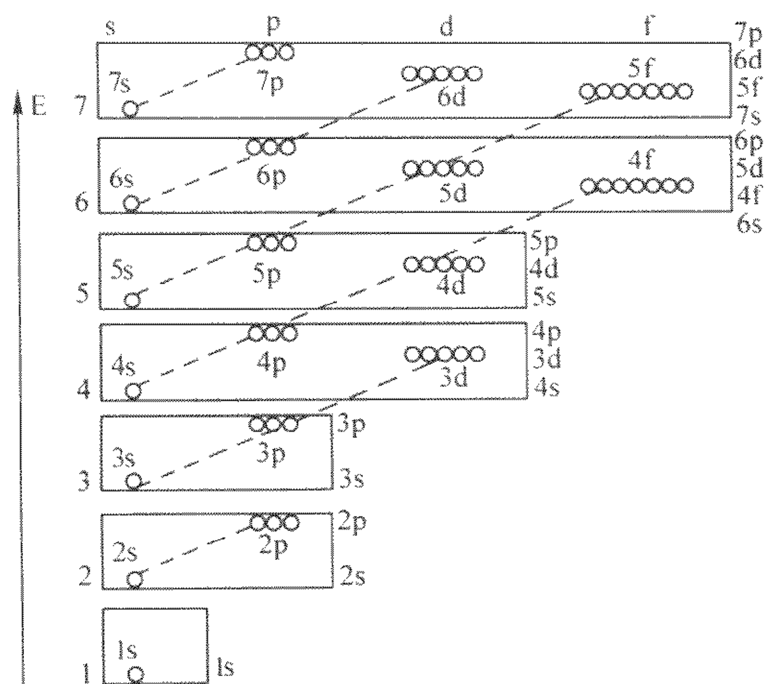


图 1.2-2　泡利的多电子原子能级示意图

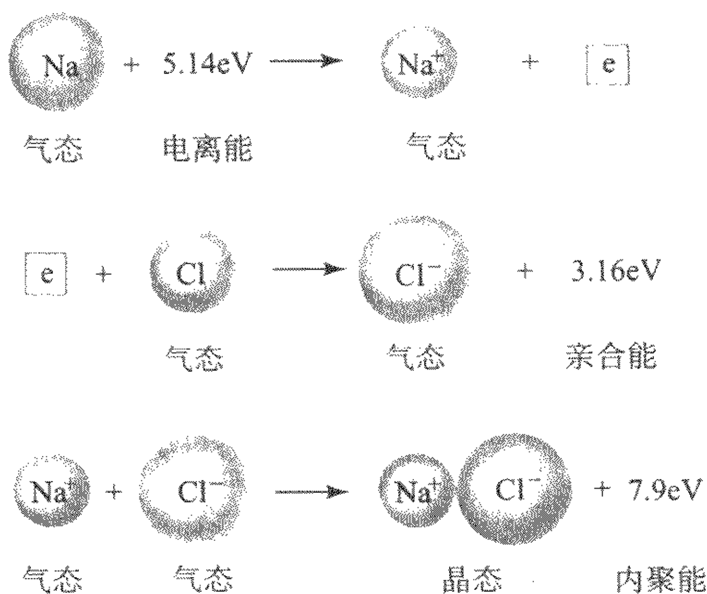


图 1.2-3　氯化钠分子形成及能量变化的示意

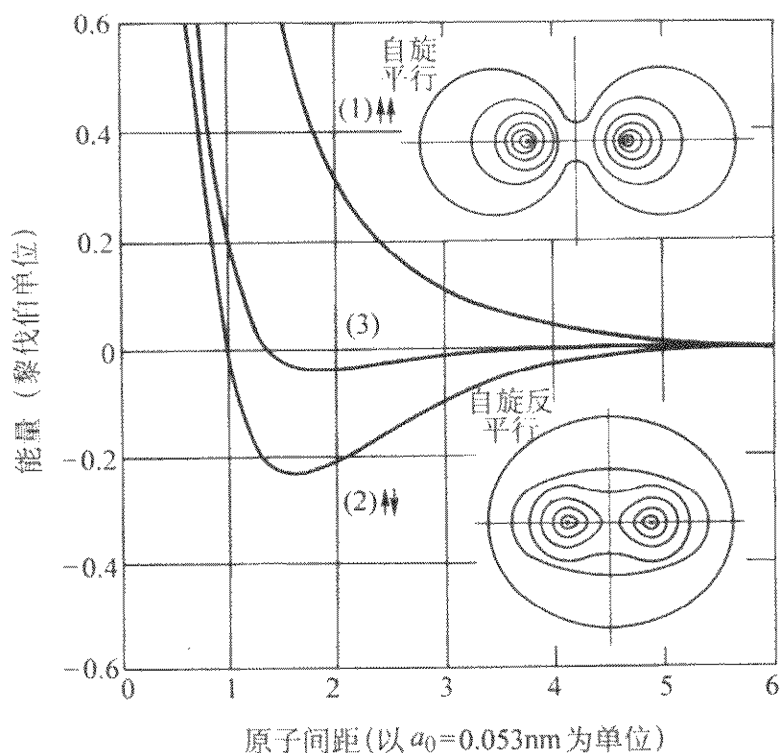


图 1.2-4　氢分子的结合能与两氢原子间距离的关系

（1）两电子自旋平行；（2）两电子自旋反平行；

（3）经典力学计算的结果

1.1.4　金属键

如果说共价键的特征是成键的两个原子共有一对价电子，那么金属键的特征则是整块材料中全部价电子的彻底共有化。原属于各金属原子的价电子在原子结合成金属晶体后可在整个晶体中运动，像云雾一样充斥于金属离子排列成的晶格之间。金属的许多特性，如良导电性，良导热性，金属光泽等都和价电子的自由运动相联系。价电子的共有化使单个原子价电子的能级演变成金属晶体的电子能带，自由电子从能带底部开始填起，直达费米能级。使金属键合在一起的吸引力主要来自电子云与金属离子的库仑作用。金属键没有明显的方向性和饱和性，通常金属具有配位数较高的面心立方、密积六方或体心立方等较密集的结构。由于金属键没有明显的方向性，金属多具有良好的范性，因为局部原子的无序排列是范性变形的必要条件。金属键也是一种强键，每个原子的结合能在 1 ~ 4 eV 之间。

1.1.5　弱键

前面介绍的三种键属于强键，它们都是原子之间的键。分子之间的键要弱得多，它们多源于偶极子之间的互作用，常称为弱键。

惰性元素的原子具有满的电子壳层，电子分布具有球对称性，因而没有永久电偶极矩。然而由于电子在作不断的运动，在许多瞬间其电荷分布偏离球对称，形成惰性元素原子的瞬间电偶极矩。这些瞬间电偶极子之间的互作用称为范德瓦尔斯吸引力。这种吸引力与两原子电子壳层间的斥力平衡则决定了两原子间的平衡距离。由范德瓦尔斯吸引力结合的两惰性元素原子间的相互作用能 U 与原子间距 r 的关系可写成

$$U(r) = -A/r^6 + B/r^{12} \qquad (1.2\text{-}1)$$

式中，前一项表达范德瓦尔斯吸引力，后一项表达两原子电子壳层间的斥力，A 和 B 是经验常数，恒为正数。该式可以满意地拟和惰性元素的试验数据。式（1.2-1）被称为 Lennard-Jones 势，常被用来描述非极性分子之间的交互作用。惰性元素氖晶体的结合能为每原子 0.02 eV，其融点为 24 K，这表明范德瓦尔斯互作用确实是很弱的。

在极性分子之间除了范德瓦尔斯互作用外更要考虑永久电偶极矩之间的互作用。特别重要和值得注意的是分子之间的氢键。以水为例，水分子中氧和氢之间的键基本上是共价键，但由于氢核的引力较弱，形成共价键的电子对显著偏向氧原子一方。同时因为水分子中两氢原子与氧原子的连线之间的夹角不是 180°，水分子具有强极性，氧原子一方带正电荷，氢原子一方带负电荷。这时水里一个水分子中的氧便会与相邻水分子中的氢形成氢键，如图 1.2-5 中的虚线所示。氢键是弱键中的较强者，但其互作用能量要较共价键小一个数量级。氢键可使一些分子液体中的分子结成断续的网络从而影响液体的性质，更重要的是氢键在聚合物和生物大分子的结合中起着关键的作用，因而越来越受到人们的关注。

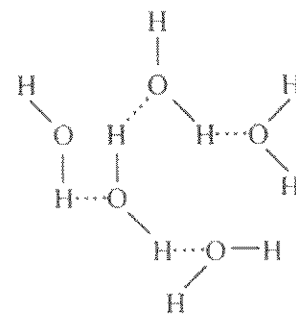


图 1.2-5　水分子之间氢键的示意

1.2　晶态、非晶态及其他

材料中原子或分子的排列方式（也称为结构）是多种多样的。许多材料具有晶态结构，其特点是原子排列具有周期性。非晶态材料则不然，它们不具有长程周期性。1984 年 Shechtman 等发现了准晶，这是一类具有 20 面体对称性并与传统周期结构不相容的新有序相。液晶则是一类其分子取向有序而位置无序的液体，兼有部分晶体和液体的性质。在本节我们将逐一进行介绍。

1.2.1　点阵与晶系

晶态的基本特征是其原子或原子团的排列具有三维周期性。组成这种三维周期结构的最小结构单元称为基元，基元

在三维空间作周期性排列便构成了晶体。在每个基元上选择一个环境相同的点，这些点的集合称为空间点阵或简称点阵。显然，点阵只是描述晶体中原子或原子团排列的一种几何抽象，它反映了晶体的对称性，并和基元一道反映了晶体的全部结构信息。为了画图的方便，以二维晶体为例，其示意见图 1.2-6。

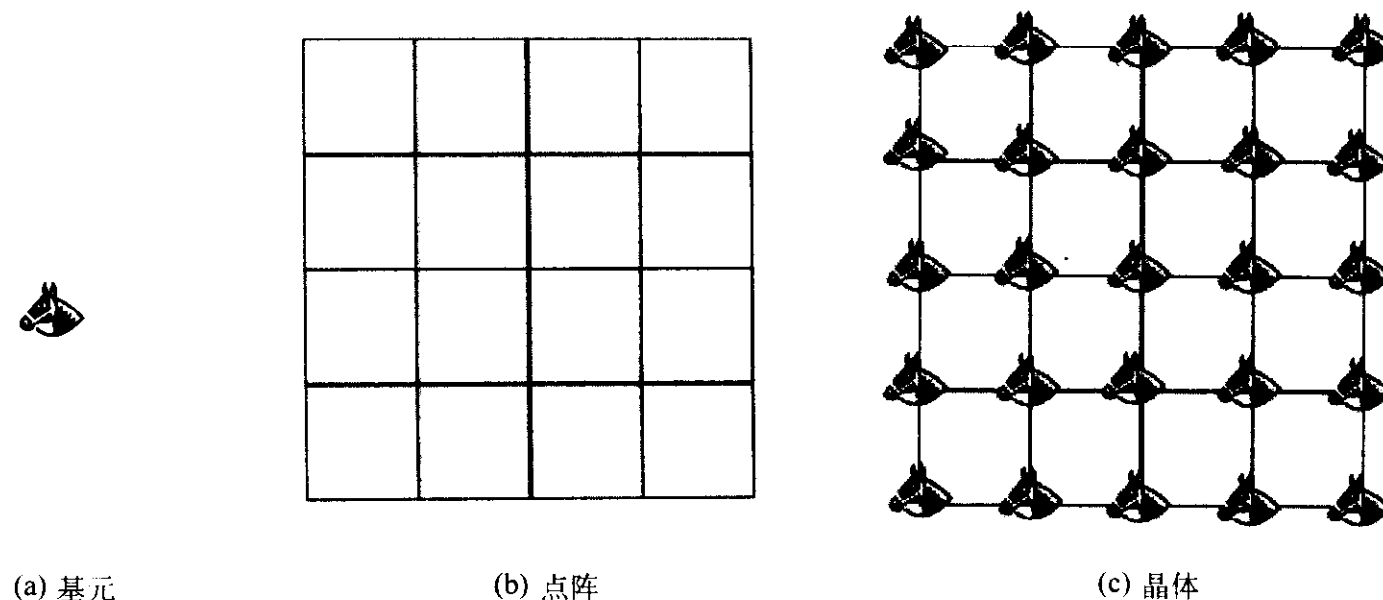


图 1.2-6　置于点阵上的基元构成晶体的示意

二维点阵可以用两个非共线的最短平移矢量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 描述，它们称为基矢。以任一格点为原点，任何其他格点的位置可以用

$$t = n_1\boldsymbol{a} + n_2\boldsymbol{b} \qquad (1.2\text{-}2)$$

表示，其中 n_1，n_2 为整数。两个基矢所确定的平行四边形称为晶胞。基矢和晶胞的选取方式是多种多样的，最小的这类平行四边形称为初级晶胞。有时人们更喜欢选择能反映晶体对称性的结构单元，称为非初级晶胞。以上的描述完全可以推广到三维，只要在式（1.2-2）中加上第三个基矢便可以了。三维面心立方和体心立方点阵中初级晶胞和反映其立方对称性的非初级晶胞（简称晶胞）选取方式的图示见图 1.2-7。

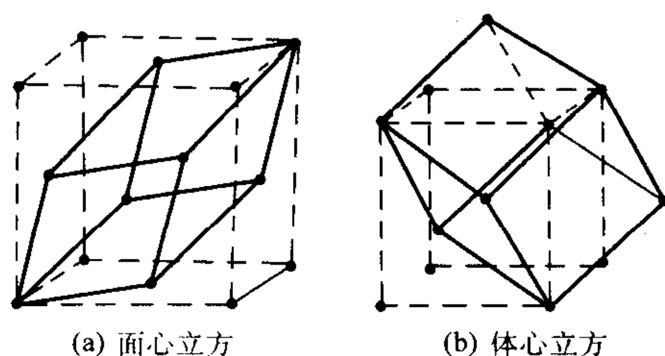


图 1.2-7　两种立方点阵中初级晶胞和反映其立方对称性的非初级晶胞（简称晶胞）的选取

点阵可按其对称性进行分类，这里所说的对称性除平移对称性外还包括其他对称性如旋转，反映，反演等。为了和平移对称性相容，这些对称性的具体种类受到严格限制。以旋转对称性为例，周期结构中只能存在 $n = 1, 2, 3, 4, 6$ 五种旋转轴，不能存在 5 次及 6 次以上旋转轴。这些可能对称性的组合决定了三维点阵只能有 14 种，称作 14 种布喇菲（Bravais）格子，分属于七个晶系。任何晶体对应的晶格都是这些布喇菲格子中的一种。布喇菲格子概括了它所代表晶格的对称性。晶系和相关布喇菲格子的图示见表 1.2-1。

布喇菲格子的格点分列在一系列相互平行的直线或平面上，这些直线或平面称为晶列或晶面。晶列的方向称为晶向。如果从某格点沿晶向到最近格点的位矢为 $h\boldsymbol{a} + i\boldsymbol{b} + j\boldsymbol{c}$，这里 h，i，j 均为正数，则该晶向便被记为 $[hij]$，它称为晶向指数。如果因晶体对称性存在多种晶列，其上格点排列完全相同，人们用 $\langle hij\rangle$ 来表示这一族晶列，例如用 $\langle 110\rangle$ 代表 $[110]$，$[1\bar{1}0]$，$[101]$，$[10\bar{1}]$ ……等晶向构成的晶向族。表示晶面的参数称为密勒（Miller）指数，它是这样定义的：考虑某晶面族中不通过原点而最靠近原点的一个晶面，它在三个主轴上的截距分别为 $\boldsymbol{a}/u$，$\boldsymbol{b}/v$，$\boldsymbol{c}/w$，则 (uvw) 称为该晶面的密勒指数。类似地，如果因晶体对称性存在多种晶面，其上格点排列完全相同，人们用 $\{uvw\}$ 来表示这一族晶面。不难证明，立方晶格中一个晶面的密勒指数和该晶面法线的晶向指数相同。

作为例子，我们对几种最重要的晶格作一些进一步的说明。惰性元素和许多金属具有密堆积结构，即面心立方（FCC）或密积六角（HCP）结构，这和等径球体的堆积类似。图 1.2-8 显示这一过程，第一层球体按最密集的六角方式排列，各球占据的位置记为 A 位。第二层球体的排布可以有两种选择，即全部占 B 位或全部占 C 位。这里我们选择占 B 位。那么第三层球体的排布仍可以有两种选择，即全部占 C 位或全部占 A 位。不难验证，如采用 *ABCABC*……的堆积方式则形成面心立方结构，纸面为该结构的 $\{111\}$ 面。如采用 *ABABAB*……的堆积方式则形成密积六角结构，纸面为该结构的底面。这两种结构中球体的空间占有率是最高的，达 74.04%，无从突破。

自然界还有许多元素具有体心立方（BCC）结构，如 Li，Na，K，Rb，Cs，Fe 等。该结构的堆砌过程示于图 1.2-9。纸面为 $\{001\}$ 面，为了保证形成立方晶格，$\{001\}$ 面上呈四方排列的球体之间不能紧靠而必须留有 $0.31r_0$ 的间隙，r_0 为球体的半径。如此反复便构成了体心立方结构。应该注意到，体心立方结构的原子密排方向不在 $\langle 001\rangle$ 而是在 $\langle 111\rangle$ 方向。体心立方结构中球体的空间占有率为 68.01%，也是比较高的。其中稍大的间隙可充填较小的球体，这是碳原子可部分溶于铁的原因。

如前所述，构成晶体的基元并不一定是单个原子，也可是分子或原子团。图 1.2-10a，b 分别显示金刚石结构和 NaCl 结构的晶胞。金刚石结构是这样构成的：由面心立方晶胞的中心到顶角引八条对角线，在其中互不相邻的四条对角线的中点各加一个原子。该结构的特点是每个原子有四个最近邻，它们正好在一个正四面体的顶角处。显然，金刚石结构中原子的空间占有率是很低的。细察金刚石结构的晶胞不难发现其中的碳原子可分为两类，位于顶角和面心上的为一类，其他的为另一类。二者的环境，或者说共价键的方向不同。金刚石结构的基元由图 1.2-10a 中 A，B 两原子共同构成，如果我们把基元的代表点取在 A 原子及其同类原子上，立即可以看出金刚石属于面心立方晶格。如果我们把基元的代表点取在 B 原子及其同类原子上结果完全相同。图 1.2-10b 中显示的 NaCl 结构看似简立方晶格，实际上它的基元由

一个 Na 离子和一个相邻的 Cl 离子组成，它也属于面心立方　晶格。

表 1.2-1　分属于七个晶系的十四种布喇菲格子

晶系	晶格类型					相关的点群
	P（初基）	I（体心）	C（底心）	F（面心）	R（菱面体）	
三斜(triclinic) $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma \neq 90°$						$1, \bar{1}$
单斜(monoclinic) $a \neq b \neq c$ $\alpha = \gamma = 90°$ $\beta \neq 90°$						$2, m, \frac{2}{m}$
正交 (orthorhombic) $a \neq b \neq c$ $\alpha = \beta = \gamma = 90°$						$222, 2mm,$ $\frac{2}{m}\frac{2}{m}\frac{2}{m}$ (mmm)
四方(tetragonal) $a = b \neq c$ $\alpha = \beta = \gamma = 90°$						$4, \bar{4}, \frac{4}{m}\ 422,$ $4mm, \bar{4}2m,$ $\frac{4}{m}\frac{2}{m}\frac{2}{m}(4/mmm)$
六角(hexagonal) $a = b \neq c$ $\alpha = \beta = 90°$ $\gamma = 120°$						$6, \frac{3}{m}\frac{6}{m}, 622,$ $6mm, \bar{6}2m,$ $\frac{6}{m}\frac{2}{m}\frac{2}{m}$ $(6/mmm)$
菱面体 (rhombohedral) $a = b = c$ $\alpha = \beta = \gamma \neq 90°$					=	$3, \bar{3}, 32, 3m$ $\bar{3}\frac{2}{m}(\bar{3}m)$
立方(cubic) $a = b = c$ $\alpha = \beta = \gamma = 90°$						$23, \frac{2}{m}\bar{3}(\bar{3}m),$ $432, \bar{4}3m,$ $\frac{4}{m}\bar{3}\frac{2}{m}(m\bar{3}m)$

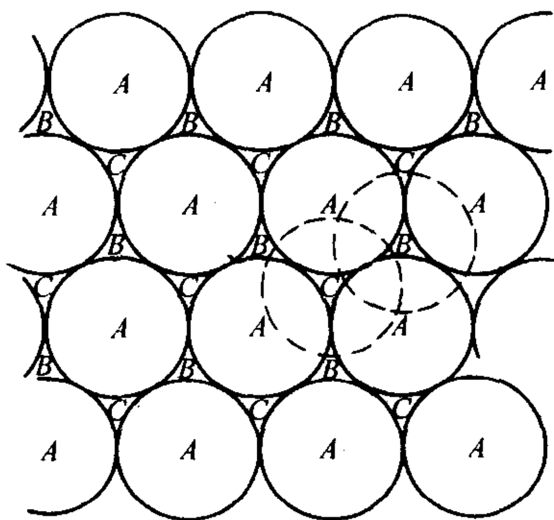


图 1.2-8　刚性球体密堆积的两种方式

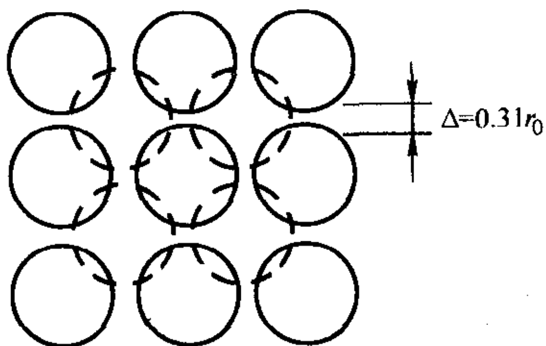


图 1.2-9　体心立方晶格的堆砌方式

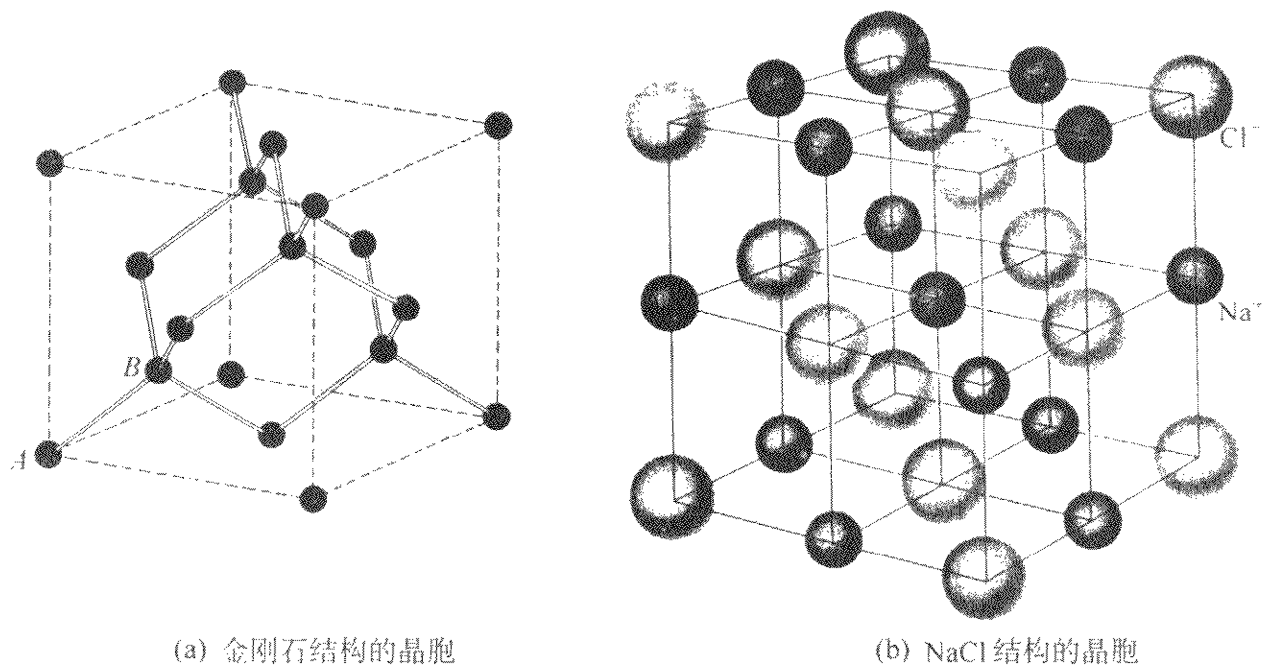
1.2.2　有序与无序

有序与无序的概念存在于物质结构的多种类型和层次之中，由于它们描述系统的混乱程度，所以它和系统的熵，从而和系统的自由能及多个物理量有密切的关系，对研究材料的结构与相变有重要意义。

让我们先以围棋子在棋盘网格上的排布为例来作一形象的说明。可以采用两种模式将一些同一种颜色的棋子排布在棋盘上，一种是井然有序地排在棋盘网格结点上，另一种是杂乱无章地撒在棋盘上。前者有序，相当于晶态，而后者无序，相当于液态或非晶态，这里的无序指的是棋子的排布完全不受棋盘网格的约束，如图 1.2-11 所示。

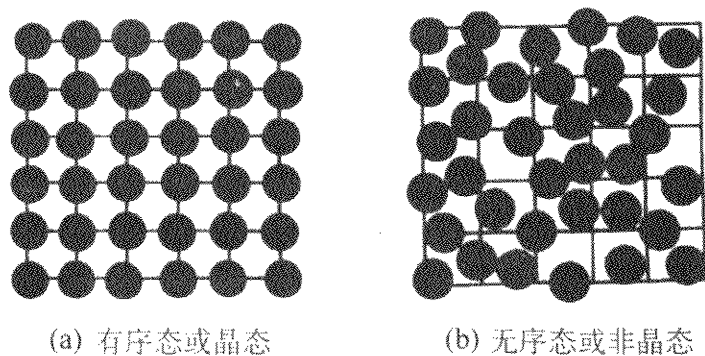
如果将数量相同的黑白两种棋子排布在棋盘网格上，除了前面所说的情况外还有更多的变化，一种方式是将棋子黑白相间，井然有序地排布在结点上，如图 1.2-12a 所示。这相当于有序合金中不同元素原子的周期性有序排列。第二种方式是将黑白棋子不加选择地排布在结点上，如图 1.2-12b 所示，这相当于晶态无序合金中不同元素原子的无序排列。第三种方式是将黑白棋子井水不犯河水地分别放置在棋盘不同区域的结点上，如图 1.2-12d 所示，这是另一种有序态，相当于二元合金中发生了相分离。对于第二种方式也可变换一种方法来表示，如图 1.2-12c 所示。这是因为无序排列实际上意味着两种棋子占据任一格点的几率相同，是黑是白纯

粹出于偶然。我们不妨认为所有棋子都是半黑半白的，用以表示黑白两种棋子占据每一格点的几率相同。这对大量原子组成的晶格既接近真实、又为数学模型的建立提供了方便。



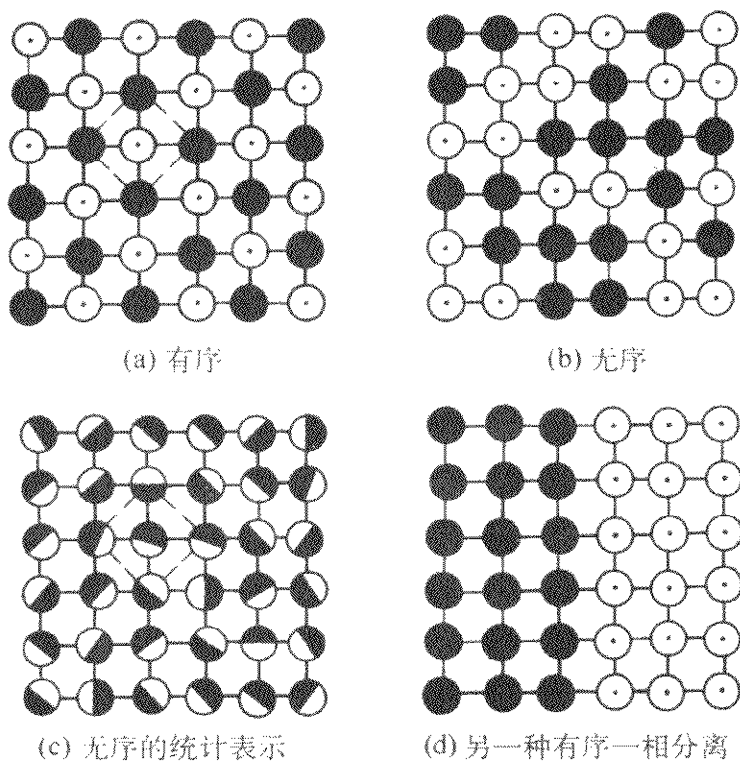
(a) 金刚石结构的晶胞　(b) NaCl 结构的晶胞

图 1.2-10



(a) 有序态或晶态　(b) 无序态或非晶态

图 1.2-11　棋盘上同色棋子排布的有序与无序



(a) 有序　(b) 无序

(c) 无序的统计表示　(d) 另一种有序—相分离

图 1.2-12　棋盘上黑白棋子分布的有序与无序

广而言之，除原子排列有序与无序外，描述材料性质的一些参量也存在有序与无序的问题，这些参量称为序参量，例如铁电体的电偶极矩，铁磁体的磁矩和结构相变中原子的相对位移等。关于这些问题在“材料的相变”一节中我们将再作讨论。

1.2.3　非晶态

晶体中原子的排列具有周期性，称为长程序。非晶态材料原子的排列不具有周期性，但由于各种原子的大小乃至成键的方向都有自身的规律，非晶态材料原子的排列也并非杂乱无章，虽然没有长程序，但存在短程序。所谓短程序的涵义是：近邻原子之间的距离，近邻原子的种类与数目，近邻原子的方位等存在某些序。当然，短程序不能惟一确定非晶态材料的结构。

前面已经介绍，在硅晶体中每个硅原子与周围四个硅原子以共价键结合形成正四面体结构。非晶硅材料中每个硅原子周围也有四个硅原子，但构成的四面体发生了畸变，或者说相邻硅原子之间的键长和键角有了一些无规起伏。非晶硅材料就是由这些畸变四面体构成的不太规则的网络，称为无规网络模型，如图 1.2-13 所示。

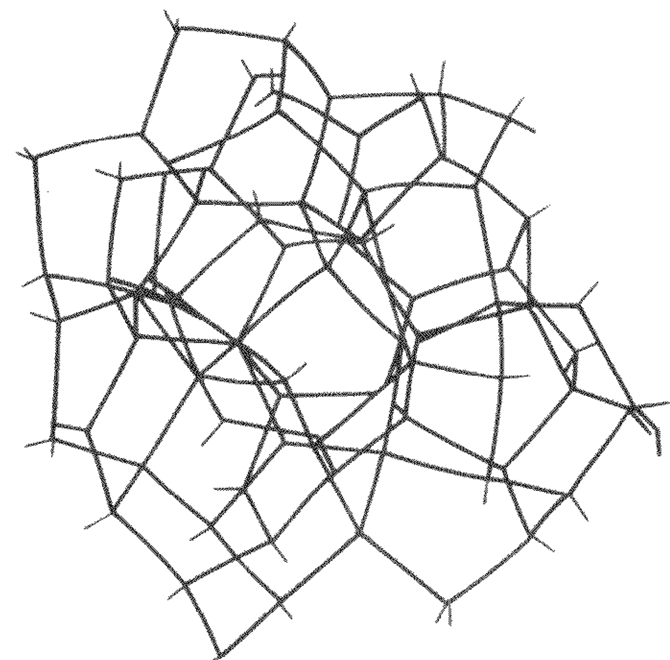


图 1.2-13　非晶硅的无规网络结构模型

非晶态金属又称金属玻璃，由于金属键没有明显的方向性，其结构可用硬球无规堆积模型来描述。这一模型用不同直径的硬球代表不同元素的原子，将这些硬球放入布袋充分混合并敦实压紧，球的排列状态便和金属玻璃中原子的排列状态相近。描述其短程序的参量是径向分布函数 $G(R)$，由同一种元素构成的金属玻璃的二维示意见图 1.2-14。在三维系统中，我们取某一原子为原点，如果中心落在半径为 R，厚度为 $\mathrm{d}R$ 的球壳内的原子数为 $\mathrm{d}n$，单位体积中的原子数为 N，那么该原子周围原子径向分布函数为

$$g(R) = \mathrm{d}n/4\pi R^2 N\mathrm{d}R \qquad (1.2\text{-}3)$$

分别以许多原子为原点重复以上计算然后求平均，便得到该材料的径向分布函数 $G(R)$，它实际上描述了在原子外围其他原子分布随距离变化的状况，是统计平均的结果。$G(R)$的具体形式可由 X 射线衍射的试验结果计算得到。非晶态材料中原子和它的第一近邻原子大都非常接近但不能互相重叠，所以在 $R < R_1 + R_2$（其中 R_1，R_2 分别为相邻两

原子的半径）处 $G(R)=0$，在 R 稍大于 R_1+R_2 处 $G(R)$ 出现一较尖锐的峰值，而 R 很大处 $G(R)=1$，标志着长程序的消失。对于晶态材料，因为长程序的存在，径向分布函数 $G(R)$ 是一系列尖锐的峰。

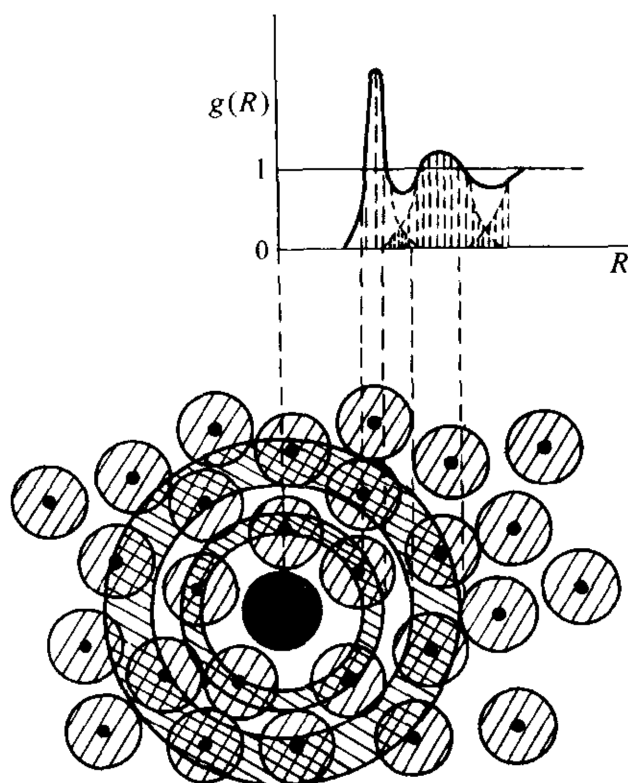
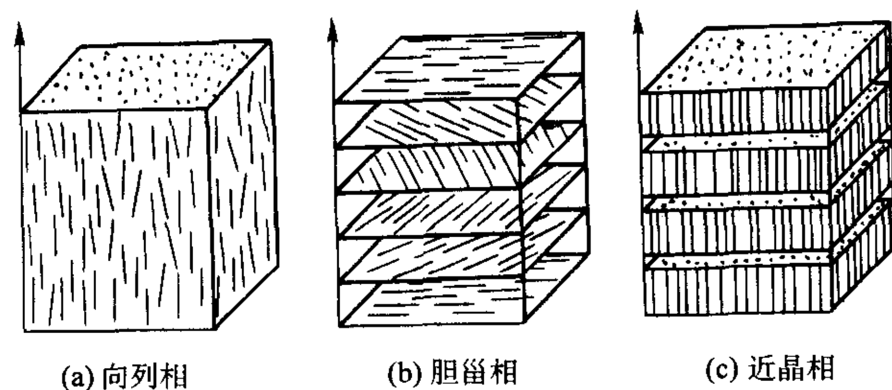


图 1.2-14 二维系统中径向分布函数的示意

1.2.4 液晶态

一些具有刚性的杆状或盘状分子的有机物质处于液态或被溶剂溶解的状态时，其分子的排列出现介乎于液态与晶态之间的情况，即分子位置无序或半无序，而分子取向仍具有长程序，这便是液晶态。分子取向有序使液晶具有明显的各向异性。属于自身可成为液态的有机物液晶称为热致液晶，因为它们多由有机液体经降温使其分子取向有序化而获得。另一种液晶是将有机物溶解于溶剂中，然后将溶液浓缩，致使溶质分子之间交互作用增强而实现有序排列，这种液晶称为溶致液晶。

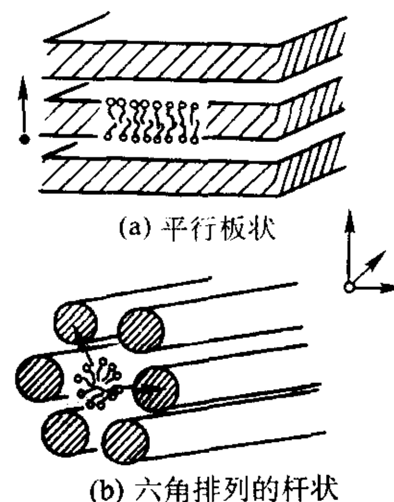
热致液晶可分为三类：向列相、胆甾相和近晶相，其结构示意见图 1.2-15。向列相液晶的特征是其分子中心的位置分布完全无序和其杆状分子的取向具有长程序，如图 1.2-15a 所示。由于分子位置分布完全无序，向列相液晶有良好的流动性。胆甾相液晶的分子是长扁平形的，沿长轴方向平行排列成层状，层平面与扁平分子的平面一致，而相邻层分子的轴向发生规则的扭转，如图 1.2-15b 所示。由于扭转层的作用，它反射的白光发生色散，透射光的偏振面发生旋转，使它得到重要的光学应用。近晶相中杆状分子相互平行地排列成层状，杆与层平面垂直或与层平面成一定的角度，在同一层中分子中心的位置分布是无序的，分子只能在本层内移动，其结构示意见图 1.2-15c。近晶相各层相互平行排列，层间还可以滑动。这三类液晶中近晶相的结构是最接近晶态的。



(a) 向列相 (b) 胆甾相 (c) 近晶相

图 1.2-15 三种液晶相中分子排列的示意

溶致液晶也是分子排列部分有序的液体。由于其中增加了溶液浓度这一变量，溶致液晶中的结构变化更为复杂，并可用相图来描述。肥皂水是一种典型的溶致液晶，它的分子是一种“双亲分子”，头部具有亲水性，尾部具有疏水性和亲油性，这是用肥皂水可洗去油污的主要原因。在肥皂水溶液里肥皂分子头部尽量亲近水，而尾部尽量躲开水，办法是肥皂分子采用尾对尾而头对水的方式聚集在一道。取决于温度和浓度的不同，聚集的形态可以是板状（图 1.2-16a），杆状（图 1.2-16b），或更复杂的形态，在这些聚集体内部分子位置并不具有周期性。由于范德瓦尔斯吸引力这些板或杆或其他聚集单元出现有序排列，这便是溶致液晶不同的相。这一过程称为“自组装”，其意义远不止在溶致液晶。在范围更广的软物质研究与制备领域它有广阔的用武之地，而软物质的研究将推动有机化学，高分子材料，乃至生命科学的发展。



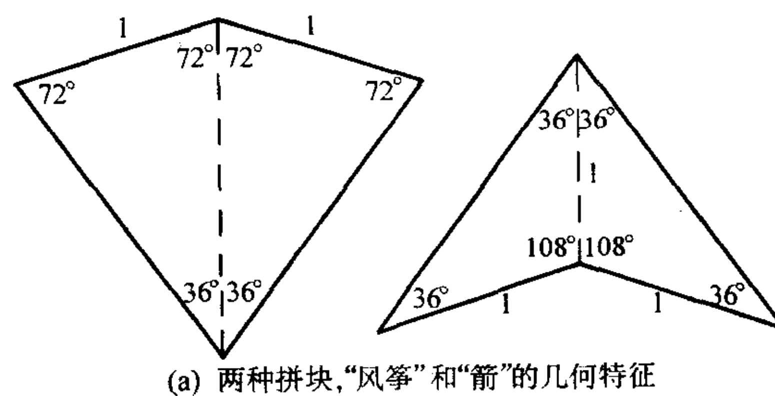
(a) 平行板状

(b) 六角排列的杆状

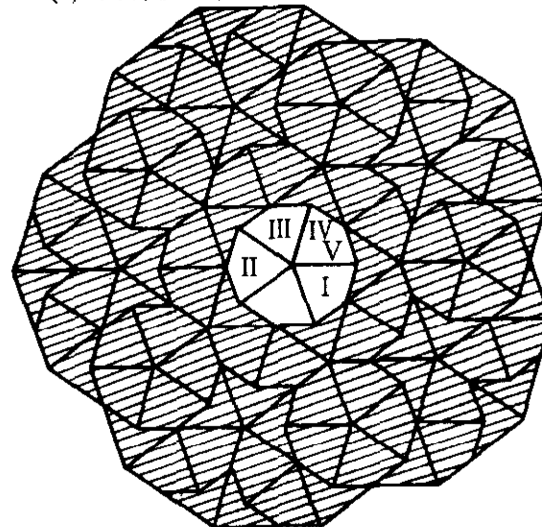
图 1.2-16 两种溶致液晶的结构

1.2.5 准晶

如前所述，从几何学的角度考虑，将许多同样的元胞在空间作平移周期性排列填满空间便形成了晶体，其结构与对称性已为人们熟知。但是，放弃了平移周期性，是否可以采用某些基本拼块无空隙地布满空间呢？答案是肯定的，以二维问题为例，1974 年 Penrose 提出，用图 1.2-17a 所示的两种拼块，即“风筝”和“箭”完全可以做到这一点。图 1.2-17b 是用“风筝”和“箭”拼出的许多种具有五次对称性的图案之一，它不具有平移周期性。



(a) 两种拼块，“风筝”和“箭”的几何特征



(b) 一种 Penrose 拼砌图案

图 1.2-17 二维 Penrose 拼砌

1984 年 Shechtman 等人在快淬的 Al-Mn 合金的电子衍射图中发现了具有五重对称的斑点分布，它和周期结构是相矛盾

的，因而在传统晶体结构理论框架内无法予以解释。Steinhardt等人提出，这种材料的结构正是Penrose拼砌的三维推广，它具有20面体对称性，并称其为“准晶”。准晶态的特点是具有长程取向序而没有长程平移对称性，它是一种新的有序相。按照这一模型计算出来的该种快淬的Al-Mn合金的衍射花样，包括斑点的位置与强度，都和试验结果吻合。

1.3　晶体缺陷

实际晶体中存在晶体缺陷，它指的是晶体中原子的周期性排列受到破坏的微小区域。通常人们按缺陷的维数将其分类。零维缺陷即点缺陷，一维缺陷主要指位错，二维缺陷即表面与界面。材料的许多性质，如强度，扩散，离子导电等，对缺陷的存在极为敏感。

1.3.1　点缺陷

点缺陷表示晶格的周期结构在某些点受到破坏，通常是由于晶格中个别原子的缺失或错位造成的。在热力学平衡态下点缺陷是存在的，这是因为点缺陷使系统的熵和内能都增加，所以一定浓度点缺陷的存在可使系统的自由能为最低，相应的点缺陷浓度称为平衡浓度。由于经历多种非平衡过程，例如淬火，机械加工，粒子辐照，合金化等，晶体中的实际点缺陷浓度可高于，甚至远高于其平衡浓度。按其原子的构成，点缺陷可分为自身点缺陷和杂质点缺陷。前者包括空位和自填隙原子，其结构示意见图1.2-18。后者包括代位杂质原子和填隙杂质原子，其结构示意见图1.2-19。在有序合金中还会存在反位原子。注意点缺陷往往伴随有周围的晶格畸变。在半导体晶体中杂质原子不仅造成晶格周期性的破坏，而且经常导致禁带中的缺陷电子能态的产生，这对半导体材料制备是非常重要的。点缺陷的移动是材料中原子扩散的主要机制之一。离子晶体中点缺陷在电场作用下的定向移动就是离子导电现象。

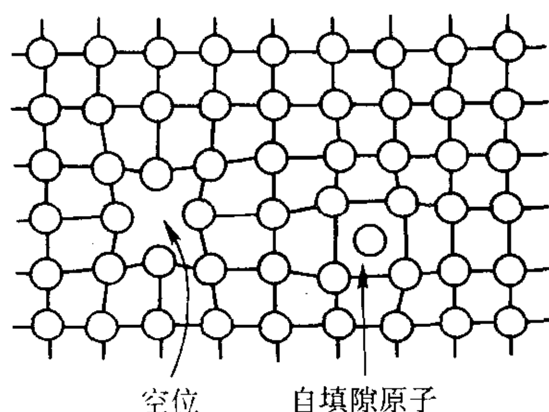


图1.2-18　二维简立方晶格中的空位和自填隙原子

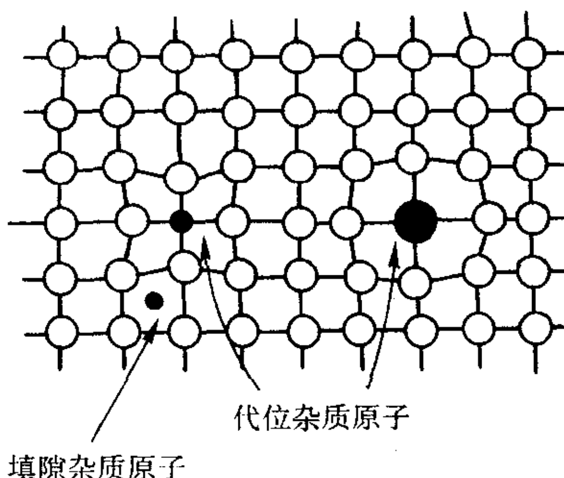


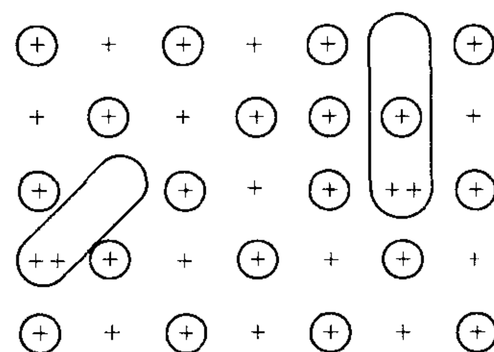
图1.2-19　二维简立方晶格中代位杂质原子和填隙杂质原子

点缺陷的平衡浓度可由热力学进行计算。以空位为例，如果空位的形成能为E_v，按Boltzmann统计，在热平衡状态下任一格点被空位占据的几率P亦即空位的平衡浓度C_0为：

$$P = C_0 = \exp(-E_v/k_B T) \qquad (1.2\text{-}4)$$

式中，k_B为玻尔兹曼常数；T为绝对温度。如果$E_v = 1$ eV，$T = 1\,000$ K，立即可以求出空位的平衡浓度$C_0 = 10^{-5}$。利用类似的方法不难求出其他点缺陷的平衡浓度表达式。

在某些情况下，由于点缺陷之间的交互作用，两个或几个点缺陷会形成点缺陷复合体，如氯化钠晶体中的二价杂质钙离子旁边总存在一个钠离子空位以保持电中性，如图1.2-20所示。而体积较大的杂质原子旁边往往伴有一个空位以减小弹性畸变。纯净的碱卤晶体对可见光是透明的，不显示任何颜色。在这些晶体中加入多余的一价金属离子或以少量二价杂质金属离子取代一价金属离子可产生阳离子-阴离子空位复合体。它们可吸收某些波长的光而使晶体显示某些颜色。这些缺陷复合体称为“色心”。



“+”表示纸面内的正离子；
“⊕”表示纸面上或下的正离子；
“++”表示双价杂质正离子

图1.2-20　氯化钠中的杂质钙离子与伴生的钠离子空位

1.3.2　线缺陷

晶体中的线缺陷主要指位错，位错是一种非平衡态缺陷，是因晶体生长条件偏离平衡态，机械加工或异质外延等过程引入的。正是位错的存在及其易动性使实际晶体的屈服强度大大低于完美晶体屈服强度的理论计算值。

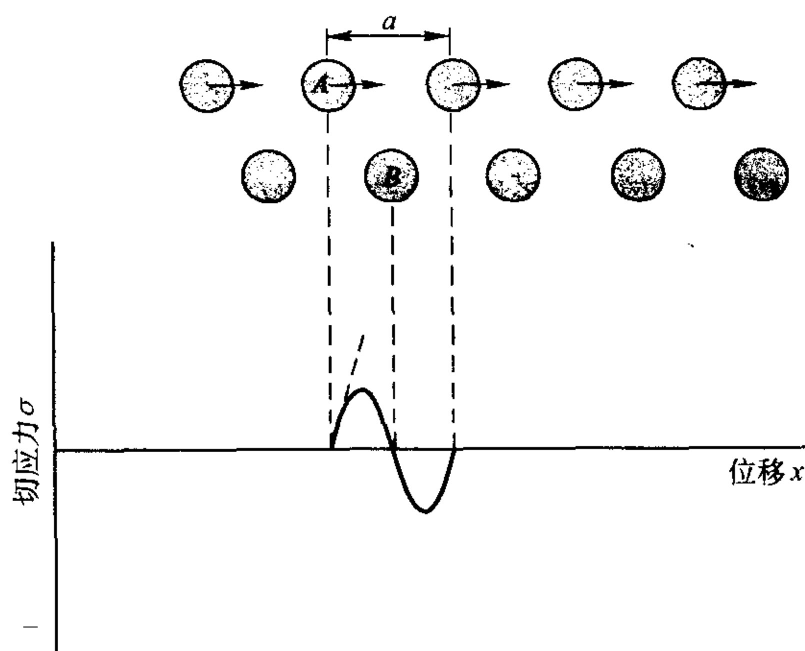


图1.2-21　晶体中相邻原子面因剪应变导致整体滑移的过程中应力与应变的关系

弗仑克尔（Frenkel）用图1.2-21所示的简单模型来估算完美晶体的屈服强度。考虑二维晶格，图示为平衡位置。对于小的弹性应变，切应力σ与两原子面相对位移x的关系可表示为：

$$\sigma = Gx/d \qquad (1.2\text{-}5)$$

式中，G为切变模量；d为相邻原子面间距。随着相对位移的增大，当A原子到达B原子上方时应力为0而系统处于不平衡状态。作为粗略近似，不妨假定应力与位移满足以下关系：

$$\sigma = (Ga/2\pi d)\sin(2\pi x/a) \qquad (1.2\text{-}6')$$

式中，a为滑移面中沿移动方向的原子间距。由式（1.2-6）可求出最大切应力σ_c出现在$x = a/4$处，若进一步假定$a = d$，不难求出：

$$\sigma_c = G/2\pi \quad (1.2\text{-}7)$$

也就是说使完美晶体相邻原子面间发生滑移所需应力约为晶体切变模量的1/6。但是实验表明，实际晶体的屈服强度比这要低$10^2 \sim 10^5$倍。Taylor，Orowan，Polanyi等指出，是晶格中的一种线缺陷，即位错使滑移所需应力大大下降。最简单的位错有两种，即刃位错和螺位错。

位错线是晶体滑移面上已发生滑移的区域和未发生滑移的区域之间的界线。可以设想以下操作在晶体中造成一条位错线：将晶体沿某一晶面切开，切开部分终止于$A-A$线，如图1.2-22a所示。将切面的两岸作一平行于切面的相对位移，若平移矢量为点阵平移矢量，则可将两岸原子重新键合起来而在切面处不留任何痕迹，只沿$A-A$线留下一条原子错排区，该线便是位错线。描述位错线的特征需要两方面的参量，即它在空间的走向和切面两岸的相对位移矢量$\boldsymbol{b}$。后者称为伯格斯（Burgers）矢量，表征位错的强度。如果伯格斯矢量$\boldsymbol{b}$为点阵平移矢量，那么上述操作过程中切面的选择并不是惟一的。可以证明，只要位错线的走向和$\boldsymbol{b}$是一定的，无论选取哪个切面进行操作所得到的位错线都是相同的。两种最基本类型的位错是刃位错和螺位错。上述操作过程中如果$\boldsymbol{b}$与位错线垂直，形成的位错为刃位错。如果$\boldsymbol{b}$与位错线平行，形成的位错为螺位错。在简单立方晶格中取$\boldsymbol{b}$为立方格子的边长，这两种操作过程的示意分别见图1.2-22b、c。刃位错的滑移有点像蠕虫的爬行。如果蠕虫永远伸直身体，它是无法爬行的。为了前进，蠕虫先将尾部前移一步同时在背部形成一个凸起，然后让这一凸起向前移动直达头部，最后头部向前移动从而完成这一以屈求伸的过程。其示意图见图1.2-23。刃位错的显著特点是它有多余的半个原子面。显然，刃位错的滑移面是惟一确定的，它便是与多余半原子面垂直的平面，或者说是由位错线与$\boldsymbol{b}$确定的平面。刃位错线有时也可作垂直于多余半原子面的缓慢移动，这种运动叫做“攀移”。攀移伴随着半原子面边缘处或者说刃位错核心处原子的扩散，例如空位的到达可使半原子面后退，自填隙原子的到达可使半原子面伸长。

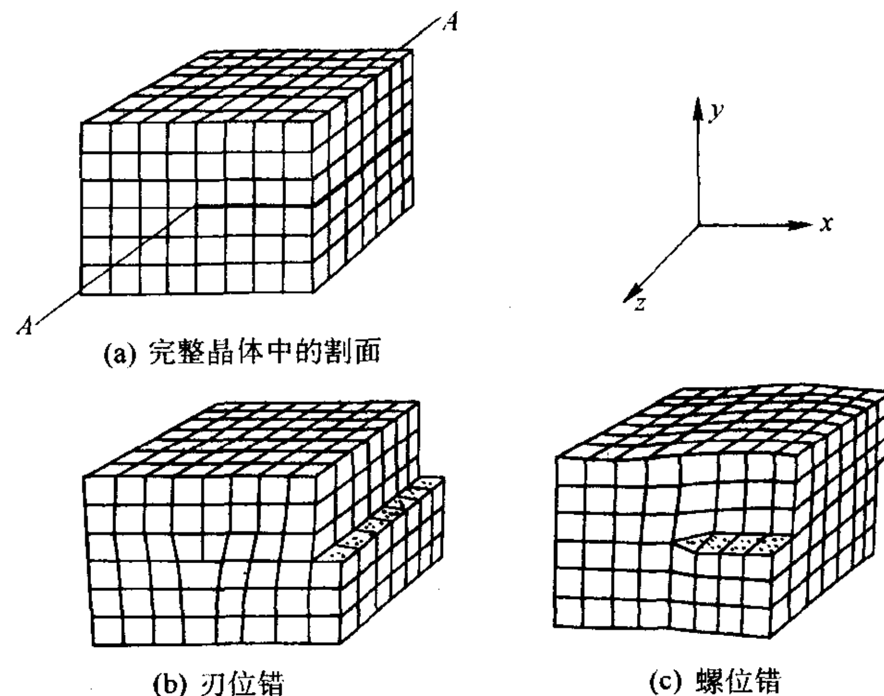


图1.2-22 两种位错线形成和滑移的示意

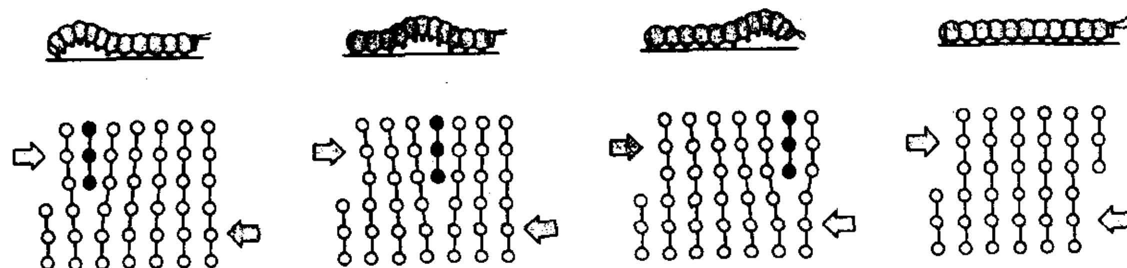


图1.2-23 刃位错的滑移与蠕虫爬行的比较

螺位错线周围原子排列的示意见图1.2-24，实际上原子层围绕螺位错作螺旋状排列，如果在原子面上绕位错线走一周，就会从一个原子面走到下一个原子面上。螺位错的滑移面不是惟一的，实际上包含螺位错线的任何面都可成为其滑移面。实际晶体中的位错线往往是混合型的，其$\boldsymbol{b}$与位错线既不垂直也不平行。图1.2-25画出了一段弯曲位错线附近原子排列的状况，可以看出，与$\boldsymbol{b}$垂直的一段位错线为刃位错，而与$\boldsymbol{b}$平行的一段位错线为螺位错，其他部分为混合位错。

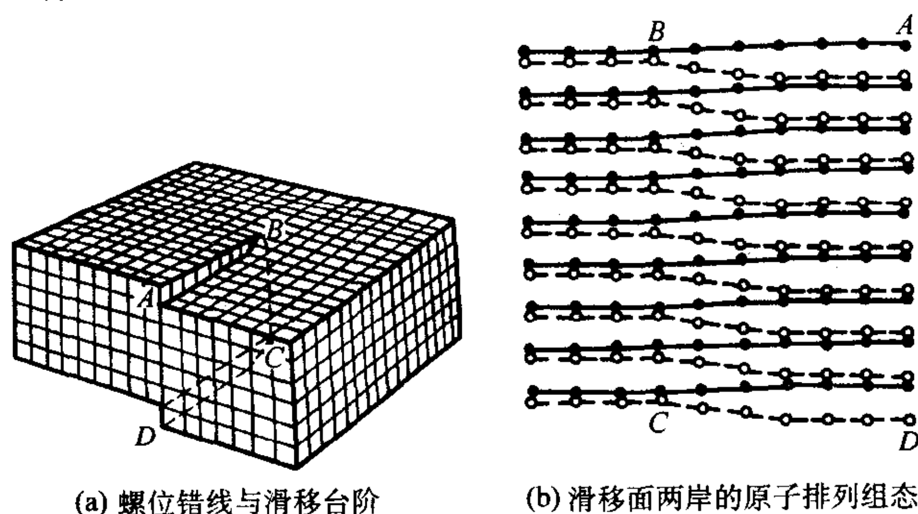


图1.2-24 螺位错线周围原子排列的示意

由以上的介绍可知，位错线受到在其滑移面内与其伯格斯矢量平行的切应力作用时会发生移动而导致滑移。从能量考虑，我们可以引入位错线受力的概念。考虑单位长度位错线在其滑移面内切应力τ作用下移动了一个小的距离dl，这意味着这段位错线扫过的面积dl的两岸发生了相对滑移b，这样切应力τ所作的功为$dw = dl\tau b$，定义dw/dl为单位长度位错线所受的力F，于是有：

$$F = \tau b \quad (1.2\text{-}8)$$

这就是说单位长度位错线所受的力等于其滑移面内沿伯格斯矢量的切应力分量与伯格斯矢量大小的乘积，其方向与位错线垂直。

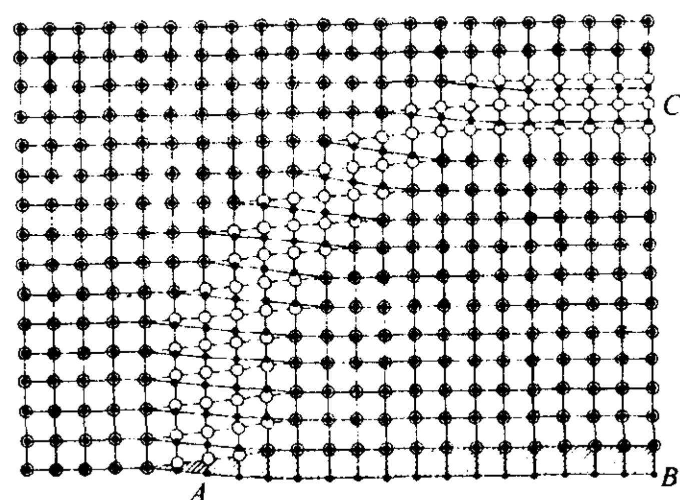


图1.2-25 一段弯曲位错线附近原子排列的状况
A处为螺位错；C处为刃位错

由于位错线周围的晶格发生畸变，位错具有能量并产生应力场。螺位错的应力场较为简单并具有圆柱对称性，图1.2-26a表示围绕螺位错的介质中半径为r的薄管状区域应

变的分布。它只有沿轴向的切应变和切应力分量，容易求出其应变 γ 和应力 σ 分别为：

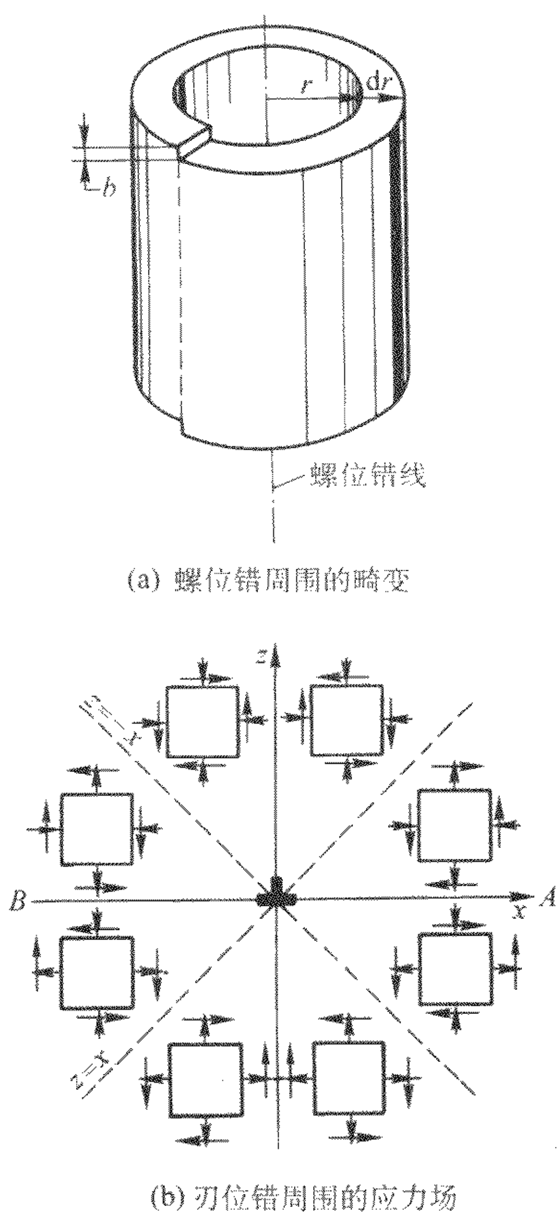
$$\gamma = b/2\pi r \quad (1.2\text{-}9)$$

$$\sigma = Gb/2\pi r \quad (1.2\text{-}10)$$

式中，G 为切变模量；b 为伯格斯矢量的大小。考虑到位错核心区原子严重错排，由弹性力学求出的式（1.2-9）、式（1.2-10）在位错核心区已不适用。同时考虑到晶体中有多种缺陷存在，一条位错线的应力场只能存在于有限空间。我们假设弹性应变存在于 r 由 r_0 到 R 的范围内，通过积分可以求出单位长度螺位错线的弹性能为：

$$E_s = \int_{r_0}^{R} (Gb^2/4\pi)\ \mathrm{d}r/r = Gb^2 \ln(R/r_0)/4\pi \quad (1.2\text{-}11)$$

式中，r_0 通常取为晶格常数；R 通常取为晶粒的尺寸，一般认为 E_s 值约为 Gb^2 左右。在典型金属材料中，其数值大约为 100 eV/nm。由于这个能量值非常高，仅靠原子的热运动产生位错的可能性极低。



(a) 螺位错周围的畸变

(b) 刃位错周围的应力场

图 1.2-26 位错应力与应变的计算

刃位错具有平面应力状态，但是没有圆柱对称性，它的应变场和应力场的计算较为困难。我们仅在图 1.2-26b 中标出各象限中正应力和切应力的方向并给出各应力分量的表达式如下：

$$\sigma_{xz} = [Gb/2\pi(1-v)]\ x(x^2-z^2)/(x^2+z^2)^2$$

$$\sigma_{xx} = -[Gb/2\pi(1-v)]\ z(3x^2+z^2)/(x^2+z^2)^2$$

$$\sigma_{zz} = [Gb/2\pi(1-v)]\ z(x^2-z^2)/(x^2+z^2)^2$$

$$\sigma_{yy} = v(\sigma_{xx}+\sigma_{zz}) \quad (1.2\text{-}12)$$

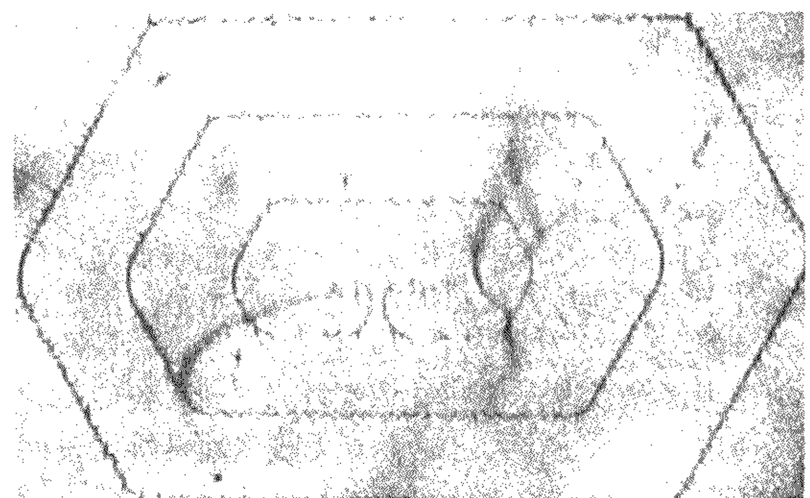
式中，v 为材料的泊松比。借助于图 1.2-26b 可以对刃位错周围的应力场分布作一些讨论。在图的上半部恒有 $\sigma_{xx}<0$，表示介质在垂直于半原子面方向受到压缩，在图的下半部恒有 $\sigma_{xx}>0$，表示介质在垂直于半原子面方向受到伸张，这显然是多余半原子面作用的结果。进一步的计算可以证明，在上半部介质的体积膨胀率为负，受到压缩，而在下半部介质的体积膨胀率为正，受到膨胀。这种应变分布导致体积小的杂质原子富集于刃位错核心区的上半部，而体积大的杂质原子富集于刃位错核心区的下半部。杂质原子的这种富集区称为科垂尔（Cottrell）气团，它使刃位错的滑移变得困难，是金属材料固溶强化的主要物理依据。由以上刃位错应力场表达式不难求出单位长度刃位错的弹性能为：

$$E_e = Gb^2\ln(R/r_0)/4\pi(1-v) \quad (1.2\text{-}13)$$

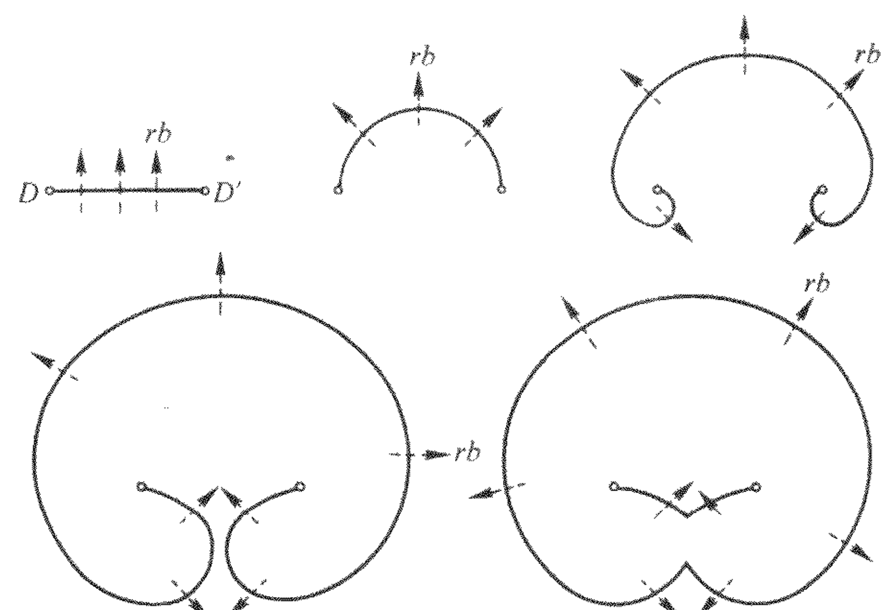
对于多数材料 v 的数值在 0.3 左右，对比式（1.2-11）和式（1.2-13），可以看出刃位错的弹性能高于螺位错。由于位错线具有很高的能量，在外界条件允许的情况下它倾向于尽量缩短自身的长度以降低能量。在各向异性介质中它还倾向于尽量让自己沿着弹性能较低的方向运动。这使位错线的行为象具有张力的弹性弦线，其张力称为位错的线张力，数值与单位长度位错线的能量相当。

我们已经看到，位错的移动最终导致它消逝在晶体表面并产生一个台阶。为了使晶体产生大的范性变形，晶体内部必须有某种机制源源不断产生新的位错。实际上人们已经发现了位错增殖的几种机制，我们介绍其中的一种，即弗兰克－瑞德（Frank-Read）位错源，如图 1.2-27 所示。

设想图 1.2-27b 中的一段位错线两端被钉扎在 DD' 两点不能移动。该段位错线在切应力 τ 的作用下向上弯曲直至与位错的线张力 T 平衡。如果 τ 足够大，位错线就会变得不稳定并绕过钉扎点直至在下部相遇，结果在 DD' 保留下一条位错线并向外发射出一个位错圈。如此反复，弗兰克－瑞德源可以不断地发射位错圈，直至前方的位错圈受阻。



(a) Dash 提供的硅中弗兰克-瑞德源的照片，位错线由铜缀饰



(b) 弗兰克-瑞德源发射位错圈的示意

图 1.2-27 弗兰克－瑞德位错源

如果伯格斯矢量 b 不是点阵平移矢量，那么切面两岸的原子就要发生错排，这种错排面称为“层错面”（下面将详细介绍），这种位错称为“不全位错”。面心立方结构中位错的伯格斯矢量通常为（a/2）[011]，它可以按以下方式分

解为两个不全位错而使总能量下降：

$$b = (a/2)[011] \rightarrow (a/6)[121] + (a/6)[\bar{1}12] = b_{p1} + b_{p2} \quad (1.2\text{-}14)$$

这三个矢量都在 {111} 面上，但是 b_{p1} 和 b_{p2} 都不是点阵平移矢量。这样这两个不全位错之间便形成了一片层错面。实际上 b_{p1} 产生的滑移造成了六方结构的原子堆垛序列，b_{p2} 产生的滑移又回到了面心立方结构。层错面有一定的能量，称为层错能，它类似于表面张力。这一张力与两个不全位错之间的斥力相平衡就决定了两个不全位错之间的平衡距离。这种两个不全位错夹着一片层错面的缺陷组态称为“扩展位错”，在实际材料中经常出现，其示意图见图 1.2-28。

关于位错的性质，还有以下几点需加以注意：

1）无论一条位错线有什么形状或者变成什么形状，它的伯格斯矢量是惟一确定的。

2）位错线不能终止在晶格内部，它要么自己形成闭合的环路，要么终止在另一位错线上形成位错网络，要么终止在晶体的表面或界面处。

3）一般说来，实际晶体中的位错线形成三维的位错网络，在位错网络的结点处各条位错线伯格斯矢量的矢量和为零。

4）位错的滑移面包含位错线自身及其伯格斯矢量。

5）当位错线在其滑移面上滑移时，滑移面两岸原子的相对位移与伯格斯矢量相同，而位错线沿垂直于自身的方向在其滑移面上运动。

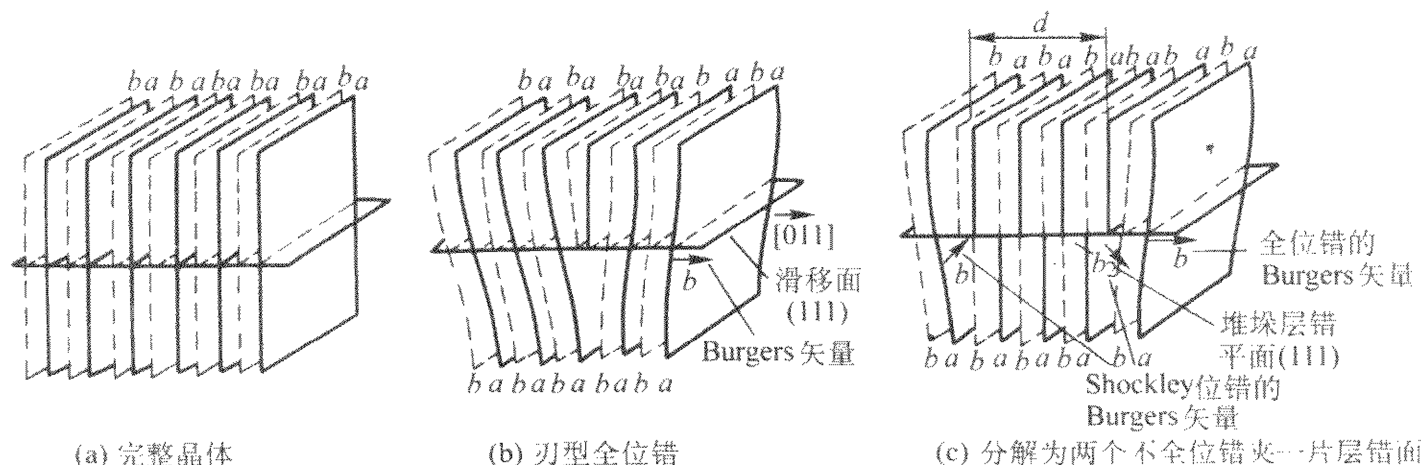


图 1.2-28　扩展位错形成过程的示意

1.3.3　表面与界面

晶体表面是平移对称性的终止处，表面原子键合状态与晶体内部不同，出现了断键。由于键合能具有负值，断键的出现使能量升高，增加单位表面积所带来的自由能的升高 σ 称为表面能。由于表面能的存在，在条件允许的情况下材料倾向于具有小的表面积，例如液滴常为球形，某些结构不太稳定的纳米颗粒也具有近似球形的外形。实际上，晶体的表面能具有各向异性，密排低指数面的表面能较低。这使在近平衡条件下生长的晶体具有规则的，由低指数面围成的多面体外形。

由于表面原子键合状态与晶体内部不同，我们没有理由认为晶体内部的原子排列方式会毫无改变地延续到表面，而这种表面原子排列方式改变的结果将导致表面能的下降而变得更稳定。表面原子排列方式与内部的差异主要有四方面，即表面驰豫，表面重构，表面偏聚和表面吸附，下面逐一进行简单的介绍。所谓表面驰豫指表面晶格结构与内部基本相同，但晶格参数略有变化。因为表面附近几层原子沿垂直于表面方向受力不对称，其晶面间距会有变化，主要是受到压缩。这种法向驰豫在金属表面较为明显。表面重构通常表现为表面晶格结构的变化，主要表现为表面超结构的出现。所谓表面超结构是指表面原子排列周期的成倍放大。最常见的表面重构有两类，即缺列型重构和重组型重构，分别示于图 1.2-29a，b，前者表现为表面原子的部分缺失，常见于金属表面。后者虽不减少表面原子的数量，但显著地改变了表面原子的排列方式，更多地见于共价晶体特别是半导体晶体表面。图 1.2-30 为由扫描隧道显微镜获得的硅 {111} 表面重构形成的 7×7 结构的照片。表面重构常同时伴有表面弛豫以进一步降低表面能。表面偏聚是指固溶体材料溶质原子在表面富集或贫化的现象。表面吸附则指环境中的气相原子分子依附于材料表面的现象，可分为物理吸附和化学吸附两类。除了超高真空中的清洁表面，材料表面上总带有吸附物，表面吸附与脱附对材料的腐蚀，化学催化及许多表面过程起着决定性的作用。

界面的种类繁多，按其结构和形成机制可以分为四类，即平移界面，孪晶界面，晶界和相界。分别介绍如下。

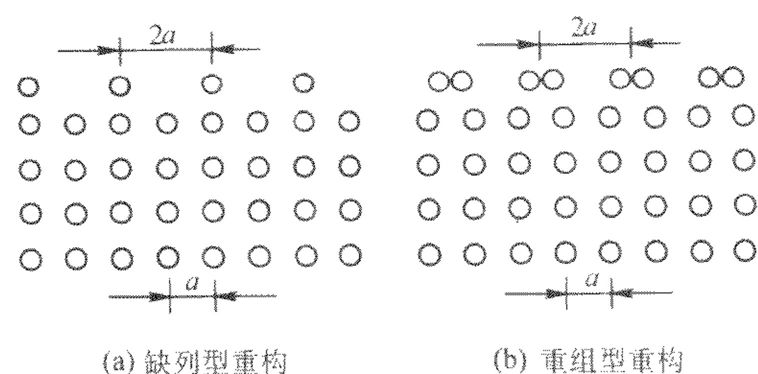


图 1.2-29　简立方晶格中的两种表面重构

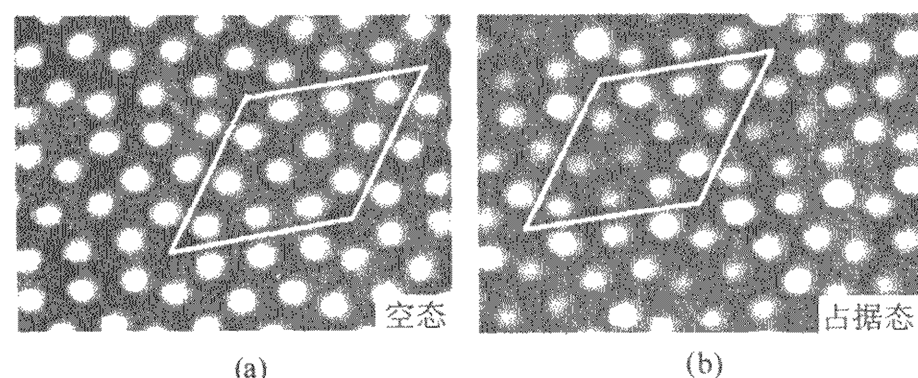


图 1.2-30　Binnig 和 Rohrer 利用 STM 获得的硅 {111} 表面重构形成的 7×7 结构显微图像

试想将晶体沿某一晶面割开，使割面两岸作一相对平移后重新结合。如果相对平移矢量不是点阵平移矢量，那么割面处将留下痕迹而成为界面，这类界面就是平移界面。一般说来，平移界面附近原子出现严重错排，界面能很高，使它们出现的可能性不大。但一些特定的非点阵平移矢量可造成低能量的界面，这便是实际材料中存在的几种平移界面，最常见的平移界面是层错面。以面心立方结构为例，我们在 1.2.1 节中已介绍过，其 {111} 面的堆垛顺序为 *ABCABC*…，如图 1.2-8 和图 1.2-31a 所示。如果在这种正常的结构中抽出或插入一个 {111} 原子面，分别如图 1.2-31b、c 所示，那么便分别形成了抽出型或插入型层错面。类似地，密积六角结构的堆积方式为 *ABABAB*…，可以插入一个 *C* 型原子层

造成一个层错面。由于层错面的出现并不改变附近原子的最近邻关系（配位数，键长，键角），只改变次近邻关系，所以它几乎不产生弹性畸变，只是使局部的电荷周期排列受到破坏，故而层错面的能量（称为层错能）较低。如果层错面终止在晶体内部，它的边界线便是一条不全位错线。

另一种常见的平移界面是反相畴界。我们在1.2.2节中已介绍了有序相的概念。一些合金的有序化导致了超结构的出现，同时也产生了有序畴。以CuAu合金为例，无序的CuAu合金为面心立方结构，Cu原子和Au原子随机地占据格点。有序化后CuAu为$L1_0$结构，Cu原子和Au原子分别占据两套四方亚点阵。二者互相套叠。而Cu原子和Au原子各占据哪套亚点阵则是随机的。于是便可能出现如图1.2-32b～d所示的情形，即有序晶体内形成若干区域，区内保持完整的有序结构，称为有序畴，而相邻有序畴间虽保持共格，但发生了非点阵矢量平移。有序畴之间的这种界面便是反相畴界。

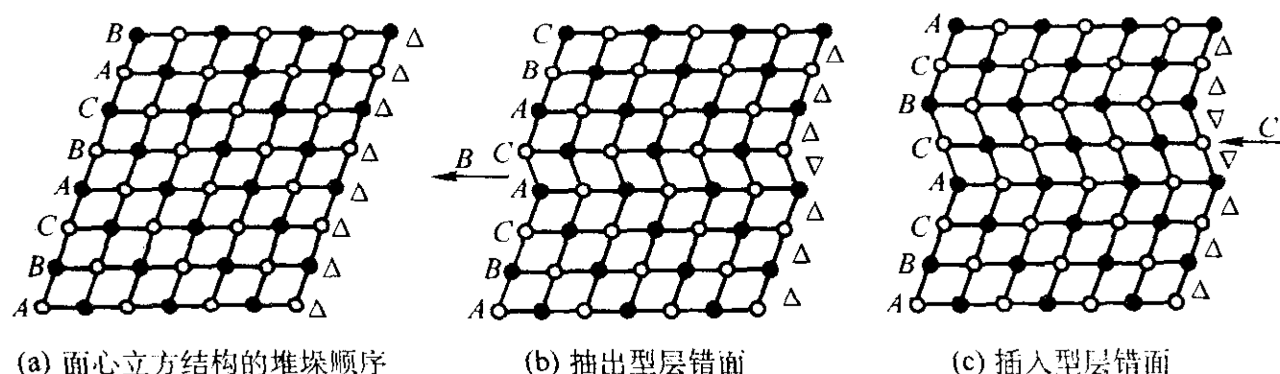
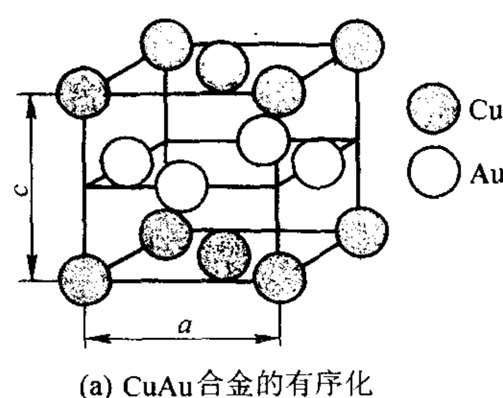
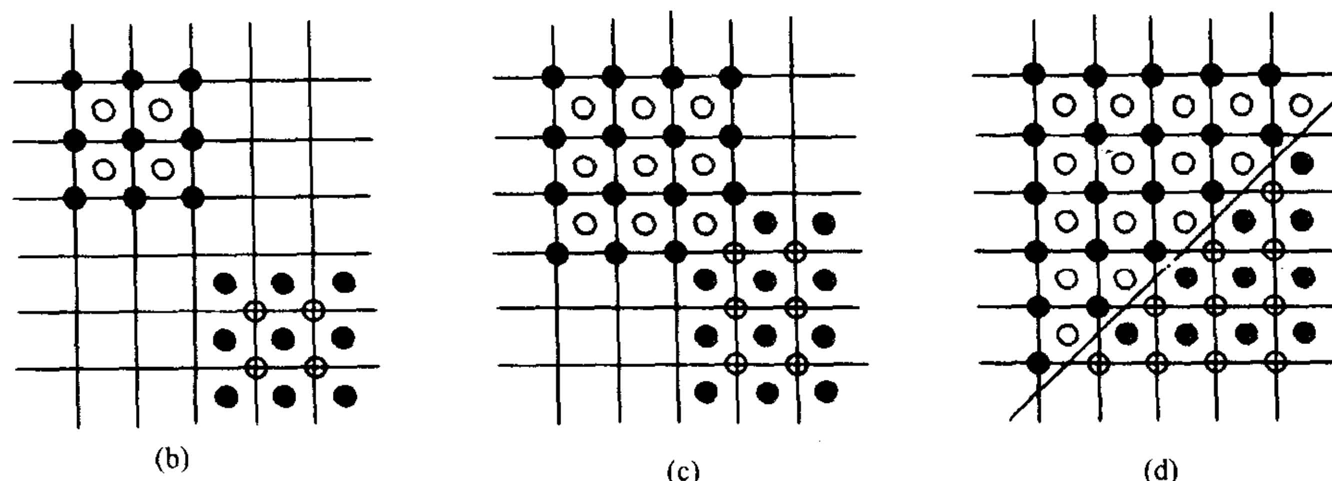


图1.2-31 面心立方结构中层错面的形成



(a) CuAu合金的有序化



(b～d) 有序化过程中反相畴界的形成

图1.2-32 有序CuAu合金中的反相畴界

孪晶界面也是一类常见的面缺陷。仍以面心立方结构为例，实际上图1.2-8所示的面心立方结构中{111}面的堆垛顺序并不是惟一的，它可以是*ABCABC*…，也可以是*ACBACB*…，所形成的都是面心立方结构，其差异只不过是晶体在空间作了一整体的旋转。试想分别取这两种堆垛模式的区域在某一{111}面相遇，形成图1.2-33所示的情形。

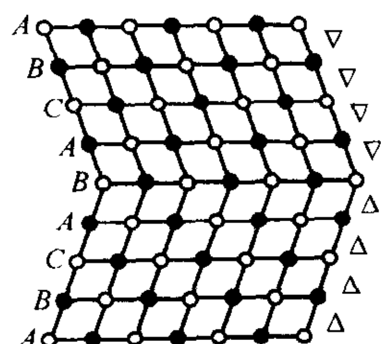


图1.2-33 面心立方结构中{111}面反映孪晶的〈110〉投影图

相遇处的这一{111}面成为一个反映面，其两侧的晶体关于此反映面成镜面对称，堆垛顺序为…*ABCABACBA*…。这时人们便说两侧的晶体呈孪晶关系，其反映面称为孪晶界面。显然孪晶界面的能量也不太高，甚至低于层错能。在强大外应力的作用下，有时晶粒的一部分可发生切变与另一部分形成孪晶关系，这种现象叫做机械孪生，也是范性形变的一种机制。通常多数材料为多晶体，因而晶界是最常见的面缺陷。为了从几何学方面描述一个平的晶界面需要五个角度，即描述晶界面的法向需要两个独立的角度，描述晶界一侧的晶体相对于另一侧晶体的方位需要三个独立的转动角度。如果晶界一侧的晶体相对于另一侧晶体绕位于晶界面内的某个轴发生了旋转，所形成的晶界称为倾侧晶界，如果晶界一侧的晶体相对于另一侧晶体绕垂直于晶界面的某个轴发生了旋转，所形成的晶界称为扭转晶界。

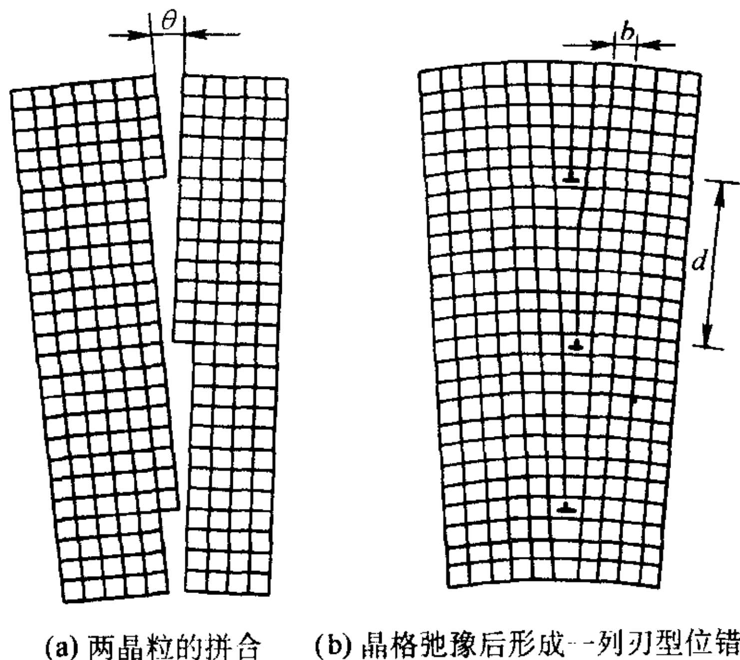
如果晶界两侧晶体的取向差别较小（例如小于2°），这种晶界称为小角度晶界。图1.2-34显示简立方晶格中一个小角度倾侧晶界的结构，它相当于一列同号的刃型位错。不难求出该倾侧晶界中刃型位错间的平均距离d近似为：

$$d = b/\theta \tag{1.2-15}$$

式中，b为刃位错的伯格斯矢量；θ为界面两侧晶体的相对旋转角。图1.2-35显示简立方晶格中一个以(001)面为界面的小角度扭转晶界的结构，它相当于一组螺型位错网络，网格间距也满足式(1.2-15)。在一般情况下，小角度晶界既非纯倾侧型，

也非纯扭转型，这时界面成为混合位错的复杂网络。

如果晶界两侧晶体的取向差别较大，则该晶界称为大角度晶界。一般的大角度晶界情况十分复杂，很难用简单的位错模型描述。现在被广为认可的大角度晶界模型是“重合位置点阵模型”。该理论认为，对某一结构的晶格，绕其某一晶轴旋转一定角度得到不同取向但结构相同的另一晶格，设想两晶格互相穿透，则两套晶格中有某些格点在一定误差范围内互相重合，这些重合的格点构成一个新的点阵，称为重合位置点阵。以体心立方晶体为例，若以〈110〉方向为转轴，转动70.5°，38.9°和50.5°时重合位置点阵有较高的密度，相应的重合位置点阵格点数与体心立方点阵格点数的比依次为1/3，1/9和1/11。当晶界面为重合位置点阵的最密排面时晶界面具有最好的匹配度，界面能也最低。由于种种因素的制约，实际晶界并不总能和重合位置点阵的最密排面相一致，但实际观测表明，即使曲面状的大角度晶界也常由一些小的低能量界面组合而成。



(a) 两晶粒的拼合　(b) 晶格弛豫后形成一列刃型位错

图1.2-34　简立方晶格中小角度倾侧晶界的结构

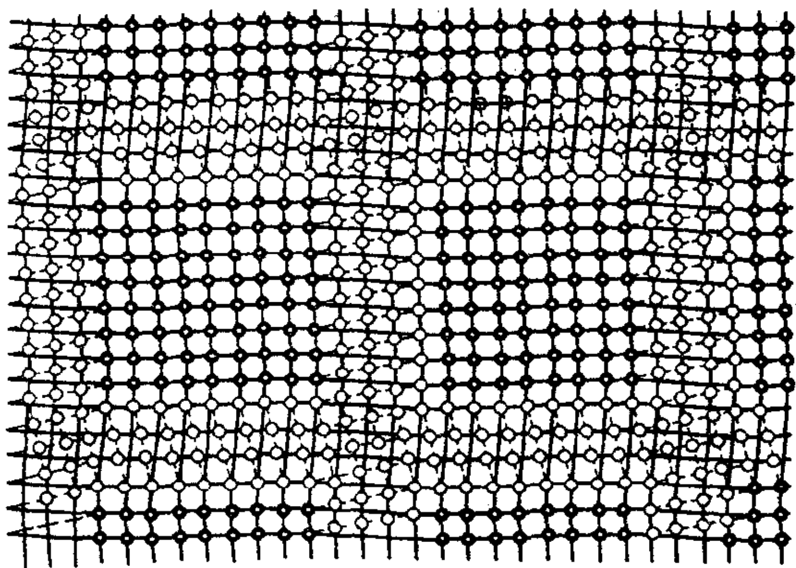


图1.2-35　简立方晶格中一个以（001）面为界面的小角度扭转晶界的结构，晶格弛豫后形成螺型位错网络

具有不同结构或不同组分的两相间的界面称为相界面，这里介绍的主要是固相间的界面。相界面可由固态相变产生，也可借助各种薄膜制备技术人工构筑而成。理解相界面的思路和晶界有颇多类似的地方，惟需注意两相的不同结构增加了相界面的复杂性。由于人工微结构材料日益增加的重要性，而外延膜和多层膜又是最重要的人工微结构材料，我们简单介绍一些相关的界面问题。相界面完全有序，两相晶格完全匹配者称为共格相界；界面处晶格常数的差异通过弛豫使错配局限于失配位错处，其余大部分区域仅有很小的弹性畸变者称为半共格相界；因晶格常数差别太大或其他原因形成的完全无序的界面称为非共格相界。薄膜在单晶衬底上外延生长时二者之间的界面往往是半共格相界。一般说来衬底和外延层的晶格常数 a_s 和 a_o 略有差异，其失配度 f 定义为：

$$f = (a_s - a_o)/a_s \qquad (1.2\text{-}16)$$

在外延生长初期膜很薄，这种失配可被弹性变形容纳，界面保持共格。随着外延层的增厚，弹性畸变能增加，当厚度超过某一临界值后，界面处出现周期排列的位错使膜的弹性畸变松弛从能量上考虑更为有利，这种周期排列的位错即为失配位错，如图1.2-36所示。相邻失配位错间的距离 D 约为：

$$D = a_o/f \qquad (1.2\text{-}17)$$

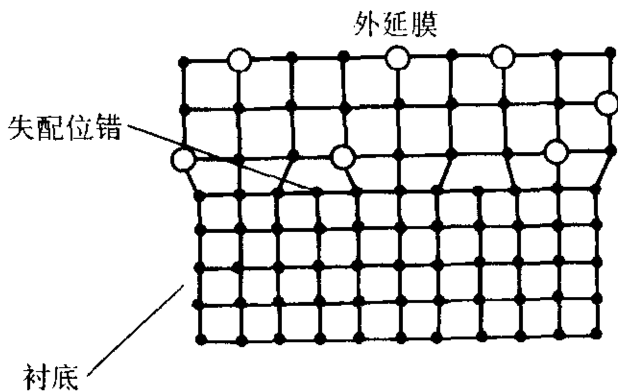


图1.2-36　薄膜外延生长形成的半共格相界

1.4　聚合物的典型结构

聚合物由许多长链大分子聚合而成，因而其结构有两个层次，即单个高分子链结构和许多高分子链聚集在一道的聚集态结构。单个高分子链结构又包括近程结构和远程结构，前者指单体的化学组成、键连方式和立体构形等，而后者指整个高分子链的形态和结构。

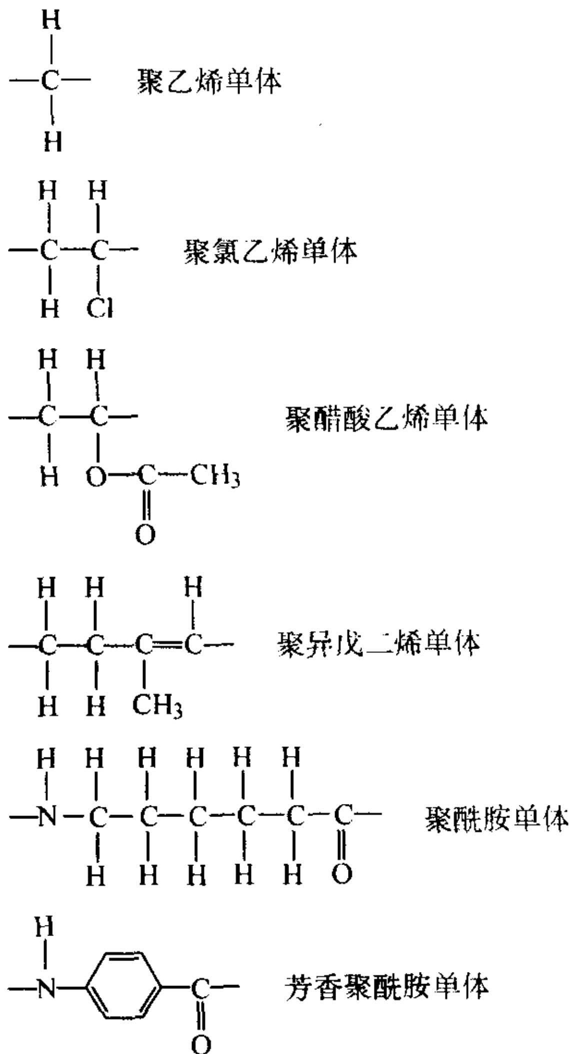


图1.2-37　一些形成聚合物的单体的结构式

先简单介绍高分子链的近程结构。构成高分子链的结构单元是单体，一条高分子链通常由 $10^2 \sim 10^5$ 个单体组成。图1.2-37给出了一些高分子单体的结构式。有些高分子链是由一种单体简单重复来构成的，也有的高分子链中单体的结构和类型不一定完全相同，这时单体的排列方式千变万化，往往会出现许多种不同的变体。共聚物是由两种或多种单体按照一定的排列方式聚合而成的。按照单体的排列方式，共聚物一般可分为无规共聚物，嵌段共聚物，接枝共聚物等，如图1.2-38所示。高分子链的远程结构通常涉及高分子链的

长短，刚度和链内旋转。高分子链大体可以分为三类，即柔性链，刚性链和螺旋链。在主链中含有苯环或杂环的高分子一般是呈棒状的刚性链。而像聚乙烯这样的高分子是柔性链，看上去像一个线团。蛋白质中的α-螺旋及核酸中的DNA则有较稳定的螺旋结构。

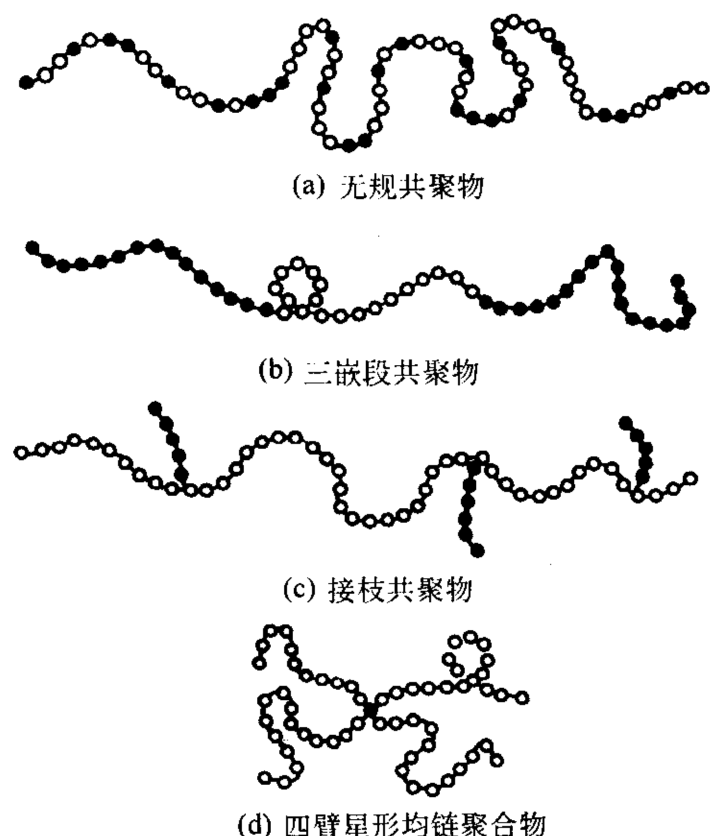


图 1.2-38　一些不同类型的聚合物

高分子聚集态结构又称作超分子结构，描述聚合物分子链的排列和堆积结构，一般分为四类，即非晶态，取向态，液晶态和晶态。关于非晶态聚合物的结构，P.T.Flory 在布朗运动粒子轨迹的启发下提出了"无规线团模型"，该模型认为，非晶态聚合物分子链的构象与在溶液中的分子链完全一样。如将聚合物分子链分为许多段，每段对应于布朗运动粒子的一步，这样聚合物链在空间任意地、不断地改变方向，成为一个无规线团。不同的线团又互相缠结，便形成了非晶态聚合物的结构。这一结构模型受到许多试验事实的支持。聚合物的取向态则是指在外力作用下分子链沿外力方向平行排列。未取向的高分子材料是各向同性的，而经过取向的材料则呈现各向异性。关于聚合物的液晶态，我们在1.2.4中已作过介绍，现不再赘述。聚合物的结晶形态往往因结晶条件而异，可分为折叠链晶体，伸展链晶体，纤维状晶和串晶四类。聚合物的结晶要在某个温度范围才能发生，但没有精确的熔点。结晶使分子链的排列有序化，可使聚合物的强度、密度、耐热性和耐化学腐蚀性等得到提高。

2　材料的相变

吉布斯首先提出物相的概念，它表示物质系统中具有相同化学组成，聚集状态及相同物理和化学性质的均匀物质部分。每种材料都是由一个或若干个相构成的。相变是指当外界约束条件（温度，压强，外场等）作连续变化并达到某定值时，物相却发生了突变，这种突变可以体现为从一种结构变为另一种结构，化学成分的不连续变化，某种物理性质的跃变或以上三者或其中二者的组合。它是控制材料结构和获得所需性质的重要方法。本节将对几种相变的起因和机制作简单的介绍。

2.1　相变热力学简介

我们首先从平衡态热力学出发考虑相变问题，取温度和压强为独立变量，对于流体系统吉布斯自由能 G 可以表示为：

$$G = G(T,\ p) = H - TS = U - TS + pV \quad (1.2\text{-}18)$$

式中，T 为热力学温度；p 为压强；H 为焓；S 为熵；U 为内能；V 为体积。对于固体，G 的表达式中还应考虑应力与应变的贡献。如存在外场，还要增加相应的项。对于可逆过程，吉布斯自由能的微分可记为：

$$dG = -SdT + Vdp \quad (1.2\text{-}19)$$

由此可以看出，熵和体积都是吉布斯自由能的一阶偏导数，即：

$$S = -(\partial G/\partial T)_p \quad (1.2\text{-}20)$$

$$V = (\partial G/\partial p)_T \quad (1.2\text{-}21)$$

由热力学第二定律可知：

$$(dG)_{T,p} \leqslant 0 \quad (1.2\text{-}22)$$

这表示在定温定压条件下吉布斯自由能不可能增加，也就是说，吉布斯自由能是定温定压条件下的热力学势。这时如果Ⅰ相的吉布斯自由能较Ⅱ相为高，那么，从Ⅰ相转变到Ⅱ相是可能的，而反过来的转变是不可能的。换言之，Ⅱ相是稳定相，而Ⅰ相和Ⅱ相吉布斯自由能的差值就是相变的驱动力。由吉布斯自由能表达式出发，可以引入三个响应函数，它们是比定压热容 c_p，等温压缩率 K_T 和热膨胀率 α，它们的表达式分别如下：

$$c_p = T(\partial S/\partial T)_p = -T(\partial^2 G/\partial T^2)_p \quad (1.2\text{-}23)$$

$$K_T = -(1/V)(\partial V/\partial p)_T = -(1/V)(\partial^2 G/\partial p^2)_T \quad (1.2\text{-}24)$$

$$\alpha = (1/V)(\partial V/\partial T)_p = -(1/V)(\partial S/\partial p)_T \quad (1.2\text{-}25)$$

在相变点上，两相的热力学势（这里为吉布斯自由能）相等，即系统的热力学势保持连续。但是作为热力学势的各阶导数（如熵、体积、比热容、压缩率等）却可能发生不连续的跃变。P.Ehrenfest 提出将热力学势对约束变量逐级求导，n 阶相变被定义为，在相变点，系统的热力学势的第 $(n-1)$ 阶导数保持连续，而其 n 阶导数则是不连续的。系统的热力学势对某些约束变量的一阶导数在相变点出现不连续性的相变称为一阶相变。若只考虑温度和压强两个约束变量的系统，在恒压下，吉布斯自由能对温度的一阶偏导数 $(\partial G/\partial T)_p = -S$，在恒温下，吉布斯自由能对压强的一阶偏导数 $(\partial G/\partial p)_T = V$，$S$ 和 V 分别代表熵和体积，发生一阶相变时，系统的熵和体积变化是不连续的，$\Delta S = \Delta H/T$，这意味着相变潜热和体积效应。一阶相变往往伴随着较大的热滞。如果在相变点系统热力学势的一阶导数保持连续，而二阶导数不连续，则此相变为二阶相变。由式（1.2-23）和式（1.2-24）可知，这时系统的比定压热容 c_p 和等温压缩率 K_T 发生跃变。类似的考虑适用于高阶相变。

2.2　凝固与熔化

2.2.1　凝固

材料由液态转变为固态的过程称为凝固。液体具有流动性，其特征是结构的无序性和原子位置的非定域性。严格意义上的固体即晶态固体在结构上是有序的，原子通常只停留在晶格座位附近，具有定域性。大量的凝固现象涉及液态转变为晶态固体的过程，在这一过程中材料结构的有序性和原子的定域性同时出现。另一类特殊的凝固现象是液体转变为非晶固体即玻璃态，它对应于液体原子非定域性的消失，而液体结构的无序性得以保留，即结构无序的液体变成了结构无序的固体。本节中我们先以液态转变为晶态固体的过程为主对凝固过程作一描述，然后对非晶态形成的一些特殊问题进行简单的讨论。

当温度下降至熔点以下，固相的自由能变得低于液相，凝固过程随之开始。由于液相的温度涨落和结构涨落可形成一些固相的晶胚。在晶胚的尺寸很小时它们的比表面积很大，过大的界面能使它们不能稳定地存在。但当其大小超过某一临界尺寸（晶核）时，就能稳定地长大成新相。成核的

驱动力是新相与母相的体自由能差，与母相的过冷度成正比，阻力是新相与母相的界面自由能。成核可以分为均匀成核和非均匀成核。这里的均匀是指新相在母相内各点出现的几率是相等的。下面以均匀成核为例介绍成核的动力学过程。考虑在液相中形成了一个半径为 r 的球形晶胚，它的出现导致的自由能变化为：

$$\Delta g(r) = -(4/3)\pi r^2 \Delta G_{LS} + 4\pi r^2 \sigma \quad (1.2\text{-}26)$$

式中，ΔG_{LS} 为生成单位体积固相时吉布斯自由能下降的绝对值，它与系统的过冷度成正比；σ 为单位面积界面能。上式中三项分别随 r 变化的情况示于图1.2-39。不难求出

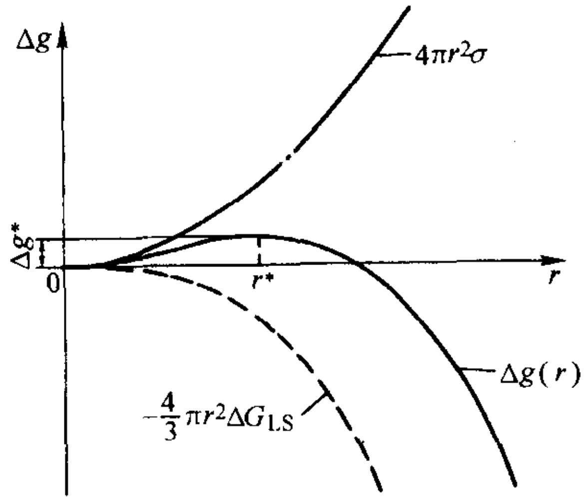


图1.2-39 成核过程中吉布斯自由能变化与其半径的关系

$\Delta g(r)$ 有一极大值，相应的临界半径 r^* 和 $\Delta g(r)$ 的极大值 Δg^* 分别为：

$$r^* = 2\sigma / \Delta G_{LS} \quad (1.2\text{-}27)$$

$$\Delta g^* = (16/3)\pi\sigma^3 / (\Delta G_{LS})^2 = (4/3)\pi\sigma r^{*2} \quad (1.2\text{-}28)$$

显然，大于临界半径 r^* 的晶胚（称作晶核）可以稳定地长大，晶核的形成必须跨越高度为 Δg^* 的势垒（称作成核功），这个过程叫作成核。成核是一个热激活的过程，单位时间内单位体积形成新相的晶核数称为成核率 N，它可表示为：

$$N = A\exp(-\Delta g^* / k_B T)\exp(-Q/k_B T) \quad (1.2\text{-}29)$$

式中，A 为常量；k_B 为玻耳兹曼常数；T 为温度；Q 为液相中原子扩散激活能。式中前一个指数项反映临界尺寸晶核的密度，后一个指数项反映液相中的原子扩散进入临界尺寸晶核使其最终达到稳定态的几率。实际的成核过程中晶相往往在模壁和杂质表面处优先成核，这叫做非均匀成核。非均匀成核的成核功更小，其理论分析与均匀成核差别不大，可参阅有关专著。

晶核形成之后便开始长大，这就是晶体在液相中的生长。晶体生长实际上是一个热量、质量和动量的输运过程。单一元素材料的凝固不涉及元素的重新分布，因而情况较为简单，只需考虑固液界面附近原子的短程迁移。界面有两种类型，即粗糙界面和光滑界面。大多数金属固液界面为粗糙界面，特点是界面处有一半左右的位置被空位占据，在微观尺度上是粗糙的，界面由几个原子厚的过渡层组成，如图1.2-40a所示。与此相反，光滑界面上只有极少的空位，保持晶体低指数面的光滑特征。对于粗糙界面，因为大量空位都可安插原子，不存在坐位不够的问题，晶体可连续向前生长。其生长速度主要受限于传热条件，与界面处的过冷度成正比。对于光滑界面，晶体生长是原子在晶面上一层层逐次填补的过程。光滑界面又可分为两种：邻位面和奇异面，前者表面法线方向与低指数晶面略有偏差，上面有许多台阶，如图1.2-40b所示，这时晶体可借台阶向前推进而生长，一般非金属晶体多为这种生长方式。奇异面则是很光滑的低指数晶面，需靠二维成核的方式产生台阶来生长。所谓二维成核是指光滑界面上在热涨落的帮助下形成长、宽可达到一定临界尺寸的凸台，然后靠台阶向前推进扫过整个界面而使该界面生长一个原子层，如图1.2-40c所示。二维成核的原理与晶相的成核过程颇有相似之处。除此之外，光滑界面可借晶体缺陷生长，当晶体内部存在结构缺陷如位错等时，常会在晶体表面有所反映，例如螺型位错的露头会在晶体表面形成永远填不满的小台阶，这时晶体生长会在表面形成螺旋状的生长丘，如图1.2-40d所示。

晶体生长的一个重要问题是生长界面的稳定性，它与界面附近温度分布密切相关。固液界面前的熔体中，温度的分布有三种形式。一种是熔体一侧离界面越远温度越高，如图1.2-41a所示。这时如果平界面出现一些凸起，那么凸起前端温度较高而过冷度较小，这使凸起受到抑制，界面保持平直和稳定。第二种是熔体一侧离界面越远温度越低，熔体为过冷熔体，如图1.2-41b所示。这时如果平界面出现一些凸起，那么凸起前端温度较低而过冷度较大，这使凸起加快向前生长，原先平坦的界面失去稳定性，越长越大的凸起还会生长出许多侧枝，这就是枝晶生长。第三种是熔体一侧界面前方温度先降后升，如图1.2-41c所示。这时由于界面前有狭窄过冷区，平界面上出现的凸起可以长期存在。但因为远处温度逐渐升高，它们不能无限制地发展，于是这种状态得以稳定下来，平界面变成了胞状界面，并产生了胞状组织。

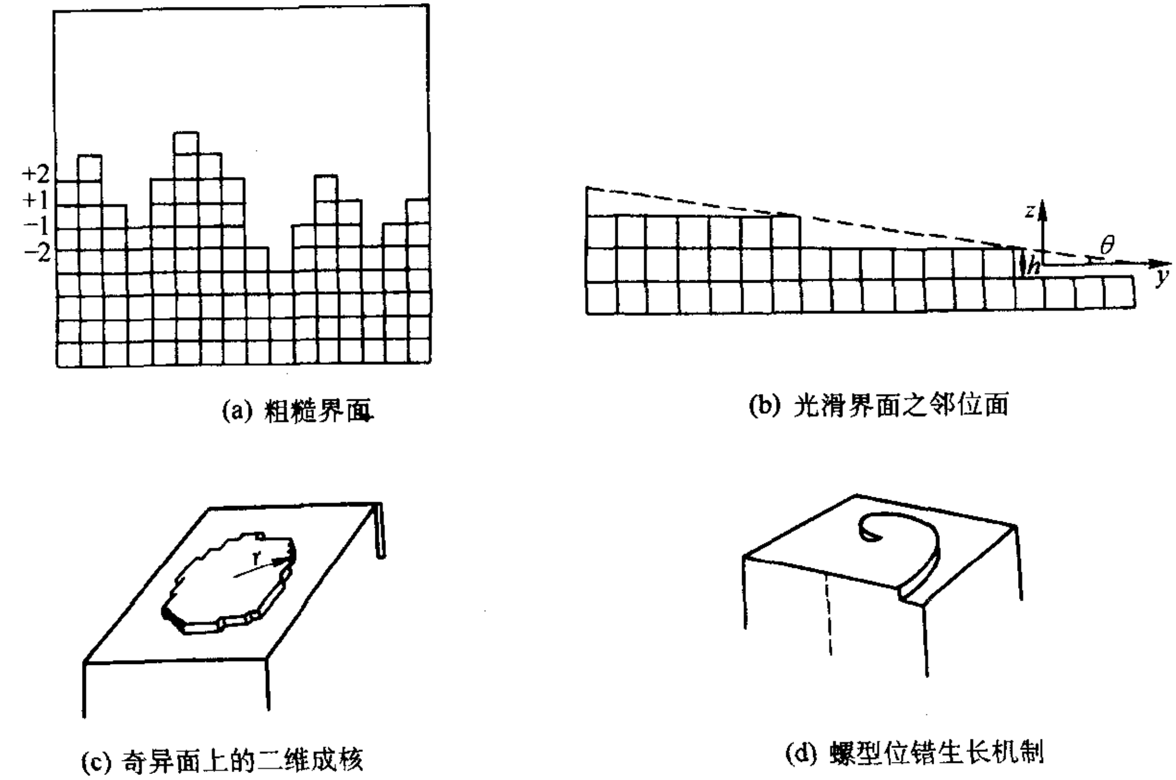


图1.2-40 界面结构与晶体生长机制

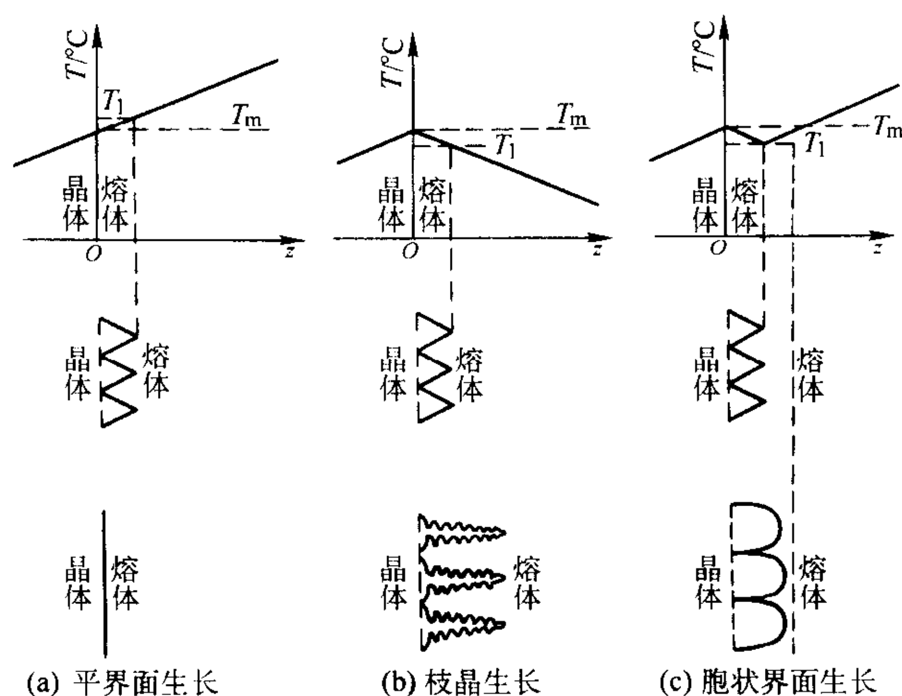


图 1.2-41　界面前方温度分布对界面稳定性和凝固组织的影响

对于两组元或多组元系统的凝固，溶质原子在液相和固相中的扩散对凝固过程和凝固组织会产生重要影响。如果固相为单相固溶体，由于溶质在液相和固相有不同的浓度，凝固过程将引起溶质的重新分布，导致晶体中溶质的偏析。如果凝固组织中有两相或多相，则称为共晶凝固。关于这些问题可参阅有关专著。

非晶态的形成与结晶动力学特别是晶态的成核有密切关系，其实质问题是如何加快冷却速度和选择材料成分以抑制晶态的出现。为了获得非晶态金属即金属玻璃，人们发展了多种快速冷却的技术。从结晶热力学和动力学的考虑出发，为了能在较低的冷却速度下获得非晶态，可按以下途径寻找解决办法：

1）增大晶态成核势垒 Δg^*，为此要选择结晶潜热较小和固液界面能较大的材料，同时尽可能将晶体中易导致非均匀成核的杂质消除干净。

2）熔体在熔点附近黏度较大，且随过冷度的加大而增加，那么原子由熔体进入晶态较为困难。晶态成核较慢，较易于形成非晶态。

3）如果结晶时要求原子作较大程度的重新分布，例如共晶成分的熔体就是如此，受扩散因素的限制，晶态成核减慢，有利于获得非晶态。

4）一般说来熔点较低的材料较易于形成非晶态，因为低熔点材料的熔化潜热较小，这意味着熔体转变为晶态的驱动力较小，因而易于形成非晶态。

2.2.2　熔化

物质由固相转变为液相的过程称为熔化。对于高温下的晶体，尽管存在不规则的热运动，原子仍然停留在晶格的格点附近。而对于液态，晶格已经不复存在，原子的位置分布是无规则的，而且原子并不定域于某个特定的坐位上，体现了液态结构中原子位置的无序性和非定域性。林德曼（F. A. Lindemann）提出了晶体熔化的判据：当原子热振动振幅 u 的均方根位移与原子间距之比值超过一定限度 δ 后，晶体即产生熔化，即：$(\langle u^2\rangle)^{1/2}/a > \delta$，$\delta$ 值因晶体结构和材料类型而有差异。对于体心立方金属 Li、Na、K 等，$\delta \approx 0.11$；对于面心立方金属 Cu、Ag、Au、Al 等，$\delta \approx 0.07$。在固相到液相的过程中，必须吸收熔化潜热，因而熔化是一级相变，在这一过程中体积及密度和熵都发生不连续变化。

2.3　固体中的扩散

固体中的扩散是指晶体中原子脱离它原来的平衡位置跃迁到另一平衡位置从而引起物质迁移的现象。从产生扩散的原因来看，可主要分为两大类，一类称为化学扩散，它是扩散物质在化学势梯度的推动下产生的扩散。一般情况下，化学势梯度的方向与浓度梯度的方向一致，原子从浓度高的区域向浓度低的区域扩散，这种扩散又叫做顺扩散；在某些发生相分离的系统中也可能发生“上坡扩散”的情况，这时化学势梯度的方向与浓度梯度的方向相反，表现为原子会从浓度低的区域流向浓度高的区域，导致相分离，这种扩散又叫做逆扩散。另一类扩散称为自扩散，它是在没有化学浓度梯度的情况下，由于空位和间隙原子的运动而产生的扩散，一般只有通过放射性同位素作为示踪原子才能观察到。

菲克定律是描述物质体系宏观顺扩散的基本定律。它是一种唯象理论，把扩散系统当做连续介质。菲克第一定律揭示了稳态条件下扩散通量与扩散系数和浓度梯度成正比的关系，其表达式为：

$$J = -D\mathrm{d}c/\mathrm{d}x \qquad (1.2\text{-}30)$$

式中，J、D 和 $\mathrm{d}c/\mathrm{d}x$ 分别为扩散通量、扩散系数和浓度梯度，负号表示扩散是沿着物质浓质降低的方向进行的。适用于恒温恒压下浓度沿 x 方向的单向一维流动，但可推广到二维和三维的情况。当扩散系数与浓度无关时，可以得到菲克第二定律：

$$\mathrm{d}c/\mathrm{d}t = D\mathrm{d}^2 c/\mathrm{d}x^2 \qquad (1.2\text{-}31)$$

式中，$\mathrm{d}c/\mathrm{d}t$ 为浓度随时间的变化率。菲克定律从本质上反映了物质系统通过扩散而趋于一致的趋势，适用于固态、液态以及气态中原子的扩散行为。

扩散的微观理论则要分析扩散原子在晶格中具体的跃迁机制，主要有空位机制，间隙原子机制和交换机制三种。空位机制是这样的：在有限温度下晶格中总存在一定浓度的空位，与空位相邻的原子可能跳进空位，使其原占据的格点变成一个新的空位，新空位的近邻原子又可跳入这个新空位使它继续运动，如此反复，使原子在晶格中迁移。由于原子跳进空位并不引起所经路径附近原子的大的位移，这种迁移比较容易实现，但受到晶格中空位浓度的限制。间隙原子机制是指扩散原子在晶格间隙位置之间跳跃而导致的扩散，间隙原子可来自弗仑克尔缺陷，但更多来自间隙式固溶体中的间隙杂质原子，铁素体中的碳原子是最常见的例子。由于间隙杂质原子比溶剂原子小得多，跃迁所需能量较小，所以扩散率比较大。交换机制则是指占据格点的相邻原子直接对调位置。由于原子对调过程中附近的原子必须退让出足够的空间，瞬间的弹性畸变是很大的，这种机制所需的激活能很大，出现的机会较少。但在实际材料，例如金在锗中的扩散过程中，已观察到这种机制的例子。除了以上三种机制之外，人们还提出和证实了环形机制，推填子机制，挤列机制，空位对机制，填隙 - 空位机制等，但它们实际上都是上述三种机制的组合或推广。

原子的扩散实际上是原子在热激活的帮助下跨越势垒由一个平衡位置到另一个平衡位置的过程，温度对扩散系数 D 的影响可表示为：

$$D = D_0 \exp\left(-Q/k_B T\right) \qquad (1.2\text{-}32)$$

式中，D_0 称为扩散常数；Q 称为扩散激活能；k_B 为玻耳兹曼常数；T 为热力学温度。对纯金属自扩散激活能 Q 近似地正比于该金属的熔点 T_m，一般有着简单的经验关系：$Q = 34T_m$。

2.4　固态相变

2.4.1　相变的分类

相变是一大类复杂的现象，我们可以从不同的侧面对相变进行描述和分类，而分类过程本身可以加深我们对相变的理解。

从热力学的角度考虑，相变可分为一阶相变，二阶相变和高阶相变。n 阶相变的定义是：在相变点，系统的热力学

势的第（$n-1$）阶导数保持连续，而其 n 阶导数则是不连续的。我们在2.1中已作过介绍。毫无疑问，热力学分析抓住了问题的本质，但它却无法清楚地给出相变具体过程的图像。

从相变动力学的角度考虑，相变可分为非匀相转变和匀相转变。吉布斯将母相中导致新相形成的结构涨落分为两类，第一类是空间范围很小而程度很大的涨落，第二类是空间范围很大而程度很小的涨落。前者，例如成核生长型固态相变，早已为人们熟知。后者，例如失稳分解，则较晚才被人们认识。我们将在下一节对它们作简单的比较和介绍。

相变也可按其结构变化来进行分类。布厄格（M. J. Buerger）概括了大量晶体在相变中结构变化的实验事实，提出相变可以区分为两种基本类型，即位移型与重构型。位移型相变表现为相变前后原子近邻的拓扑关系保持不变，相变过程不涉及化学键的破坏，只是键的长度，键与键之间的角度发生变化。相变所对应的原子位移较小，结构变化的途径也明确可辨，新相与母相之间存在明确的晶体学位向关系。位移型相变所经历的势垒甚小，相变潜热很小或完全消失。马氏体相变即属于位移型相变。如果相变过程中母相的结构被拆散为许多小单元，这些小单元重新组合起来形成新相，即称为重构型相变。在重构型相变中涉及大量化学键的破坏，在重新组合后，新相与母相之间在晶体学上没有明确的位向关系，原子近邻的拓扑关系也要产生显著的变化，这类相变经历了很高的势垒，相变潜热很大，因而相变的进程相当迟缓。

2.4.2 沉淀反应

外界条件的变化例如温度的下降可使固溶体变得过饱和。过饱和固溶体是不稳定的，会析出富含溶质的相，称为沉淀相，这一过程叫做沉淀反应，它是一种重构型相变。沉淀相的出现和增加有成核生长型和失稳分解型两种机制，分别介绍如下。

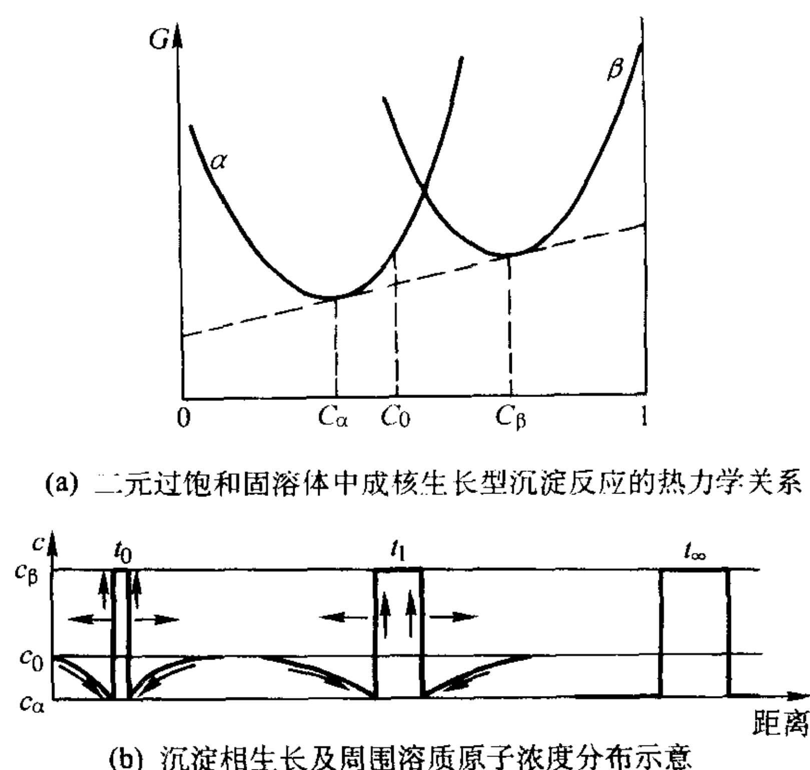
考虑一个二元过饱和固溶体，其母相 α 相和沉淀相 β 相的吉布斯自由能与成分的关系如图1.2-42a所示。设B组元在过饱和母相中的初始浓度为 c_0。因为在 c_0 处有：

$$\partial^2 G/\partial c^2 > 0 \qquad (1.2\text{-}33)$$

母相中程度甚小的成分涨落将导致吉布斯自由能的增加，因而系统是稳定的。但如果过饱和母相分解为B组元浓度分别为 c_α 和 c_β 的母相和沉淀相，则由于系统的吉布斯自由能将下降而变得不稳定。由于 c_β 和 c_0 差别很大，沉淀相的出现只能依赖第一类吉布斯涨落，即空间范围很小而程度很大的涨落来实现，这就是β相成核的过程。沉淀相成核与2.2.1中介绍的熔体中晶相成核的过程类似，但是还要考虑两个新的因素：一是弹性畸变能的贡献，由于弹性畸变能使自由能增加，它的出现使沉淀反应需要更大的过冷度。二是沉淀相成核需要溶质原子的长程扩散，所以进行的缓慢得多。最后求出的成核率的表达式仍为式(1.2-29)，只是一些物理量的表达式有所不同。类似地，沉淀相实际的成核过程中新相也往往在杂质、表面、界面、位错等缺陷处优先成核，亦即非均匀成核。沉淀相成核之后，由于其中富含溶质元素，它的周围会形成一个溶质原子的贫化区和溶质原子的浓度梯度，这使得远处的溶质原子向沉淀颗粒扩散并提供溶质原子使沉淀颗粒长大，直至沉淀反应接近完成。沉淀相生长及周围溶质原子浓度分布示意见图1.2-42b。

如果固溶体成分位于其吉布斯自由能曲线拐点之内，即：

$$\partial^2 G/\partial c^2 < 0 \qquad (1.2\text{-}34)$$

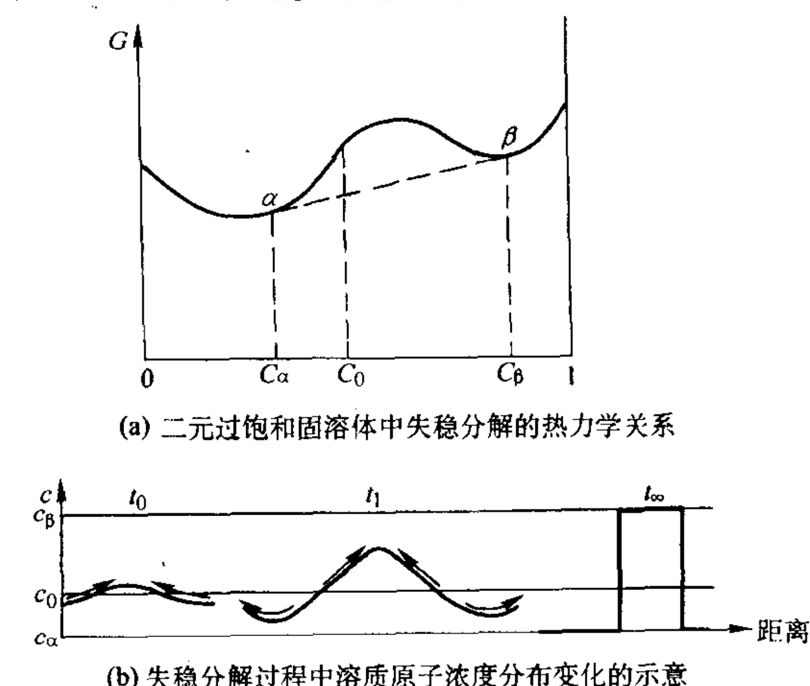


(a) 二元过饱和固溶体中成核生长型沉淀反应的热力学关系

(b) 沉淀相生长及周围溶质原子浓度分布示意

图 1.2-42

当固溶体内部发生微小的浓度起伏即可引起系统自由能的下降，其母相 α 相和沉淀相 β 相的吉布斯自由能与成分的关系如图1.2-43a所示。这使得这种偶然出现的成分偏离会自发



(a) 二元过饱和固溶体中失稳分解的热力学关系

(b) 失稳分解过程中溶质原子浓度分布变化的示意

图 1.2-43

增强而无需任何临界晶核，这意味着系统失去了稳定性。小的浓度起伏一旦出现，其幅度会借助逆扩散逐渐增大并迅速席卷整个系统，逐步完成沉淀反应过程直至达到两相的平衡浓度。这种机制称为失稳分解，其成分分布随时间变化的情况见图1.2-43b。它的特点是：

1）失稳分解属于吉布斯第二类涨落（大范围内程度甚小的涨落）所导致的相变。

2）失稳分解的初期两相之间没有明晰的界面，两相之间成分是连续过渡的，与明晰界面的界面能相对应，人们引入“梯度能”来表达浓度梯度所带来的吉布斯自由能的增加。失稳分解的后期两相分别逐渐达到其平衡浓度。

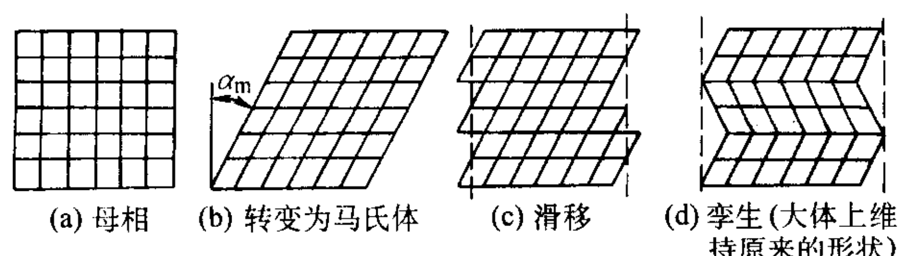
3）绝大多数发生失稳分解的固溶体母相与沉淀相有着结构上共格的关系。

4）由于晶体的弹性各向异性，失稳分解往往产生调制结构，受扩散的限制其周期较小，多在几十纳米的量级。

2.4.3 马氏体型相变

马氏体型相变是一种典型的位移型相变，以晶格切变为主而无成分变化。最初马氏体型相变是指钢中面心立方结构的奥氏体经快速冷却转变为体心四方结构的马氏体的相变，

它是最常见和最重要的马氏体相变。后来人们把具有同样特征的相变都称为马氏体型相变。马氏体型相变有以下明显特征：相变过程中无明显的组元扩散或成分变化，近邻原子间的相对位移不超过一个原子间距；相变引起晶格畸变，相变前抛光的表面上在相变后会出现浮凸；马氏体相和母相之间有严格的位向关系；马氏体往往在母相的特定晶面上生成，这类特定晶面构成了母相与马氏体的界面并被称为“惯态面”，它们在相变过程中不发生应变，也不发生转动；在马氏体内部通常含有亚结构；马氏体相变是可逆的，但逆相变有一定的热滞；马氏体相变速率极高，常可达声速量级。图 1.2-44 是晶粒的一部分发生马氏体型相变并形成亚结构的示意，母相的一部分发生切变成为马氏体结构，但为了与周围的晶粒互相协调，它必须通过滑移或孪生让自己的轮廓没有太大的改变，这便形成了亚结构。



(a) 母相 (b) 转变为马氏体 (c) 滑移 (d) 孪生(大体上维持原来的形状)

图 1.2-44 马氏体转变及其亚结构形成的示意

通常人们用马氏体转变前后保持平行的两个对应的晶面和这两个对应的晶面中各一个马氏体转变前后保持平行的晶向来描述马氏体型相变的几何关系。金属中常见的马氏体型相变例如：碳钢由面心立方到体心四方转变满足库尔久莫夫 - 萨克斯关系，即 $\{111\}_P // \{011\}_M$，$\langle 01\bar{1}\rangle_P // \langle 1\bar{1}\bar{1}\rangle_M$。铁镍合金在低温下由面心立方到体心四方转变满足西山关系，即 $\{111\}_P // \{011\}_M$，$\langle 11\bar{2}\rangle_P // \langle 0\bar{1}\rangle_M$。其中脚标 P 表示母相，脚标 M 表示马氏体相。马氏体型相变不仅可以发生在金属中，也可以发生在非金属材料之中。

3 材料的力学性质

材料的力学性质指材料在外应力作用下的行为，包括弹性，屈服强度，硬度，断裂等。材料的力学性质是结构材料的基础，它和材料的其他物性也有着密切的关系。

3.1 弹性

材料在外力作用下发生变形并在外力去除后恢复原状的性质称为弹性。弹性直接源自材料原子间的交互作用。以 Lennard-Jones 势式（1.2-1）为例讨论弹性的由来，该势能曲线见图 1.2-45a。当材料受压应力时原子互相排斥，当材料受拉应力时原子互相吸引，应力大小为：

$$F(r) = \partial [U(r)] / \partial r \tag{1.2-35}$$

原子间距为 a_0 时势能最低，此时有：

$$U(a_0) = -A/a_0^6 + B/a_0^{12} = U_{min} \tag{1.2-36}$$

$$\partial [U(r)] / \partial r|_{r=a_0} = 6A/a_0^7 - 12B/a_0^{13} = 0 \tag{1.2-37}$$

利用这些关系，由应力（F/a_0^2）与应变（$r/a_0 - 1$）的比值可以求出此系统的弹性模量 E 为：

$$E = CU_{min}/a_0 \tag{1.2-38}$$

式中，C 为与 A，B 有关的常数。显然，势阱越深，晶格参数越小，则弹性模量越大。

在弹性形变范围内，材料的应力与应变成正比，其比值称为弹性模量，这一规律称为胡克定律。正应力 σ 与正应变 e 的比值称为弹性模量，用符号 E 表示并满足：

$$\sigma = Ee \tag{1.2-39}$$

切应力 τ 与切应变 γ 的比值称为切变模量，用符号 G

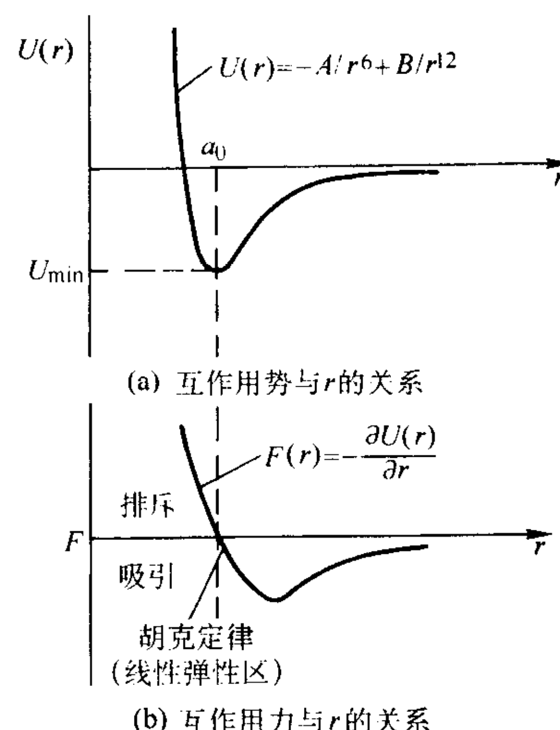


图 1.2-45 两原子交互作用的 Lennard-Jones 势

表示并满足：

$$\tau = G\gamma \tag{1.2-40}$$

静水压 p 与体积应变 $\Delta V/V$ 的绝对值之比称为体积模量，用符号 κ 表示并满足：

$$p = \kappa (\Delta V/V) \tag{1.2-41}$$

材料在正应力作用下沿应力方向伸长（或缩短），则同时在垂直于应力的方向缩短（或伸长），垂直于应力的方向上缩短的量与沿应力方向伸长的量之比称为泊松比，用符号 ν 表示。以上三种弹性模量和泊松比这四个量中只有两个是独立的，它们之间存在以下关系：

$$G = (E/2) / (1+\nu) \tag{1.2-42}$$

$$\kappa = E/3(1-2\nu) \tag{1.2-43}$$

弹性模量表征材料抗拒变形的能力，是材料性能的一种重要指标，多数无机材料的弹性限度在 0.2% 左右，超过这个限度便会发生范性变形。

有些材料例如橡胶显示出很高的弹性，其特点是弹性模量小而变形量可以达到很大。这类材料大多由长链状大分子构成。在不受外力的情况下分子链处于松塌状态，如 1.4 节的“无规线团模型”所描写。在受到张拉时分子链可在较大程度上被拉直，从而导致很大的应变。高分子材料还具有“黏弹性”。理想弹性体的应变与时间无关，而理想黏性体应变与应力作用的时间成线形关系。高分子材料介乎二者之间，这是由无规弯折的长链大分子在应力作用下弹性伸展和大分子之间相对滑移造成的黏性流动共同造成的。

3.2 屈服与范性

图 1.2-46 是材料样品拉伸试验中的应力-应变曲线。它分为两个区域，Ⅰ是弹性区，满足胡克定律，Ⅱ是范性变形区，随着应变的增加应力不呈线性增加甚至有所下降，应力松弛后材料也无法恢复原长度。划分Ⅰ区和Ⅱ区的应力就是材料的屈服强度，标志着范性变形的开始。有些材料没有明显的屈服点，工程上一般取永久变形值达 0.2% 时所对应的应力为该材料的屈服强度。屈服现象与材料中位错的形成，运动和增殖密切相关，关于位错的易动性我们在 1.3.2 节中已作介绍。金属单晶的拉伸试验表明，范性变形后的单晶表面总是出现许多台阶。这证实了范性变形的位错滑移机制。位错的滑移总是沿一定的滑移面和滑移方向发生的。图 1.2-47 画出了一个圆柱形六方单晶沿（0001）面滑移的情况。这里滑移面与样品横截面斜交，二者夹角为 φ，滑移面内的滑移方向与样品轴线之间的夹角为 χ。容易求出，在样品受轴向正应力 σ_0 作用时，滑移面内沿滑移方向的分切应力 τ 为：

$$\tau = \sigma_0 \cos\varphi \cos\chi \tag{1.2-44}$$

如果这一分切应力大到足够使该滑移系统中的位错运动和增殖，那么范性变形便会发生。一些重要晶体的滑移系统见表1.2-2。

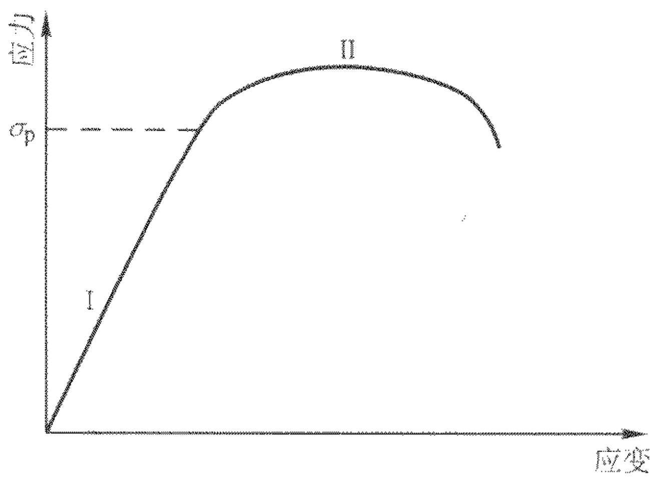


图1.2-46 典型的应力-应变曲线（图中σ_p为屈服强度）

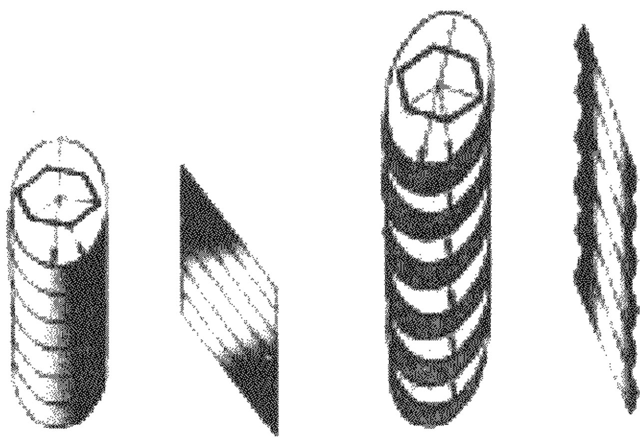


图1.2-47 圆柱形六方单晶沿（0001）面滑移的情况

表1.2-2 一些重要金属晶体和无机非金属晶体的滑移系统

材 料	点阵类型	温度/℃	滑移面	滑移方向
Al，Cu，Ag，Au，Ni，Pb	面心立方	20	{111}	〈110〉
Al		$\frac{T}{T_m}>0.72$	{100}	〈110〉
α-Fe，Mo，Nb	体心立方		{110} {112} {123}	〈111〉
α-Fe+4% Si			{110}	〈111〉
Zn，Cd，Mg	六角	20	(0001)	〈11$\bar{2}$0〉
Mg		$\frac{T}{T_m}>0.54$	{10$\bar{1}$1}	〈11$\bar{2}$0〉
Mg+14% Li		20	(0001) {10$\bar{1}$0}	〈11$\bar{2}$0〉
α-Ti		20	{10$\bar{1}$0} {10$\bar{1}$1}	〈11$\bar{2}$0〉
C（钻石），Si，Ge			{111}	〈1$\bar{1}$0〉
NaCl，LiF，MgO，NaF		低温	{110}	〈1$\bar{1}$0〉
NaCl，LiF，MgO，NaF		高温	{110} {001} {111}	〈1$\bar{1}$0〉 〈1$\bar{1}$0〉 〈1$\bar{1}$0〉
TiC		高温	{111}	〈1$\bar{1}$0〉
PbS，PbTe			{001} {110}	〈1$\bar{1}$0〉〈001〉
CaF_2，VO_2		低温	{001}	〈1$\bar{1}$0〉
CaF_2，VO_2		高温	{001} {110} {111}	〈1$\bar{1}$0〉〈1$\bar{1}$0〉
C（石墨），Al_2O_3，BeO			{0001}	〈11$\bar{2}$0〉
TiO_2			{101} {110}	〈10$\bar{1}$〉 〈001〉
$MgAl_2O_4$			{111}	〈1$\bar{1}$0〉

3.3 断裂

材料的断裂情况十分复杂，大体上可分为延性断裂和脆性断裂两大类。前者在断裂前试样已经历了大量的范性变形，后者表观上没有明显的范性变形迹象，只有仔细观察断口才能发现少量范性变形的痕迹。断裂实际上是裂纹形成和扩展的过程，我们简单地介绍格里菲斯（Greafith）的裂纹发

展理论。该理论认为：材料中有微小裂纹存在引起应力集中，使得断裂强度下降。对应于一定尺寸的裂纹，存在一临界应力 σ_c，当外应力小于 σ_c 时裂纹不能扩展，只有当外应力大于 σ_c 时裂纹才能扩展并导致断裂。

考虑一薄板状试样，中间有一长度为 $2a$ 的裂纹。设板受到与裂纹垂直的张应力 σ 的作用，可以认为裂纹周围直径为 $2a$ 的区域内应力被松弛。被松弛掉的弹性能为弹性能密度乘以面积，即：$-(\sigma^2/2E)\pi a^2$。更准确的计算求出的值为：

$$U_1 = -\pi a^2 \sigma^2 / E \qquad (1.2\text{-}45)$$

其中 E 为弹性模量。同时，裂纹增加的表面能为：

$$U_2 = 4a\gamma^* \qquad (1.2\text{-}46)$$

式中，γ^* 为单位长度裂纹的表面能。图 1.2-48 画出了 U_1，U_2 和 U_1+U_2 和裂纹长度的关系。显然，如果裂纹的长度超过了 U_1+U_2 的极大值所对应的长度，它便可以扩展，反之，则不能扩展。换言之，如果已存在一个宽度为 $2a$ 的裂纹，可以由 $d(U_1+U_2)/da=0$ 求出一个临界外应力 σ_c，并且有：

$$\sigma_c = (2E\gamma^*/\pi a)^{1/2} \qquad (1.2\text{-}47)$$

当外应力大于这一临界值时裂纹便可以扩展。这就是格里菲斯的裂纹发展理论。

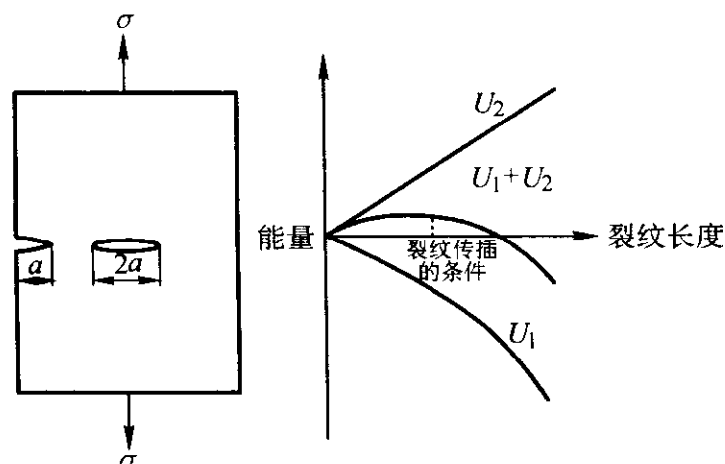


图 1.2-48 格里菲斯裂纹及其能量关系

3.4 高聚物的黏性流变

高聚物同时具有弹性和黏性，液体的流动规律是研究高聚物黏性的基础。理想的黏性液体在极小的切应力作用下即会开始流动，在流速不大时黏性液体的流动以层流的方式进行，层之间存在一定的速度梯度。为了维持这样的流动，层与层之间要保持一定的切应力。液体内反抗这一切应力的内摩擦力就反映了材料的黏性。单位面积液层上所受的切应力 σ 与层间的速度梯度成正比，即有：

$$\sigma = \eta \mathrm{d}v/\mathrm{d}x \qquad (1.2\text{-}48)$$

式中，$\mathrm{d}x$ 是两液层之间的距离；$\mathrm{d}v$ 是两层液体之间的速度差；η 为黏度系数。该式又称为牛顿流动定律，相应的液体叫作牛顿流体。从应力和应变的角度来看，牛顿流动定律可写为：

$$\sigma = \eta \mathrm{d}\gamma/\mathrm{d}t \qquad (1.2\text{-}49)$$

式中，γ 为切应变。该式可用来描述高聚物的黏性。高聚物的黏性和弹性均和温度有非常强的依赖关系。

4 材料的其他物理性质

4.1 基本概念

4.1.1 量子特征

20 世纪初之前人们认为粒子，例如电子及中子，和波，例如光波及其他电磁波，是完全不同的两回事。前者是定域的，一个一个的。而后者是广延的，连续的。但是在 20 世纪初人们相继发现，原本被认为是粒子的电子具有波动性，可以发生衍射，而光波的能量是一份一份的，即“光子”。这使得物理学家不得不重新审视粒子性和波动性，它导致了量子力学的诞生。

1924 年德布罗意（L. de Broglie）提出了电子的波粒二象性假说：一个动量为 p，能量为 E 的粒子与一个单色的波相联系，该波的波矢 k 和圆频率 ω 可以表示为：

$$k = p/h \qquad (1.2\text{-}50)$$

$$\omega = E/h \qquad (1.2\text{-}51)$$

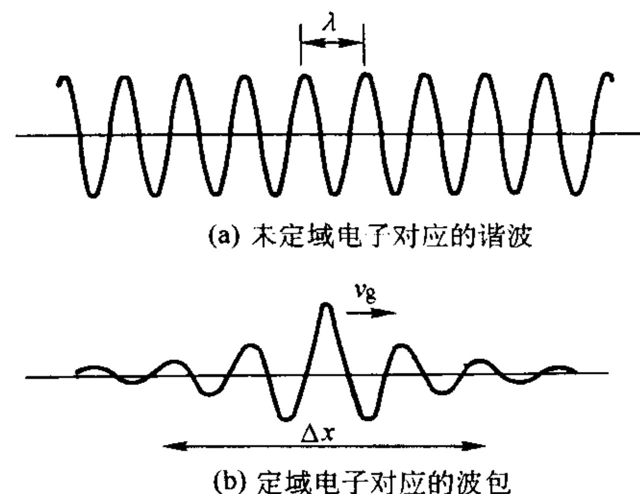
式中，h 为普朗克常数，其数值为 1.0544×10^{-34} J/s。这种波被称为德布罗意波，这个粒子的德布罗意波波长 λ 为：

$$\lambda = h/2\pi p \qquad (1.2\text{-}52)$$

根据这一假说，可以用一个物质波来描述一个粒子。自由粒子具有确定的动量和能量，对称性要求其德布罗意波振幅在整个空间都相同，是具有恒定振幅的谐波，如图 1.2-49a 所示。而一个束缚于空间区域 Δx 内的粒子对应于一个波包，如图 1.2-49b 所示。为了使波包在空间定域，它必须由不同波长的波叠加而成，以使得波的振幅在波包内很大，而在波包之外很快衰减为零。粒子的速度等于波包的群速度。如果波包扩展在 Δx 区域，则组成此波包的干涉波的波数的值必须处于 Δk 范围内，并且有 $\Delta x\Delta k \sim 2\pi$，由于 $\Delta p = h\Delta k$，故有：

$$\Delta x \Delta p \sim 2\pi h \qquad (1.2\text{-}53)$$

该式称为海森伯（Heisenberg）不确定原理，它反映了自然界的一个基本事实，即：一个粒子的位置和动量是不可能同时准确测知的，Δx 和 Δp 分别称为位置和动量的不确定度，Δx 越大，则 Δp 越小，反之亦然。



(a) 未定域电子对应的谐波

(b) 定域电子对应的波包

图 1.2-49

“波粒二象性假说”很难由人们的直觉来理解，这并不奇怪，因为普朗克常数非常小，所以在宏观世界人们不能直接感知量子特征。在材料科学中，通常人们将原子系统仍当作经典粒子系统，电子系统才表现出明显的量子特征。但是，量子力学效应产生的一些宏观量子效应，如超导电性，却是人们可以直接感知的。

4.1.2 固态电子论简介

固体中有大量电子。如果电子间的作用可以忽略，便是独立电子近似。固态电子论的中心问题是电子波在离子实组成的周期势场中的传播。

对于简单金属，可以采用最简单的自由电子近似。整块金属被当作势阱，价电子被看作在这个势阱中自由运动的电子气，类似于量子力学中电子处于三维方形势阱中的情况，这时有平面波解。若金属为边长 L 的立方体，利用周期性边界条件可求出波矢 k 的三分量只能为 $2\pi/L$ 的整数倍，即出现了分立的量子态。电子的能量为 $E = h^2k^2/2m$，其中 m 为电子质量，能量也相应地出现了量子化。如以 k 的三分量为坐标轴构筑 k 空间，电子的等能面为球面。考虑到由于泡利原理每个能态上只能有两个电子其自旋分别向上和向

下，所以电子按 k 由小到大的顺序填充各能态。由波矢和能量的关系可以得到能态密度 $D(E)$ 的表达式为：

$$D(E) = (V/2\pi^2)(2m/h^2)^{3/2}E^{1/2} \quad (1.2\text{-}54)$$

全部电子都填入之后，被占区域为一个球，称为费米球。球面称为费米面，其能量称为费米能。自由电子论可以满意的解释绝大多数金属的导电性。自由电子近似导出的抛物线型能态密度曲线见图 1.2-50。

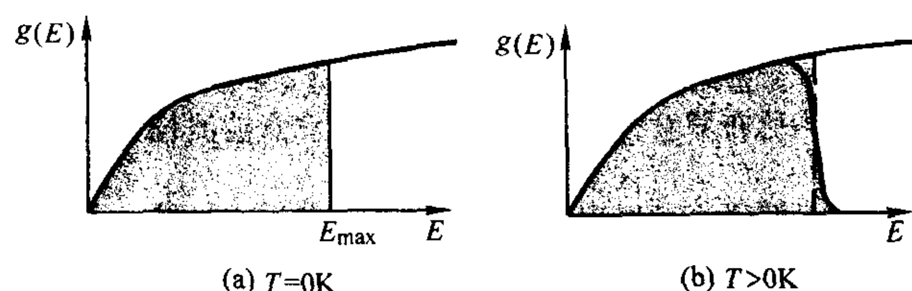


图 1.2-50 自由电子近似导出的抛物线型能态密度曲线

晶体中周期性排列的离子实对于晶体中的价电子来说是一个周期性的势场，自由电子近似完全忽略了它的存在，是过于简单了。对在周期势场中运动的大量电子进行量子力学处理就导致了能带理论。它指的是晶体中的电子态分布在某些限定的能量范围内，或者说这些电子态组成了互相分隔的若干能带。关于能带出现的物理原因我们可以从两个角度来理解。一方面，可以从原子能级和分子轨道的角度来考虑。当原子相距较大而孤立存在时，原子的每个电子都处在特定的能级上。当两原子间距缩小到晶体中相邻原子间距时，两原子同一能级的电子能态将发生重叠，而这是泡利原理不能允许的，于是该原子能级（又叫原子轨道）分裂为两个能级（又叫两个分子轨道）。对于 N 个原子的分子体系，每个原子能级将分裂为 N 个分子轨道。当 N 非常大时，由原来一个原子能级分裂出来的这些分子轨道就形成了一个接近连续的能带。能带之间是禁止被占据的禁带。12 个原子系统中能带形成的示意图见图 1.2-51。

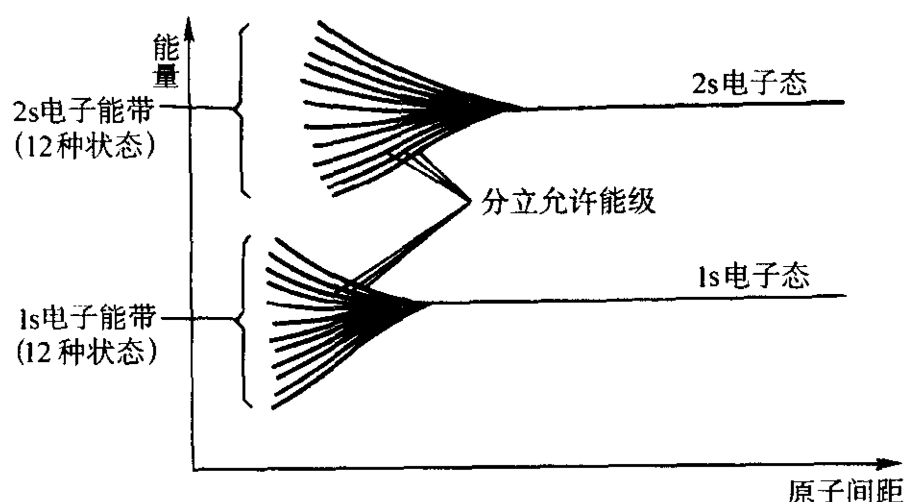


图 1.2-51 12 个原子系统中能带的形成

另一方面，也可以由自由电子近似出发来考虑能带的形成。自由电子近似中电子按 k 由小到大的顺序填充各能态，从某种意义上说，这就是一个大能带。考虑晶体中周期性排列的离子实产生的周期性的势场后，情况将发生变化。那些与周期性的势场满足布喇格反射条件的电子波将不能被激发，相应的那些能态也就不能被占据，这就是禁带。禁带的出现便造成了分立的能带。

能带论的巨大成功在于它对各类晶体显著不同的电学性质给出了简单的解释。金属、绝缘体和半导体的典型能带结构见图 1.2-52。金属中能量最高的能带，即价带只被电子部分充满，这种部分充满的价带又称为导带。在热激活下费米面附近的电子很容易被激发到能量稍高的能态而运动。这使得金属有良好的导电性和导热性。绝缘体的价带是全满的，而上面的能带是全空的，其间隔着很宽的禁带。由于价带被充满，其中的电子没有活动的余地，除非部分跃过禁带到达上面的能带。而因为禁带很宽，这种跃迁难于进行。这种能带结构造成了绝缘体的特征。半导体的能带结构与绝缘体类似，但价带上面的禁带很窄。热激活可以把少量价带顶部的电子激发跃入上面导带的底部并在价带顶部留下相同数量的空穴，这就是本征半导体。顾名思义，半导体的电导界于金属和绝缘体之间。

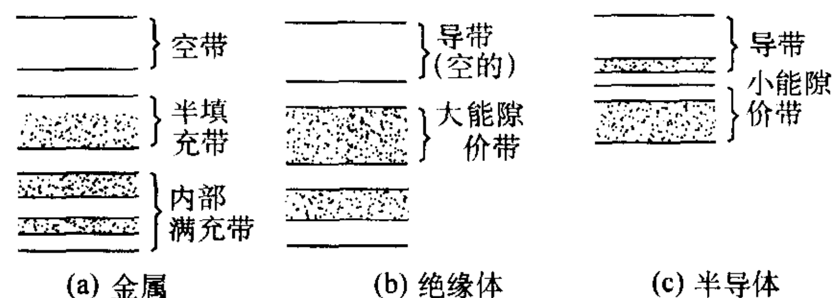


图 1.2-52 各类能带结构

4.1.3 晶格振动理论和声子的概念

在有限温度下组成晶格的原子或离子要发生热振动。由于邻近原子的交互作用，各原子的振动会发生耦合，致使这种振动以格波的形式在晶格中传播。由原子的交互作用势可知，原子间距较大时互相吸引，原子间距很小时互相排斥。在平衡位置附近，原子间作用力与偏离平衡位置的位移方向相反而大小成正比。这使我们可以用图 1.2-53 所示的模型来模拟晶格。考虑 N 个原子组成的一维的原子链，设原子质量为 M，弹簧力常数为 k，原子间距为 a，由量子力学并利用周期性边界条件可以求出该系统共有 N 支格波，或者说共有 N 种振动模式。每一种振动模式类似于量子力学中的一个一维谐振子，第 q 种振动模式的频率为 ω_q，其能量 ε_q 的表达式为：

$$\varepsilon_q = (n + 1/2)\ \hbar\omega_q \quad (1.2\text{-}55)$$

式中 $n = 0, 1, 2, 3\cdots\cdots$。人们将式中一份量子化的能量 $\hbar\omega_q$ 叫作一个“声子”，换言之，该一维系统中相应于 N 种振动模式共有 N 种声子，每种振动模式有许多能级，相邻能级的能量差为该种声子的能量。当电子或光子与晶格发生能量交换时，交换的能量只能以某种声子的能量为单位，称为“吸收一个声子”或“发射一个声子”等等。显然，声子与电子那样的物质粒子不同，它被称为“准粒子”，是晶格振动中一份量子化的能量。对于由 N 个原子组成的三维晶体，类似的计算可得出，系统共有 $3N$ 种振动模式，N 种为纵振动模式，$2N$ 种为横振动模式，每种振动模式能量的表达式仍为式（1.2-53）。这里介绍的晶格振动理论在长波极限的情况下与经典弹性波计算的结果相同，因为这时晶格可被近似看作连续介质。

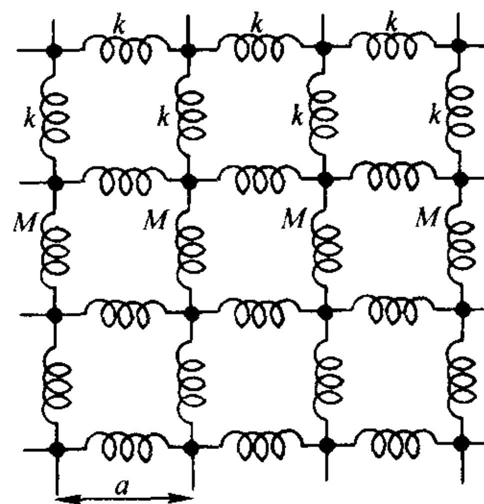


图 1.2-53 用于晶格振动计算的二维晶格模型

晶格振动理论的一个重要应用是成功地解释了固体比热容。爱因斯坦粗略地认为 N 个原子的振动是互相独立的，可视为 $3N$ 个一维谐振子并具有相同的频率 ω_E。这一理论成功地给出了固体在高温下的比热容，但计算得到的低温下的比热容与实验不符。德拜（Debye）具体考虑了 $3N$ 个一维谐振子频率的分布，成功地给出了固体在各种温度下的比热容。

4.2　材料的导电性

材料的电导是材料中的载流子在外场或外力的作用下定向迁移的过程，涉及载流子的种类，数量和它们在外场及外力作用下的行为。不包括超导体在内，各类材料电导率的差别可达 10^{25} 倍，其机制也各有不同。在此分别作简介如下。

4.2.1　金属导电性

前节已经介绍，金属电导是建立在自由电子论或固体能带理论基础上的。依据量子力学计算，在严格周期势场中运动的电子保持在一个本征态中，不会受到散射，因而也不会有电阻。然而这与实际金属中的情况不符，问题出在实际金属中晶格的严格周期性受到晶体缺陷和晶格振动的破坏。纯金属单晶在室温下电阻主要来自晶格原子振动对电子的散射。散射几率与晶格原子振动振幅的平方成正比。在较高温度下，散射几率和由此带来的电阻率因而与温度成正比。晶体缺陷造成的散射基本与温度无关，而是正比于缺陷密度，它们是较低温度下剩余电阻产生的主要原因。金属中电流的传输符合欧姆定律，即：

$$j = \sigma E \qquad (1.2\text{-}56)$$

式中，j 为电流密度；σ 为电导率；E 为电场强度。

4.2.2　半导体导电性

半导体的导电性主要取决于其中载流子的浓度和迁移率。半导体中的载流子分为两类，即电子和空穴，电子带有负电荷，而空穴带有正电荷。关于本征半导体中电子和空穴的形成机制，我们在4.1.2节中已做过介绍，其电子和空穴的密度相等而且随禁带宽度和绝对温度之比而呈指数下降。室温下本征半导体中的载流子浓度很低，例如在硅中只有 $10^{10}/cm^3$，其本征电导率很低。在半导体中产生电子和空穴的另一种方法是掺杂，例如用五价的磷代替硅中部分四价的硅原子，磷有五个价电子，除了用四个电子和周围的硅原子形成共价键外还多一个电子。该电子很容易进入导带，磷因而被称为施主杂质。而用三价的硼代替单晶硅中的部分硅原子，便可以掺入空穴，硼被称为受主杂质。掺有施主杂质的半导体叫做n型半导体，掺有受主杂质的半导体叫做p型半导体。掺杂半导体中载流子的类型和浓度取决于掺杂元素的种类和浓度。

载流子在外电场作用下将发生漂移，在同一电场作用下电子和空穴的运动方向相反。载流子漂移速度 v 与电场强度 E 比值的绝对值称为其迁移率 μ，即：

$$\mu = |v/E| \qquad (1.2\text{-}57)$$

考虑到半导体中有两种载流子，即电子和空穴，其电导率 σ 应包括二者的贡献，即：

$$\sigma = j/E = (n\mu_n + p\mu_p)\, q \qquad (1.2\text{-}58)$$

式中，n，p 分别为电子和空穴的浓度；μ_n，μ_p 分别为电子和空穴的迁移率；q 为电子电量。与金属有些类似，影响载流子迁移率的是各种散射作用，主要是晶格热振动的散射和晶格缺陷特别是杂质原子的散射。晶格热振动的散射随温度的上升而增强并导致迁移率的下降。

对于本征半导体，温度上升导致迁移率下降，但载流子浓度上升更快，因而电导率随温度上升，如图1.2-54所示。掺杂半导体中不仅有本征载流子而且有掺杂载流子，温度越高，前者贡献越大，掺杂浓度越高，后者贡献越大。掺杂半导体电导率与温度的关系随掺杂浓度的变化而呈现不同的规律。一种轻掺杂半导体和一种重掺杂半导体电导率与温度的关系见图1.2-55。

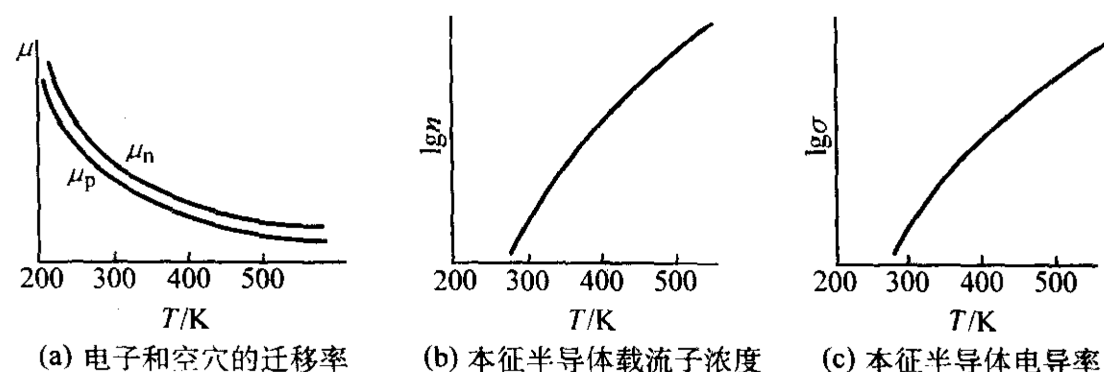


图1.2-54　几种输运参数随温度的变化

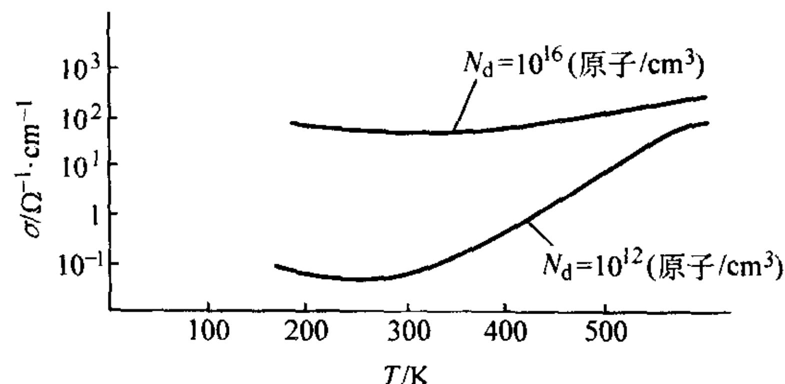


图1.2-55　一种轻掺杂半导体和一种重掺杂半导体电导率与温度的关系

4.2.3　离子电导

离子电导是材料中的离子在外电场作用下定向运动的结果，它与晶格缺陷密切相关，主要通过空位和填隙原子的迁移来实现。由于离子晶体中存在正负离子的空位，正负离子的填隙原子，各种杂质离子以及它们组合而成的缺陷对，其中空位和填隙原子的浓度又有平衡浓度与过饱和浓度之分，离子电导是这些缺陷贡献的叠加，应对各类缺陷的贡献分别进行计算。具体某一种缺陷对电流密度的贡献应是其浓度，所带电荷，沿电场方向移动速度三者的乘积。如果第 i 种缺陷的浓度为热平衡浓度，可以求出它对电导的贡献 σ_i 为：

$$\sigma_i = (N\nu_i a^2 q_i^2 / k_B T)\exp[-(g_i + U_i)/k_B T] \qquad (1.2\text{-}59)$$

式中，N 为单位体积内的离子数；ν_i 为该种缺陷的振动频率；a 为晶格常数；q_i 为该缺陷所带电量；k_B 为玻耳兹曼常数；T 为绝对温度；g_i 和 U_i 分别为该缺陷的迁移激活能和形成能。

对于大多数离子晶体，其室温下的离子电导率很低，是良好的绝缘体。但也存在一类“快离子导体”，即使在室温下也有很高的离子电导率（$>0.01\ \Omega^{-1}\cdot cm^{-1}$），可与液体电解质相比。快离子导体可按其结构特点分为两类；一类存在亚晶格无序，其晶格由两套结构组成，一套是由不易运动的离子构成的刚性亚晶格，另一套是可运动离子构成的高度无

序类似于液体的亚晶格。刚性亚晶格为可运动离子提供了大量能量相等的间隙位置，这些间隙位置少部分被可运动离子无序地占据，大部分空着。可运动离子在间隙位之间穿行的势垒很低。这类快离子导体中可运动离子的半径都比较小，如 Ag^+，Cu^+，Li^+，Na^+，F^-，O^{2-} 等。其典型材料如 AgI，$RbAg_4I_5$，CuI 等。另一类快离子导体结构中存在一系列松散的层或管道，使某些离子极易在其中运动，其例子是具有框架结构或钨青铜结构的一些晶体。

4.2.4 超导电性

材料在一临界转变温度 T_c 以下表现出零电阻和完全抗磁性的性质称为超导电性。其中前一种现象发现于 1911 年，后一种现象又称为迈斯纳（Meissner）效应，发现于 1933 年。直至 20 世纪 80 年代初，人们发现的超导体多为金属和金属间化合物，其最高的超导转变温度为 23.2 K，这些超导体被称为常规超导体。1986 年 G. Bednorz 和 K. A. Mueller 发现某些氧化物如 La-Ba-Cu-O 具有超导电性，而且转变温度高达 30 K 以上，被称为高温超导体。其后人们发现了一系列高温超导体，其最高转变温度已高达 135 K，而且似乎还未到顶。此外人们还发现了重费米子超导体和有机超导体，这三类超导体被称为非常规超导体。2001 年人们又发现了一种新的超导体 MgB_2，其转变温度为 40 K，但具有常规超导体的特征。超导材料已有重要应用，更有极其广阔的应用前景。

关于常规超导体超导电性的微观物理机制，巴丁（Bardeen），库珀（Cooper），施里弗（Schrieffer）在 1957 年提出了著名的 BCS 理论，其大意是：当温度降低到超导转变温度以下，由于电子－声子相互作用，费米面附近的电子可以相互吸引，形成库珀对。库珀对中电子的能量比正常费米能更低。费米面附近大量库珀对的形成导致了总能量最低的状态称之为 BCS 基态。库珀对彼此通过电子－声子相互作用而成为一个整体，因此它们可以用一个多电子波函数来描述，库珀对所对应的超导态是一种整体行为。在超导态下，组成库珀对的电子也受到散射，但它们通过电子－声子交互作用在散射过程中保持其质心的动量不变，也就是它们对电流的总贡献不变。当超导体温度增加时，晶格平均动能增大，就可能拆散一些库珀对使其成为正常电子。由于超导基态与最低激发态之间存在着能隙 Δ，拆散一个库珀对并把两个电子激发到正常态至少需要 2Δ 能量。能隙也是温度的函数。计算表明温度愈高，能隙愈小。当温度上升到临界温度 T_c 以上时，能隙变为零，库珀对也就不能存在，超导电性随之消失。应用能隙概念可以说明超导电性的起因，也可解释临界温度、同位素效应、电子比热和红外吸收等与超导相关的实验现象。

BCS 理论在解释非常规超导体的超导电性方面遇到了困难，目前还没有找到一个完整的理论模型来解释高温超导体的超导电性。重费米子超导体和有机超导体的情况也是如此。

4.3 材料的磁性

4.3.1 磁性的来源

原子的磁性是材料磁性的来源和基础，我们首先简介孤立原子的磁性，然后再讨论原子在晶格中的磁性行为。原子的磁矩有三个来源，即电子在其轨道上运动产生的磁矩，电子自旋产生的磁矩和原子核的磁矩。其中原子核的磁矩比前二者要小 2 000 倍左右，对于宏观磁性而言一般可以忽略不计。

电子的轨道磁矩与其轨道角动量成正比，角量子数为 l 的轨道电子的轨道磁偶极矩 μ_l 为：

$$\mu_l = [l(l+1)]^{1/2}\mu_B \tag{1.2-60}$$

式中，μ_B 为波尔磁子，在 CGS 制中其表式和数值为：

$$\mu_B = qh/2m = 0.927\,3\times10^{-27}\ \text{J/g} \tag{1.2-61}$$

式中，q，m 分别为电子的电荷和质量。电子轨道磁矩的 z 方向分量 μ_{lz} 为：

$$\mu_{lz} = m_l\mu_B \tag{1.2-62}$$

式中，m_l 为磁量子数。

电子的自旋量子数为 1/2，由此可求出其自旋磁矩 μ_s 和轨道磁矩的 z 方向分量 μ_{sz} 分别为：

$$\mu_s = (3)^{1/2}\mu_B \tag{1.2-63}$$

$$\mu_{sz} = \pm(1/2)\mu_B \tag{1.2-64}$$

原子的总磁矩 μ_J 及其 z 方向分量 μ_{Jz} 可借助量子力学由其所含各电子的轨道磁矩和自旋磁矩求出，其表达式分别为：

$$\mu_J = g_J[J(J+1)]^{1/2}\mu_B \tag{1.2-65}$$

$$\mu_{Jz} = g_J m_J \mu_B \tag{1.2-66}$$

式中，J 为总角量子数；m_J 为总磁量子数；g_J 为朗德因子。关于计算的细节可参阅有关磁学的专著。计算表明，原子中内部满壳层电子的总角动量和总磁矩都等于零，只有当原子中有未被填满的壳层时它才有不为零的总磁矩。以上关于原子磁矩的介绍也完全适用于离子。原子结合成晶体后其电子态和磁性与孤立原子相比发生了变化，其原因有二：一是键合使外层电子发生变化，例如共价键和离子键往往造成满壳层结构，金属键产生大量共有的传导电子等；二是晶体电场效应，即一个离子中的轨道电子的运动受临近离子产生的电场的作用而变化。

4.3.2 磁性的分类

材料在外磁场作用下被磁化，单位体积的磁矩总和称为磁化强度 M，磁化强度与外磁场 H 的比值 χ 称为磁化率，它和磁感应强度 B、磁导率 μ 在 CGS 制中由以下三式决定：

$$\chi = M/H \tag{1.2-67}$$

$$B = H + 4\pi M = \mu H \tag{1.2-68}$$

$$\mu = 1 + 4\pi\chi \tag{1.2-69}$$

物质的磁性可分为五类：抗磁性，顺磁性，反铁磁性，铁磁性和亚铁磁性。前三种很弱，后两种为强磁性。

抗磁性是一切物质都具有的一种磁性，其磁化率为负值且绝对值很小，大约为 $10^{-6}Z$（Z 为该元素的原子序数）。它主要来自局域电子的抗磁性，其机制是电子轨道磁矩在外磁场作用下作以磁场方向为轴的进动，这种进动产生的磁矩总是与外磁场反向，导致了抗磁性。自由电子也有抗磁性，但它小于自由电子的顺磁性，因而一般无关大局。

顺磁性物质的磁化率恒为正，主要有三个来源：一是局域电子的顺磁性，来源于外磁场对原子或离子原本无序的固有磁矩的取向作用。只有含有固有磁矩不为零的原子或离子的材料才有这种顺磁性。其磁化率小，但一般较抗磁性为大，与温度的关系满足居里－外斯定律：

$$\chi = C/(T - T_p) \tag{1.2-70}$$

式中，C 和 T_p 为常数；T 为热力学温度。第二个来源为非定域电子的顺磁性，主要存在于金属材料中。金属导带中正负自旋电子在磁场中能量不同，在磁场作用下正负自旋电子数变得不等，导致了自由电子的顺磁性。第三个来源为范弗莱克（Van Vleck）效应，它产生于外磁场对电子云的形变。

如果说抗磁性和顺磁性只代表独立磁矩集合的性质，那么反铁磁性，铁磁性和亚铁磁性则反映大量磁矩的合作现象。图 1.2-56 显示了除抗磁性以外的四种磁性中的磁矩分布和磁化率随温度变化的情况。让我们先看铁磁性，典型的铁磁材料是铁，钴，镍。如图 1.2-56b 所示，在居里温度 T_c 以下其磁矩有自发平行排列的趋势。在 T_c 以上铁磁体具有

顺磁性，其磁矩是无规取向的，磁化率与温度的关系满足居里－外斯定律。

典型的反铁磁材料是铬，锰及氧化铬，氧化锰，氧化钴等，其磁化行为如图1.2-56c所示。在转变温度 T_N（称作尼尔点）以上磁矩是无规分布的，具有顺磁行为。但在尼尔点以下，磁矩是交替反平行排列的，其合成磁化强度为零，属于弱磁性。

图1.2-56d显示的是亚铁磁性，铁氧体如四氧化三铁是典型的亚铁磁体。在 T_c 以上磁矩无规排列，但在 T_c 以下形成反平行排列，不过由于两子格上的磁矩大小不等，所以出现了净磁化强度，它和铁磁性一样属于强磁性。

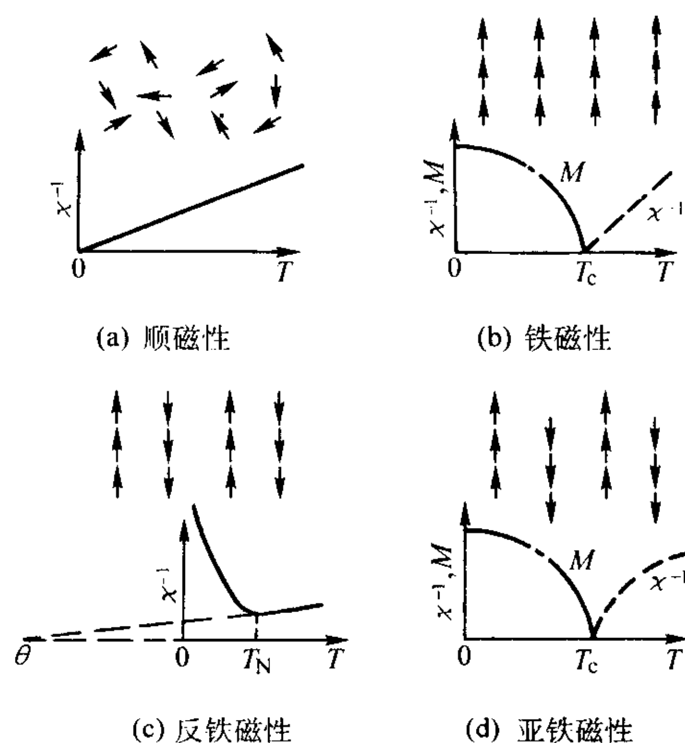


图1.2-56 磁矩分布和磁化率随温度变化的情况

反铁磁体，铁磁体和亚体磁体都是磁有序结构。这种有序结构来源于晶体格点上的离子磁矩之间以及离子磁矩与非定域的电子磁矩之间直接或间接的交互作用。正是这些交互作用造成了磁矩或平行排列，或反平行排列，其细节可阅读磁学方面的专著。

4.3.3 磁化过程，软磁与硬磁

在居里温度以下强磁体内自发磁化的分布形成多磁畴结构。一个磁畴指磁体内部磁化方向相同的一个微小区域，在区域内各原子的磁矩平行排列。磁畴之间的界面称为畴壁，其厚度在数十纳米的量级，畴壁中原子的自发磁化方向实现逐渐过渡。未加磁化的磁体中由于大量磁畴的磁化方向无序排列，在无外磁场时其总磁化强度为零。在外磁场作用下磁畴调整其大小和方向导致磁化。磁化有三种机制：一是畴壁移动，即磁化平行于外磁场的畴长大，反平行于外磁场的畴缩小；二是磁畴转动，即各磁畴的磁化方向逐渐转向与外磁场平行的方向；三是内禀磁化，即在各磁畴的磁化方向都趋于与外磁场一致后，材料的磁化还可随外磁场的继续增大而缓慢地线性增加。其示意见图1.2-57。

畴壁移动和磁畴转动都要消耗能量，这导致了磁化过程的不可逆性即磁滞回线，如图1.2-58所示。OA 表示对于未磁化的样品施加磁场 H，随 H 增加磁化强度不断增加，当 H 增加到 H_s 时磁化强度达到饱和强度 M_s。达到饱和以后，再减小磁场，磁化强度并不是可逆地沿原始的磁化曲线下降，而是沿着图中 AB 变化。在 B 点磁场已减为零，但磁化强度并没有消失。只有当磁场沿相反方向增加到 $-H_c$ 时，磁化才变为零，H_c 称为矫顽力。继续增加反向磁场到 $-H_s$ 可以使磁化强度达到反向的饱和。这时如果再由 $-H_s$ 增加到 H_s，磁化强度将完成如图示的回线，称为磁滞回线。不同的铁磁材料的磁化曲线可以有很大的差别。例如，软磁材料要求高磁化率（即高 M_s）和低矫顽力，有些材料 H_c 只有800 A/m量级。而硬磁材料则要求高磁化率和高矫顽力，最好的硬磁材料如钕铁硼合金，其矫顽力高达800 kA/m以上。

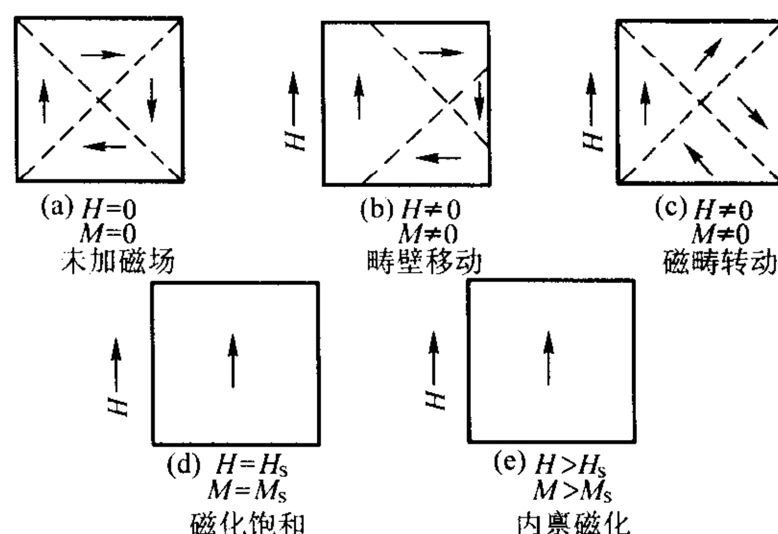


图1.2-57 三种磁化过程的示意图

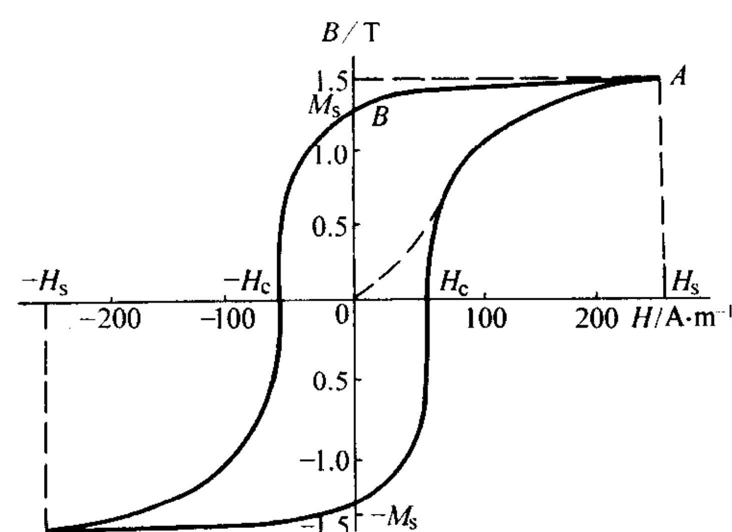


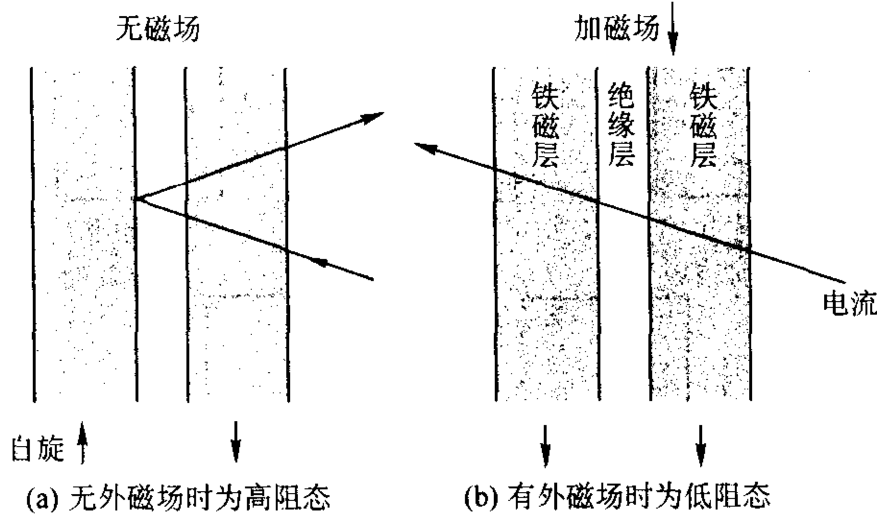
图1.2-58 磁滞回线

4.3.4 巨磁电阻效应简介

磁电阻指磁场使材料的电阻发生变化的现象，定义为外磁场导致的电阻改变 ΔR 与原电阻 R 之比。一般强磁体的磁电阻来源于磁畴中磁电阻的各向异性，其数值很小，在百分之几的量级，称为正常磁电阻（OMR）。1988年人们在纳米金属铁/铬多层膜中发现了巨磁电阻，其 $\Delta R/R$ 为负值，绝对值较正常磁电阻高1～2个量级，被称作巨磁电阻（GMR）效应。其后人们又发现了两种新的巨大磁电阻效应：一是钙钛矿型锰氧化合物中的庞磁电阻（CMR）效应，其 $\Delta R/R$ 值可高达 10^6%；二是铁磁体/绝缘体/铁磁体中的隧道结巨磁电阻（TMR）效应，在应用方面有巨大吸引力。为了解释以上四种磁电阻效应，人们提出了自旋相关导电的理论。

过去人们在分析导电问题时只考虑电子的电荷而不考虑电子的另一个属性自旋。自旋相关导电理论则提出，在巨磁电阻材料中电阻与载流子的自旋有关。不同的材料又有不同的自旋相关导电机制，例如自旋相关散射，自旋相关隧穿和自旋极化电子导电等。以铁磁体/绝缘体/铁磁体隧道结巨磁电阻效应为例，无外磁场和有外磁场情况下隧道结的图示分别见图1.2-59a、b。在无外场时两铁磁层极化反平行，因为这样自由能较低。考虑电子由右侧隧穿进入左侧，右侧铁磁体极化向下，其费米面附近自旋向下电子的态密度较大，自旋向上电子的态密度较小，而左侧铁磁体极化向上，其费米面附近自旋向上电子的态密度较大，自旋向下电子的态密度较小。由于隧穿过程中电子的自旋不变，这就造成右侧费米面附近许多自旋向下的电子无法隧穿进入左侧而被反射，这便是隧道结的高阻态。如果在隧道结上加一个足够大向下的外磁场，隧道结两侧铁磁体的极化变为都向下，这时两侧费米面附近自旋向下电子的态密度都较大而自旋向上电子的态密度都较小，由于隧穿过程中电子的自旋不变，这就使右侧

自旋向下和向上的电子都可较方便地不改变自己的自旋而隧穿进入左侧，这便是隧道结的低阻态。由于隧道结中发生的是量子隧穿过程，结中绝缘层必须很薄，厚度仅为1~2 nm。



(a) 无外磁场时为高阻态 (b) 有外磁场时为低阻态

图 1.2-59 磁性隧道结电子的自旋相关隧穿

4.4 介电性与铁电性

4.4.1 介电性

材料可按其对外电场的响应方式区分为两类，一类以电荷长程迁移即传导的方式对外电场作出响应，这类材料称为导电材料。另一类以感应的方式对外电场作出响应，即沿电场方向产生电偶极矩或电偶极矩的改变，这类材料称为电介质，这种现象称为电介质的极化。广义而言，一些导电材料也可表现出介电性，但这不是电介质物理的研究范畴。通常，绝缘体都是典型的电介质。电介质又可分为非极性电介质和极性电介质两大类。前者由非极性分子组成，在无外加电场时分子的正负电荷重心互相重合，不具有电偶极矩。只是在外加电场作用下正负电荷出现相对位移，才出现电偶极矩。后者由极性分子组成，即使在无外场时每个分子的正负电荷重心也不互相重合，具有固有电矩，它与铁电性有密切关系。电介质的极化有3种主要基本过程，即材料中原子核外电子云畸变产生的电子极化；分子中正、负离子相对位移造成的离子极化和分子固有电距在外电场作用下转动导致的转向极化。此外，与材料中的缺陷相关联还有一些其他极化机制，但其重要性不如前3种。

描述一个中性分子电极化状态的参量是电偶极矩 p，其表达式为

$$\boldsymbol{p} = q\boldsymbol{l} \tag{1.2-71}$$

其中 q 为分子中正电荷的总量（显然其中负电荷的总量为 $-q$）；$\boldsymbol{l}$ 为正负电荷重心之间的位矢，其方向由负电荷重心指向正电荷重心。介质的电极化强度是描述电介质极化状态的基本物理量，其定义为该电介质单位体积内电偶极矩的矢量和，即：

$$\boldsymbol{P} = \sum \boldsymbol{p}/V \tag{1.2-72}$$

表征电介质在静电场作用下介电性质的主要参量为介电常数 ε，由电学知识可知，电介质电位移矢量 $\boldsymbol{D}$ 与外加电场 $\boldsymbol{E}$ 的关系为

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} = \varepsilon\varepsilon_0 \boldsymbol{E} \tag{1.2-73}$$

其中 $\boldsymbol{P}$ 为电极化强度，在电场较小的情况下有

$$\boldsymbol{P} = \chi\varepsilon_0 \boldsymbol{E} \tag{1.2-74}$$

式中，χ 为介电极化率；ε_0 为真空介电常数，其数值为$\varepsilon_0 = 8.85\times10^{-12}$ F/m。由上述两式可求出

$$\boldsymbol{D} = \varepsilon_0 (1+\chi) \boldsymbol{E} \tag{1.2-75}$$

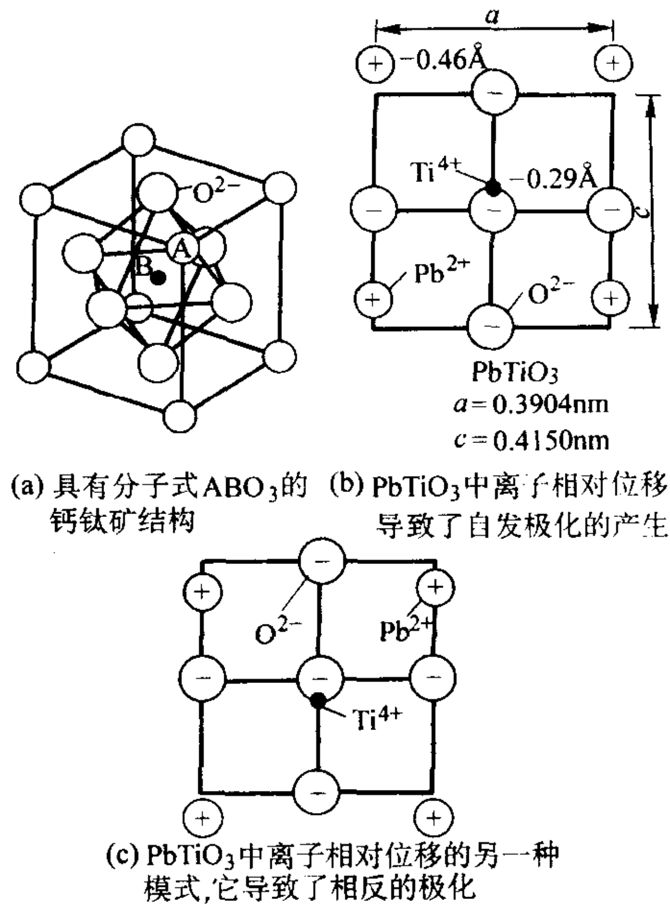
由此可知，介电常数 ε 与介电极化率 χ 只差一个常数1。在均匀极化的电介质内部其各分子电偶极矩的正端总是和另一电偶极矩的负端相连，正负端的电荷恰好互相抵消而使介质内呈现电中性，但在电介质的表面这种抵消不再能实现。电偶极矩正端露头处便呈现正电荷，负端露头处呈现负电荷，由于这些电荷束缚于电偶极子，不能自由移动，这种电荷称为束缚电荷。束缚电荷的面密度 σ 满足

$$\sigma = \boldsymbol{P}\boldsymbol{n}_0 \tag{1.2-76}$$

式中，$\boldsymbol{n}_0$ 为表面外法线上的单位矢量。这些面束缚电荷在电介质内部产生一个电场，称为退极化电场，其方向与 $\boldsymbol{P}$ 相反。

4.4.2 铁电性

铁电体是一类具有自发极化，而且这种自发极化可以在外电场作用下改变方向的电介质。因此，自发极化及其变化便成了铁电体研究的核心问题。通常晶体的铁电性只存在于某一特定温度之下。在此温度之上铁电体变成顺电体，其自发极化消失，铁电相与顺电相之间的转变称为铁电相变，该温度称为居里温度 T_c。铁电相变可按其结构转变的特征分为两类，即位移型铁电相变和有序无序型铁电相变。



(a) 具有分子式ABO_3的钙钛矿结构 (b) $PbTiO_3$中离子相对位移导致了自发极化的产生

(c) $PbTiO_3$中离子相对位移的另一种模式，它导致了相反的极化

图 1.2-60 钙钛矿结构的铁电转变

许多铁电体具有钙钛矿结构，图1.2-60为钙钛矿结构晶体原胞的示意图。金属离子A位于立方体的8个顶角上，氧离子位于六个面心位置上并构成氧八面体的6个项点。金属离子B位于原胞的中心，同时也位于氧八面体的中心。显然，没有畸变的钙钛矿结构具有对称中心，其正负电荷重心互相重合，因而不具有自发极化，也不可能有压电性。钙钛矿结构的一个重要特征是其结构的不稳定性。某些这类晶体如钛酸铅（$PbTiO_3$）在高温下具有标准的钙钛矿结构，但随着温度下降和晶格中离子振动减弱，位于氧八面体中心的B离子即钛离子变得不稳定而出现向氧八面体某一顶角方向移动的倾向，并终于在490℃发生位移型铁电相变，由顺电相变为铁电相。在此过程中离子移动的细节见图1.2-60b。显然该图中画出的只是钛离子移动的一种可能方式。考虑到顺电相钛酸铅具有立方结构，实际上钛离子的位移可以向着氧八面体六个顶角中的任意一个发生。在钛离子移动的同时位于原胞顶角上的铅离子也向相同的方向移动，其后果是该原胞由立方结构变成了四方结构，同时原胞中正电荷的重心不再与负电荷的重心重合，从而导致了自发极化的出现。自发极化的方向与四方铁电相的 C 轴一致。

铁电体表现的宏观电极化强度 $\boldsymbol{P}$ 与外电场 $\boldsymbol{E}$ 密切相关。实际上 $\boldsymbol{P}$ 是铁电体中电畴的自发极化与感应极化的总和。

由于电畴在电场作用下转向的过程总是滞后于电场，$\boldsymbol{P}$ 与 $\boldsymbol{E}$ 呈现出看上去类似于磁滞回线的复杂的“电滞回线”关系。对于一多畴铁电单晶体，一般说来，外加电场前它不表现出宏观电极化。若沿其某一极化轴方向施加一外电场，其极化矢量 $\boldsymbol{P}$ 与外加电场 $\boldsymbol{E}$ 的关系如图 1.2-61 所示。在外加电场之初，外电场很小时，$\boldsymbol{P}$ 随 $\boldsymbol{E}$ 线性增加，如 OA 段所示。此时 P 的增加主要来自感应极化和“弹性”的可逆畴壁移动。当电场进一步增加，新畴成核与不可逆畴壁移动造成 $\boldsymbol{P}$ 随 $\boldsymbol{E}$ 的增加而迅速增加，如 AB 段所示。当电场增至 B 点所在处的数值，晶体已变成单畴，自发极化对 $\boldsymbol{P}$ 的贡献趋于饱和。在此之后 $\boldsymbol{E}$ 继续增加时 $\boldsymbol{P}$ 线性增加，这种增量完全来自感应极化的贡献，如 BC 段所示。此后，若所加电场强度减小，极化强度 $\boldsymbol{P}$ 将循 CBD 曲线减小。在电场降至 0 时，仍保留有相当大的电极化强度 OD，称为剩余极化 P_r。将 CB 外推与 P 轴相交于 K 点，OK 表征自发极化 P_s。此后若电场反向强度增加，极化将逐渐减小。至 F 点，极化下降为 0。OF 表征的电场强度称为矫顽场，它表示将一个单畴铁电晶体变成宏观极化为 0 的多畴晶体所需的外加电场强度。反向电场继续增大，极化反向并沿 FG 曲线变化直至饱和。电场在大于正负饱和值之间循环一周过程中，电极化强度与电场强度沿封闭曲线 $CBDFGHC$ 变化。这一曲线称为电滞回线，它是铁电体的标志。近年来铁电薄膜在高密度不挥发铁电存储器中得到了重要应用。一般说来铁电体都具有压电性和热释电性。

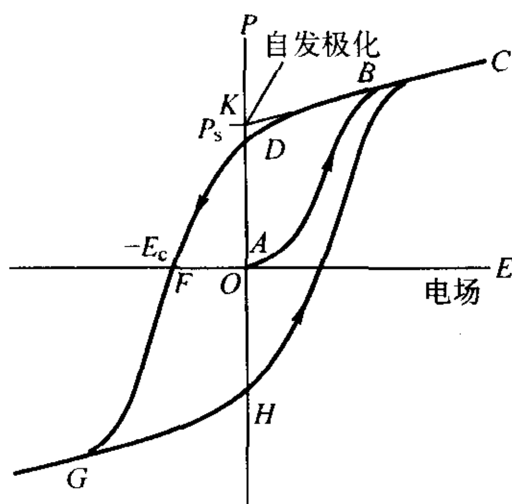


图 1.2-61 典型的铁电体电滞回线

4.4.3 压电性

某些介质在机械力作用下发生电极化或电极化的变化，这样的性质称为压电性。电极化的改变导致介质与极化方向垂直的两端面出现等量反号的束缚电荷变化，看起来这是由于压力造成了电荷的变化，这就是压电性一词的由来。具有压电性的介质称为压电体。上述过程的逆过程为：在压电体的适当方向上施加外电场会导致压电体发生应变，这种现象称为逆压电效应。

介质具有压电性的条件是其结构不具有对称中心，在 32 类点群中有 20 类点群不具有对称中心，属于这 20 类点群的电介质才可能是压电体。不具有对称中心的晶胞在应力作用下出现极化和表面束缚电荷的示意图见图 1.2-62。铁电体具有自发极化，它们不仅结构上没有对称中心，而且在无外力作用时晶胞的正、负电荷重心就不互相重合。在应力作用下，其自发极化一般都要发生变化，因而常见的铁电体都具有压电性。压电晶体和陶瓷最重要的应用是作为机电换能器和电声转换器。其主要功能系数为机电耦合系数 K，定义为：

K^2 = 由逆压电效应获得的机械能/储入的电能总量

或 K^2 = 由压电效应转换的电能/储入的机械能总量

此系数越大，材料的压电耦合效应越强。

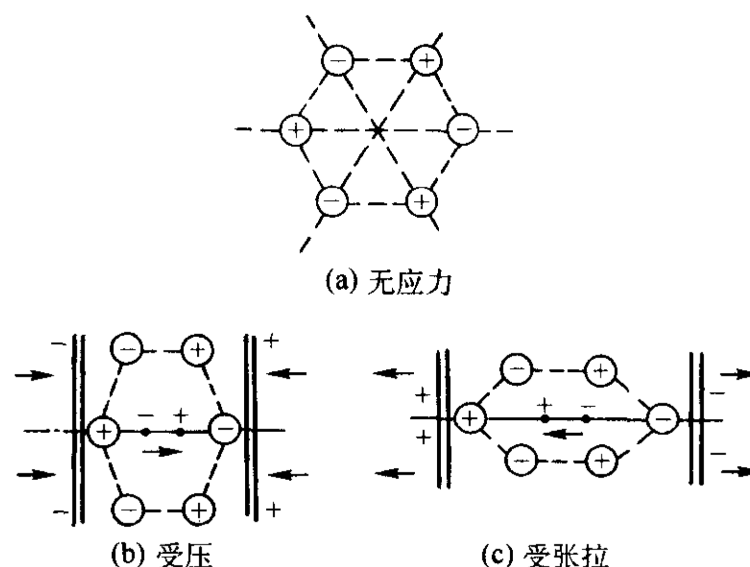


图 1.2-62 压电体晶胞受压，张应力出现极化和面束缚电荷

4.4.4 热释电性

热释电性指的是某些电介质的电极化随温度改变的性质。这些电介质称为热释电体。实际上热释电体是有自发极化的电介质。32 种点群有 10 种点群可出现自发极化，属于这些点群的电介质可具有热释电性。一般说来，铁电体都具有热释电性。考虑一个单畴的铁电晶体或经过极化处理的铁电陶瓷，与极化矢量垂直的相对的两个表面附近分别出现正、负束缚电荷。通常，这些束缚电荷被等量而符号相反的自由电荷所屏蔽，总体表现出电中性。当温度改变时，电极化强度发生变化，原先的自由电荷不再正好可屏蔽束缚电荷，于是表面上就表现出电荷的存在。如果这时在相对两表面装上电极，在连接两电极的导线上就会有电流流过。材料的热释电性质用热释电性系数 $\boldsymbol{p}$ 来表示，若热释电晶体的温度改变 ΔT，则其极化强度的改变 $\Delta \boldsymbol{P}$ 可由下式给出：

$$\Delta \boldsymbol{P} = \boldsymbol{p}\Delta T \qquad (1.2\text{-}77)$$

由于 $\boldsymbol{P}$ 为矢量，T 为标量，热释电系数 $\boldsymbol{p}$ 应为一矢量，与自发极化的方向一致，其单位为 C/ ($m^2 \cdot K$)。热释电材料的最重要的应用是用于红外探测和成像的热释电传感器和热释电成像焦平面。

编写：刘治国（南京大学）

第3章　基础材料与新材料的现状与发展

1　概述

材料可以广义地分为基础材料与新材料两大类，但基础材料体系中会有新材料衍生出来，而新材料体系中也存在向基础材料的转化。

基础材料习惯上称为传统材料，一般是指那些已经成熟并且在工业中已批量生产并大量应用的材料。但“传统材料”的提法容易与“落后”相混淆，故“基础材料”的提法更确切。例如钢铁材料、有色金属材料、建筑材料、化工材料等。

新材料一般是指近期发展起来或正在发展的具有优异性能和特殊功能并具有应用前景的材料，例如新能源材料、超导材料、纳米材料、生物医用材料、生态环境材料等。

本章介绍主要基础材料与新材料的现状与发展，鉴于基础材料与新材料之间存在着一些衍生与转化的关系，故在基础材料介绍中也会涉及到一些新材料，例如钢铁材料中的“超级钢”；而在新材料的介绍中也会涉及到一些已经转化为基础材料的实例，像复合材料中的玻璃纤维增强塑料（玻璃钢）。

由于基础材料和新材料种类繁多，而本章的篇幅有限，所以介绍只能是纲要性的，着重介绍综合概况和发展趋势，更具体的信息和数据，要依靠《大典》各卷以及其他手册和专著来提供，对于有些资料和数据在本章中的引用谨向有关作者致谢。

2　钢铁材料

2.1　分类概况

钢铁材料包括铁、铁合金、非合金钢、低合金钢、合金钢和高合金钢等。本节主要介绍结构用钢材。钢是一种合金，其性能主要取决于铁与碳相互作用后生成物的结构、数量和分布状态。铁－碳系平衡相图就是研究和描述碳钢的组成、加工工艺、组织结构和性能之间关系的依据（见图1.3-1）。

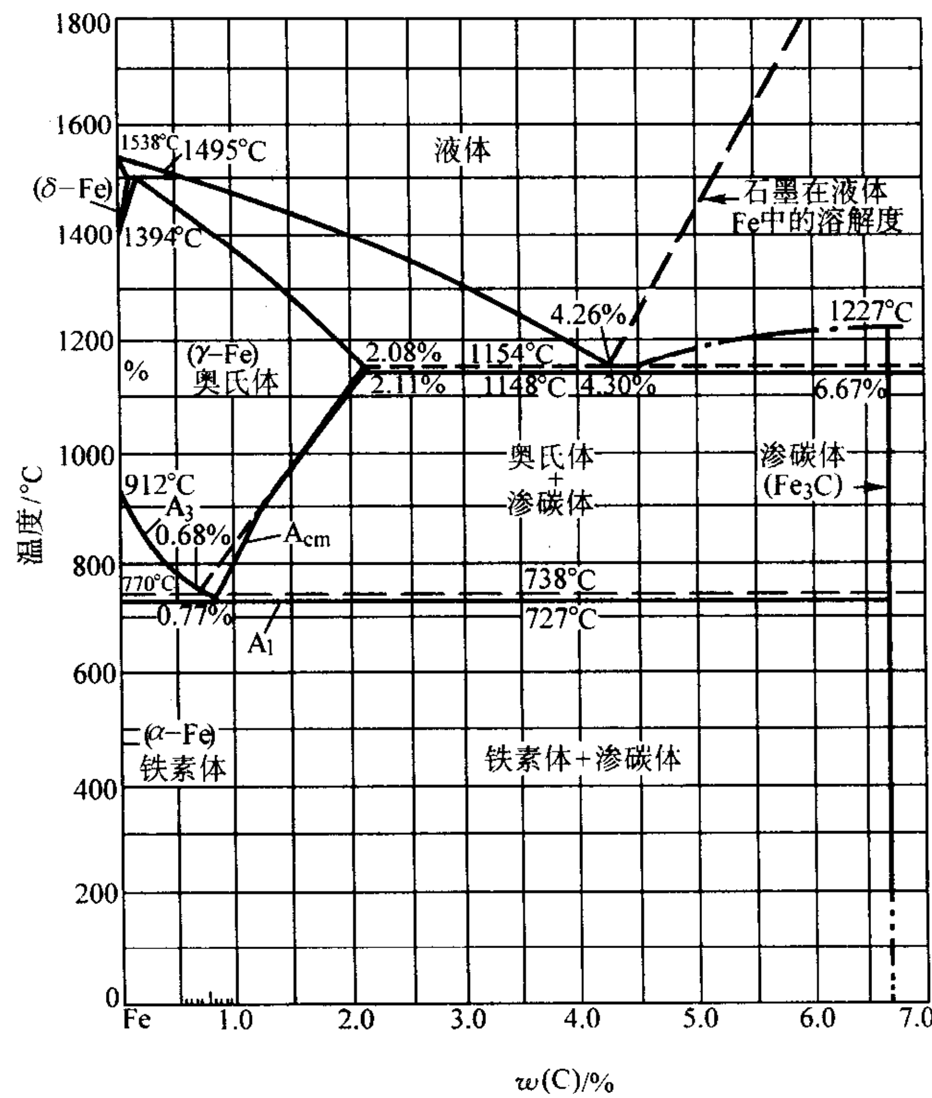


图1.3-1　铁－碳系平衡相图

按用途分类，钢可以分为：①建筑和工程结构用钢；②机械制造用钢；③工模具钢；④特殊性能钢。

按钢的材型分类，钢可以分为铁道用钢、型材、钢带、薄板、中厚板、特厚板、无缝钢管、焊接钢管、线材、钢丝、锻件等。

按冶炼方法分类，钢可以分为：①沸腾钢；②镇静钢；③半镇静钢；④连铸钢；⑤洁净钢；⑥平炉钢；⑦转炉钢；⑧电弧炉钢；⑨真空冶炼钢；⑩等离子熔炼钢；⑪电渣重熔钢等。

按钢中碳含量的不同，钢可以分为超低碳钢、低碳钢、中碳钢和高碳钢；按钢中添加合金元素的多少，钢可以分为非合金钢、微合金钢、低合金钢、中合金钢和高合金钢；按加工状态和金相组织钢还可以分为热轧钢、冷轧钢、退火钢、正火钢、调质钢、亚共析钢、过共析钢、铁素体钢、奥氏体钢、贝氏体钢、马氏体钢等。

2.2　产业现状

1949年中国产钢15.8万吨，1978年产钢3 178万吨，1989年产钢超过6 000万吨，而从1996年起年产钢超过1亿吨，并且已连续8年钢产量和钢材产量都居世界第一位。1990—2003年中国粗钢产量增长情况见图1.3-2。据中国钢铁工业协会快报统计，2003年中国钢铁全行业产钢22 234万吨，同比增长22.38%；产生铁20 231万吨，同比增长19.65%；产钢材（含重复材）24 119万吨，同比增长24.3%。

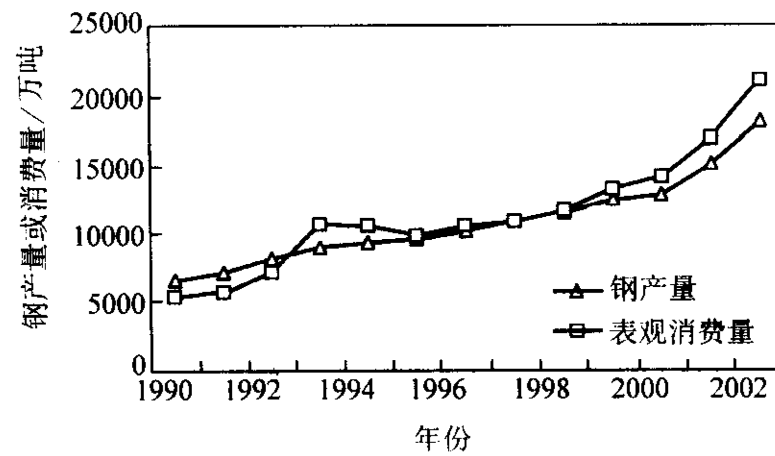


图1.3-2　中国近年来的钢产量与消费量

2003年我国钢产量首次突破2亿吨，连续第八年钢产量居世界第一位，也是全球第一个年产钢量突破2亿吨的国家，在中国钢铁工业发展历史上具有里程碑重大意义。

2003年钢产量超过500万吨的有13户大企业，合计产钢9 789万吨，占全国钢产量的44%。其中宝钢1 987万吨、鞍钢1 018万吨、武钢843万吨、首钢817万吨、本钢720万吨、唐钢608万吨、马钢606万吨、邯钢606万吨、攀钢534万吨、包钢525万吨、华菱519万吨、济钢505万吨、沙钢502万吨。此外，钢产量在200万～500万吨的有21户大企业，合计产钢5 549万吨，占全国总量的25%。

我国钢铁工业结构调整、技术改造和技术进步都获得了

明显的进展。

初步统计，通过新建和技术改造，我国热连轧宽带钢轧机生产能力由 2002 年的 3 200 万吨上升到 2003 年的 4 608 万吨，增加 1 408 万吨，增长 44%；冷轧宽带钢轧机能力由 1 129 万吨上升到 1 414 万吨，增加 285 万吨，增长 25.2%；镀锌板机组能力由 330 万吨上升到 583 万吨，增加 253 万吨，增长 76.6%；彩涂板机组能力由 446 万吨上升到 873 万吨，增加 427 万吨，增长 95.7%。冷轧不锈钢薄板机组能力不完全统计由 70 多万吨上升到 100 多万吨，增加 30 多万吨，约增长 42.8%。同时又有一批新的板带材轧机、涂镀层生产线及型线材生产线开工建设。由于新建机组多数设备制造拖期，调试时间长，当年大部分能力没有发挥。

占全国钢产量 85.26% 的重点统计的大中型钢铁企业连铸比由 2002 年的 92.89% 上升到 95.30%。钢材综合成材率平均达 94.93%，同比提高 0.37%。因铁、钢比下降，吨钢综合能耗由 807 kg 标准煤降到 770 kg 标准煤，下降 3.8%。高质量要求、高附加值产品产量有显著增长。其中，低合金中厚板产量增长 62%，压力容器板增长 65%，锅炉钢板增长 48.6%，集装箱板增长 40%，管线钢板增长 45%，桥梁钢板增长 26%，造船板增长 10%，不锈钢材增长 59.3%，轴承钢材增长 39.1%，齿轮钢材增长 22%，模具钢材增长 31.5%，合金结构钢材增长 19.9%。节能、节水、环境保护先进技术得到了进一步推广。

2003 年我国进口钢材 3 717 万吨，同比增长 1 268 万吨，增长率为 51.8%。同期我国出口钢材 696 万吨，同比增长 27.5%。2003 年我国成为最大的铁矿石进口国，全年进口总量为 14 812 万吨，同比增长 32.87%。

目前我国钢铁工业存在的主要问题是由于快速增长而引起的铁矿石、焦炭、焦煤等原料和燃料供应紧张，价格和海运费暴涨，部分电力供应短缺，铁路运输和港口接卸能力不足。资源、能源、交通运输等制约发展的因素短期内难以根本缓解，为此应该增强忧患意识，抑制盲目投资和低水平产能扩张。此外，钢铁工业产业结构调整还需深入进行下去，要积极促进企业深化改革、改制、重组，以进一步提高产业集中度，实行我国钢铁工业可持续发展的战略。

2.3　发展趋势

在国际上，从 1993 年开始，日本进行了两年的社会调查，提出了为提高钢铁材料与其他基础材料之间的竞争力，也为了迎接日本战后 50 年基础设施的重建，需要有新的发展目标，即钢铁材料的“强度翻番、寿命翻番”。1997 年日本政府投资 1 000 亿日元安排了十年的“超级钢”基础研究计划，相应其他国家也安排了类似项目，我国于 1998 年也安排了第一期为期 5 年的 973 计划项目“新一代钢铁材料重大基础研究”，参见表 1.3-1。

表 1.3-1　各国新一代钢铁材料的研究发展计划

年份	国别	项目	备注
1997 年	日本	“超级钢”基础研究	科技厅，十年项目
1997 年	日本	“超级金属”计划	解决控制制造技术
1998 年 3 月	世界钢铁协会	“超轻钢车身”	35 家西欧、北美钢厂、汽车厂参加
1998 年	韩国	“21 世纪高性能结构钢”	十年项目
1998 年 10 月	中国	“新一代钢铁材料的重大基础研究”	第一期五年项目
2001 年 7 月	欧盟	“超细晶粒钢”	英、德、意、比等国
2002 年 1 月	日本	“环境友好型超微细晶粒钢的基础技术研究”	经济产业省

中国“新一代钢铁材料重大基础研究”973 项目中提出了五个主要的科学问题：①高洁净度的化学冶金学问题；②超细晶组织（微米级的铁素体或贝氏体；微米级的贝氏体/马氏体条束）的形成理论；③非平衡态的物理金属学问题；④微合金化理论的新认识；⑤相关技术的基础理论。在上述科学问题的基础上提出了新一代钢铁材料应具备的三个基本特征。

1）超细晶　将传统 10 ~ 100 μm 的晶粒度进行超细化研究，对不同类钢分别达到超细化的尺寸范围见表 1.3-2，其强度增量见表 1.3-3，各种强化效果示意图见图 1.3-3。

表 1.3-2　不同钢类确保“强度翻番”的超细晶化尺寸范围

钢类	组织	现有强度 /MPa	目标强度 /MPa	超细晶尺寸/μm
低碳钢	铁素体 + 珠光体	σ_s 约 200	σ_s 约 400	（α）约 5
低（微）合金钢	低碳贝氏体或针状铁素体	σ_s 约 400	σ_s 约 800	（α）1 ~ 2
合金结构钢	回火马氏体或贝/马复相钢	σ_b 约 800	σ_b 约 1 500	（γ）约 5

表 1.3-3　各种产品可能达到的最小晶粒直径和相应的强度增量

产　品	晶粒直径 d/μm	$d^{-1/2}$/mm	强度增量 $kd^{-1/2}$/MPa
厚板	8 ~ 10	10 ~ 12	174 ~ 209
薄板	5	14	244
正火板材和棒材	5	14	244
带材	3 ~ 4	16 ~ 18	278 ~ 313
实验室最佳正火材	3 ~ 4	16 ~ 18	278 ~ 313
实验室最佳控轧材	1 ~ 2	25	435

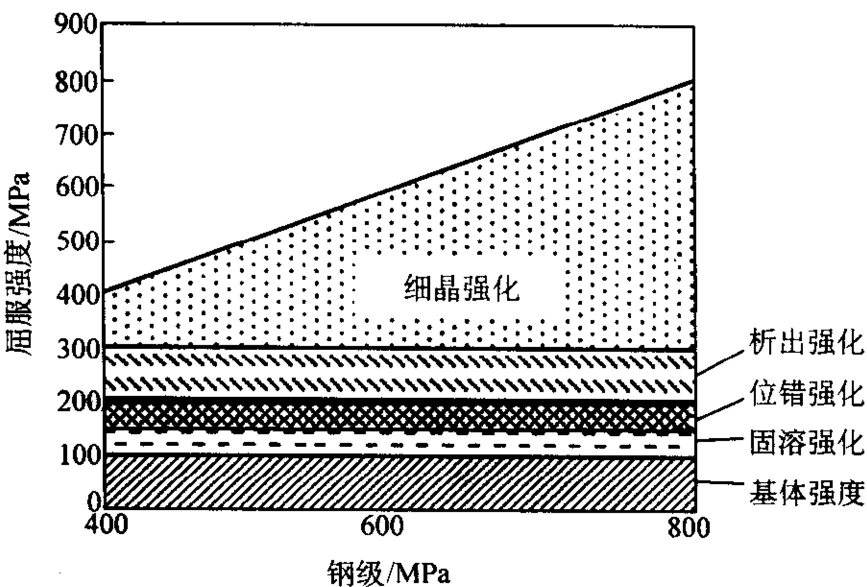


图 1.3-3　各种强化效果示意图

2）高洁净度　洁净度是指钢材内部杂质含量和夹杂物形态能满足使用要求，我们把它称为“经济洁净度”。由于

钢的强度翻番，材料在使用时承受了更高应力，使裂纹形成和扩展的敏感性增加。按照断裂力学的基本概念，在相同条件下，受力愈大，要求临界裂纹尺寸（以夹杂物大小作为内在不可避免的裂纹）愈小。新一代钢铁材料应有更高的洁净度。

3）高均匀性　钢液凝固过程中，由于传热规律造成顺序凝固。无论模铸还是连铸都带来低熔点元素的宏观偏析。随后的高温加热及大变形量轧制，都难以消除偏析。现代冶金的发展趋势是流程愈来愈紧凑，过程愈来愈快，材料组织向非平衡发展。为改善钢的均匀性，在凝固过程中应尽可能减少柱状晶，争取获得全等轴晶的钢坯，在杂质总量不变的情况下，提高均匀性相当于提高洁净度。

此外，新一代钢铁材料除上述三个基本特征外，从经济性和生产应用的要求考虑，还必须满足以下三个条件：

1）低成本、高性能　强度翻番，成本基本不增加；

2）强韧性良好配合　强度翻番后，韧性和塑性基本不下降；

3）符合可持续发展方针　合理利用资源，少用或不用合金化元素，注意环境保护和废钢的回收利用。

新一代钢铁材料重大基础研究项目根据五个科学问题中所描述的高洁净度应用基础理论、超细组织形成理论、所涉及的非平衡态物理金属学问题研究以及在新一代钢铁材料研发中重新认识或深入微合金理论研究基础上，形成了新一代钢铁材料的研发路线，其重点是超细晶的组织形成理论，它包括四个创新性的基础理论研究结果。

1）低碳（铁素体＋珠光体）钢的形变诱导铁素体相变理论，简称DIFT理论。为了增加相变驱动力，在更大过冷度下产生更小临界形核尺寸、更大形核速度、更细小的新相晶粒尺寸。在与DIFT同一热力学原理和更快动力学基础上，发展了形变强化铁素体相变，简称DEFT相变。

2）薄板坏连铸连轧流程中纳米析出相的发现及引进的超细化现象。

3）微合金钢（低碳贝氏体或针状铁素体为主的钢类）的形变诱导析出（简称DIP）和中温控制。

4）合金结构钢中的抗延迟断裂理论。

在上述基础理论研究结果的指导下，已经研制出新一代钢铁材料的原型钢种并获得了初步的示范工程应用。在碳素结构钢方面（目前约占钢材总消费量的50%左右），已由目前的屈服强度200 MPa级提高到了400 MPa级，这类材料的主要应用对象是建筑用钢和汽车用钢的更新换代；低合金结构钢（目前约占钢材消费总量的25%左右），已由目前的屈服强度为350～400 MPa级提高到700～800 MPa级，这类材料的主要应用对象是建筑、耐候钢、管线钢、工程机械和船舶用钢；合金结构钢（目前主要用于机械制造用钢，约占钢材总消费量的6%～10%）的抗拉强度由目前的800 MPa级提高到1 500 MPa级，重点解决抗延迟断裂性能过低这一技术难题。以上三类钢的传统晶粒度和新一代钢铁材料的超细组织晶粒度对比见图1.3-4。当然，上述工作只是围绕我国新一代钢铁材料的研究发展中的第一步，针对“强度翻番”问题还需要做更深入细致的研究工作，而在工程化和产业化方面的任务更为艰巨。同时在“强度翻番”的基础上，还需要进一步开展“寿命翻番”的研究发展工作。

钢种	传统晶粒度	超细晶粒度
普通低碳钢		
微合金钢		
合金结构钢	（γ晶粒尺寸）	（γ晶粒尺寸）

图1.3-4　三类钢材的传统晶粒度和超细晶粒度对比

3　有色金属材料

3.1　分类概况

俄罗斯和东欧一些国家将金属分为黑色和有色两大类。黑色金属指铁、铬、锰及其合金，其外观多半为深黑色或灰黑色；1958 年中国参照前苏联的分类方法，将铁、铬、锰以外的 64 种金属定为有色金属。美、英、日本和西欧则将金属分为铁和非铁金属两种，非铁金属是指铁以外的所有金属。在本书中我们沿用黑色金属与有色金属的分类方法。

有色金属的主要分类方法是按其密度、价格以及在地壳中的储量及分布情况而分为 5 种：

1）重有色金属　密度大于 4.5 g/cm³ 的有色金属，包括铜、铅、锌、镍、钴、锡、锑、汞、镉、铋等。

2）轻有色金属　密度小于 4.5 g/cm³ 的有色金属，包括铝、镁、钛、钠、钾、钙、锶、钡等。

3）贵金属　包括金、银和铂族金属（铂、钯、铑、铱、锇、钌等)。因其在地壳中含量很少（如银的丰度为 10^{-7}，金和铂的丰度为 5×10^{-10})、提取困难、价格比一般金属贵而得名。

4）稀有金属　自然界中含量少（大部分金属的丰度小于 0.01%），分布稀散或难以从原料中提取的金属。包括稀有轻金属，如铍、锂、铷、铯；难熔金属，如锆、铪、钒、钽、铌、钨、钼；稀散金属，如镓、铟、铊、锗、铼、硒、碲；稀土金属，钪、钇和镧系元素的统称，共有钪、钇、镧、铈、镨、钕、钷、钐、铕、钆、铽、镝、钬、铒、铥、镱、镥 17 种元素；放射性金属，如钫、镭和锕系金属中的锕、钍、镤、铀以及人造的锝和钷等。

5）半金属　物理化学性质介于金属与非金属之间的金属，如硅、碲、硼、砷等。

除上述主要分类方法外，有色金属还可以按生产方法分为以下 4 种：

1）冶炼产品　以冶炼方法得到的各种有色金属及其化合物，如铝锭、铜锭、各种氧化物及盐类等。

2）塑性加工产品　以塑性加工方法得到的各种板、带、管、棒、型、线和锻件等产品。

3）铸造产品　以铸造方法获得的各种有色金属和合金铸件。

4）金属粉末　以机械方法或雾化方法生产的有色金属及合金粉末如铝粉、铜粉和镍粉等。

有色金属及合金的生产工艺主要是冶炼和加工两大过程。冶炼过程是利用有色金属及其化合物的特性来进行的，在制取金属的还原过程之前，原料经过高温或水溶液等预处理，还原方法有碳还原、氢还原、金属还原、热离解、电解还原等。10 种主要有色金属的冶炼方法见表 1.3-4。有色金属及合金的加工过程包括塑性加工、铸造和粉末冶金等，其产品包括多种塑性加工半成品（板、带、管、棒、型、线材等)、锻件、铸件及粉末制品等。

表 1.3-4　10 种主要有色金属的冶炼方法

金属	原料	预处理	金属还原	金属精炼
Al	氧化矿	碱浸，化学分离（Al_2O_3）	熔盐电解	三层熔盐电解
Mg	海水 氯化物 碳酸盐矿	化学分离（$MgCl_2$） 煅烧（MgO、CaO）	熔盐电解 真空下用 Fe - Si 还原	真空蒸馏 熔剂及 SF_6 处理

续表 1.3-4

金属	原料	预处理	金属还原	金属精炼
Ti	氧化矿	氯化，化学分离（$TiCl_4$） 化学分离（TiO_2）	镁还原 铝还原（Ti - Fe）	碘化法 电弧炉熔炼
Cu	硫化矿 氧化矿	焙烧，冰铜熔炼（$Cu_2\cdot S\cdot FeS$） 硫酸浸出（Cu_2SO_4）	转炉熔铜 用铁置换 电解	电解
Pb	硫化矿	焙烧，烧结（PbO）	碳还原 利用相互反应还原	电解 熔析 碱性精炼
Zn	硫化矿	焙烧，烧结（ZnO） 焙烧，硫酸浸出（$ZnSO_4$）	碳还原（蒸馏） 电解	精馏
Ni	硫化矿 氧化矿 混合硫化矿	升焙烧，冰铜熔炼，焙烧（NiO） 同硫化矿熔炼，还原焙烧 - 浸出 加压氨浸	碳还原（Ni，Ni - Fe） 电解 高压氢还原	电解 Ni（$CO)_4$ 的分解
Sn	氧化矿	焙烧	碳还原	电解
Sb	硫化矿	熔析（Sb_2S_3） 挥发焙烧（Sb_2O_3）	铁还原 碳还原	熔剂处理 电解
Hg	硫化矿		蒸馏（脱硫）	再蒸馏

有色金属及合金的应用包括结构材料应用和功能材料应用两方面。结构材料主要用于航空、航天、汽车、火车、船舶和建筑等领域。功能材料则主要用于信息产业、机电、能源等领域。有色金属及合金在国民经济各部门的广泛应用见图 1.3-5。

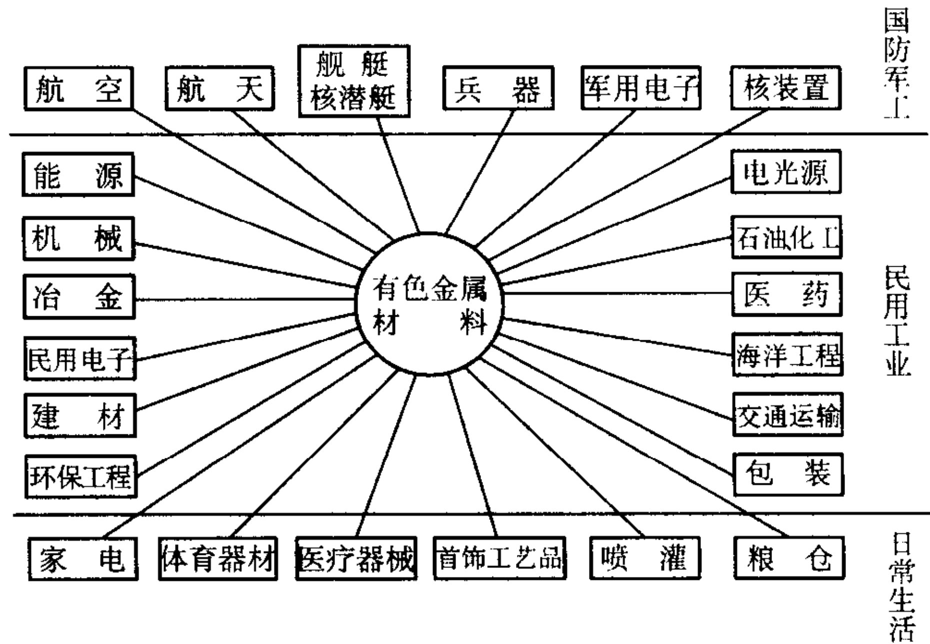


图 1.3-5　有色金属材料在国民经济各部门的广泛应用

3.2　产业现状

2003 年中国 10 种常用有色金属（铝、铜、镁、钛、锌、锡、锑、铅、钨、钼）的总产量达到 1 205 万吨，连续第二年跃居世界首位。我国十几年来 10 种有色金属总产量的增长情况见图 1.3-6。六种有色金属精矿产量实现恢复性生产增长，2003 年铜、锌、铅、镍、锡、锑六种精矿总产量（金属量）306 万吨，同比增长 9.7%，实现了恢复性增长。钨、

钼两种精矿产量小幅度增长（3%～4%），全年生产钨精矿折合量6.7万吨，钼精矿折合量6.8万吨。氧化铝、粗铜产量继续增长，2003年我国氧化铝总产量达609.4万吨，同比增长11.2%。由于电解铝投资过热，产能迅速膨胀，2003年总产能已超过830万吨，目前电解铝投资过热现象已基本得到扼制，但氧化铝缺口仍较大。2003年我国生产矿产粗铜130.8万吨，同比增长12.3%。铜材和铝材的产量大幅度增长，2003年铜材产量329.7万吨，同比增长32.6%；铝材产量363万吨，同比增长33.3%。我国铜材和铝材的产量已与占世界第一位的美国的产量基本相当。据不完全统计，我国铝及铝合金材有18个大类，200多种合金，2 400多个品种，13 630多个规格；铜及铜合金材有250多个牌号，1 000多个品种；镍材有7大类，18个牌号，50多个品种；锡材有7类，64个牌号；稀有金属材料有近500个品种，40 000多个规格，其中钛材有40多个牌号，数百种规格；难熔金属6个品种，有上百个牌号。硬质合金有8大类，320多个品种。

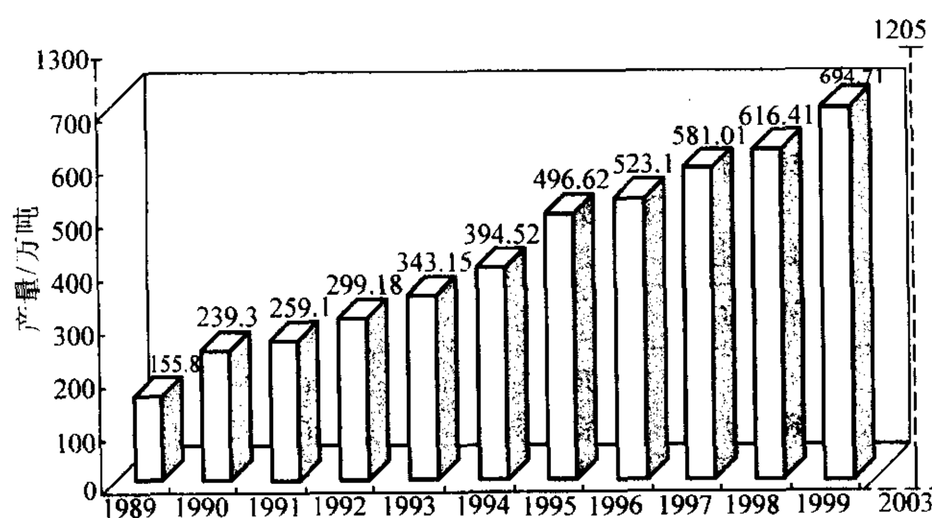


图1.3-6 我国十几年来10种有色金属总产量

3.3 发展趋势

有色金属材料及加工技术正朝着高性能、高质量、高效率和低成本的方向发展。

新型高性能铝合金是世界主要发达国家重点研究发展方向，如$\sigma_{0.2} \geqslant 500$ MPa，密度小于2.4 g/cm^{3}的高强可焊铝合金，$\sigma_{0.2} \geqslant 600$ MPa的超高强铝合金以及新型耐热铝合金和铝基复合材料等。铝合金加工技术正朝着连续化、短流程和近终形的方向发展，世界上最大的热轧机已达5 588 mm宽(美国)，热轧铝板尺寸最大可达5 000 mm×270 mm×3 000 mm。采用1～2台可逆热轧开坯机加上3～6机架热连轧机组成高度自动化、连续化的热轧生产线，这是当今铝板带箔材加工技术的主流方向，锭坯重达27 t，热轧速度为5～8 m/s，单位宽度卷重达10 kg/mm，卷径已达2 400 mm。有些采用计算机控制，带有X射线测厚、激光对中、VC辊等系统，热轧卷材厚公差小于0.75%。在冷轧方面，据不完全统计，全世界已有4～6辊大型单机架冷轧机（>1 400 mm）200余台，6机架以下冷连轧线数十条。连铸连轧技术迅速发展，包括连铸连轧、连续铸轧、连续铸挤和半固态复合铸轧等。最大挤压机为350 MN立式反向挤压机，可生产ϕ1 500 mm以上的管材。在重型锻压设备方面有俄罗斯的750 MN、法国的650 MN、美国的450 MN和中国的300 MN水压机。铝合金的熔炼炉已有150 t的圆形、可倾斜、可开盖、计算机控制的燃气炉。在国内也已研发了许多相当于2000系和7000系的新型铝合金，开展了B96、7055型的超高强铝合金研制，引进了6 t级Al－Li合金熔铸生产线。国内又装备了60 t级的熔炼炉，热效率达50%以上，熔化速率为9～10 t/h。大型热轧机有2 800 mm和2 000 mm，辊宽≥1 200 mm的大型现代四辊冷轧机有数十台，引进并自行设计制造连铸连轧机组数十台套，引进箔材生产线多条。采用先进技术的多机架（1+4）铝板热连轧生产线已经建设。我国自行研制成了万吨级油压双驱动挤压机可用于生产大截面铝型材和大直径管材。在国家“863”计划中也开展了快凝AlFeVSi系，过共晶系耐热铝合金研制以及纤维增强和颗粒增强铝基复合材料研究。在国家重大基础研究计划中，也列入“提高铝材质量的基础研究”项目，其中针对我国高硅含量铝土矿资源而自主研究的反浮选法取得了重要进展，此外在超常铸轧工艺基础和设备研制方面也取得了明显进展。

在铜合金材料技术方面，大规模集成电路用引线框架铜合金经过攻关已形成CuFeP、CuCrZr、CuNiSn、CuAg等系列产品，强度为纯铜的3～4倍，导电率可达60%～80%IACS。电解铜箔已达到6～9 μm的水平，新引进美国CB公司生产线，阴极辊ϕ2 134 mm×1 600 mm，可生产13～25 μm，宽1 500 mm的铜箔。采用当代国际先进技术的江西铜业公司铜箔厂和日本三井公司珠海铜箔厂已开工建设。研发成功了具有中国特色的同步齿环材料新系列HQC－1、HQC－2。正在开展特殊黄铜弹性材料、新型铜基形状记忆合金以及阻尼铜合金的研制。

在稀有金属材料技术方面，鉴于其与国防工业密切相关，是各先进国家竞相发展的关键材料。如$\sigma_b \geqslant 1$ 250 MPa的高强高韧钛合金，600℃以上的钛合金及钛铝金属间化合物，新型阻燃、耐蚀钛合金，高性能的锆合金、铌合金和钽钨合金以及高精度铍材等。稀有金属的加工技术正朝着优质和大型化方向发展。俄罗斯的钛锭尺寸已达到ϕ1 500 mm×4 000 mm，单锭重达30 t。美国的钛锭也已重达23 t。奥地利普兰西金属公司生产的最重钼锭已达6 t，最大钨锭重达1.2 t。先进的熔炼技术不断涌现，如日本的空心阴极等离子束炉已炼出重达3 t的扁锭，不经开坯可直接进行轧制。冷炉床熔炼技术已达商业水平，可熔炼重达25 t的均匀、无夹杂、高质量钛锭，日本已生产出470 mm×150 mm×2 000 mm扁锭。挤压的钛无缝管重达9.1 t，大尺寸钛厚板可达25 mm×1 200 mm×3 783 mm，钛箔材为0.002 5 mm×600 mm，钛丝可达ϕ0.002 mm。钽板尺寸可达（0.127～1.000）mm×1 092 mm×91 500 mm（钽卷），1 220 mm×3 660 mm的钽－铜和钽－钢复合板也已问世。

在稀土材料技术方面，中国已研究开发出具有世界先进水平和自主知识产权的北、南两大体系的全套生产技术流程。由于近年来高附加值的高纯单一稀土开发成功，使高纯单一稀土出口量和经济效益大幅度提高，年出口量达千吨，创汇达数亿美元。高技术稀土功能材料的研究和开发在国家“863”计划和“973”计划中都作了相应安排，如稀土永磁、荧光、发光、磁光、储氢以及超磁致伸缩等方面都已取得了明显进展，Ni－MH电池、NdFeB永磁和农用稀土方面的应用已形成一定规模。

在贵金属材料技术方面，日本和德国采用全浸式平辊拉丝工艺，美国采用液体金属挤压工艺生产大规模集成电路元件用ϕ18～25 μm的键合金丝。世界贵金属特种复合电接材料的尺寸精度已达公差0.05 mm，贵金属复合位置公差0.1 mm，覆层最小厚度已达1.5 μm的水平。我国集成电路用键合金丝已形成Y型，C型和FA型键合金丝年产2 000万米（200 kg）的能力，基本满足国内需求。贵金属复合材料尺寸精度已达到国际先进水平。已建立了贵金属浆料生产线，研发了多种气敏元件、汽车尾气净化用三元催化剂、钯基透气净化材料、AgMgNi长带、贵金属催化剂以及贵金属药物等，在电子、化工、汽车、能源、军工以及医药等工业中发挥了重要作用。

4　化工材料

4.1　分类概况

化工材料在这里主要是指高分子材料，即以高分子化合物为基础而制得的材料。按照来源高分子材料可以分为天然高分子、半合成高分子和合成高分子三大类。天然高分子材料包括棉、麻、木、蚕丝、天然橡胶等，半合成高分子包括硫化胶、醋酸纤维、黏胶纤维等，合成高分子材料包括聚乙烯、聚丙烯、尼龙、聚酰亚胺、环氧树脂等。高分子是由成千上万个小分子单体通过加聚或缩聚反应以共价键结合起来的长链分子，高分子材料是由大量的单个长链高分子聚集而成。由于长链高分子所固有的特点（分子量高、分子量的不均匀性、高分子链结构及其柔顺性等），高分子材料与金属及陶瓷相比具有质轻、价廉、易加工成形等许多特色与优点，并在人类的衣、食、住、行和信息、能源、农业、航空航天等许多工业领域和农业领域获得了广泛的应用。

按照应用，高分子材料还可以分成塑料、橡胶、纤维、涂料、胶粘剂和功能高分子材料等。

1）合成树脂及塑料　在一定意义上说，塑料、树脂和聚合物是同义词，即以合成树脂或化学改性的天然高分子为主要成分，加入填料、增强剂和其他添加剂，在一定压力和温度下成形的高分子材料。塑料的分子间次价力（8.7～21 kJ/mol），模量（150～3 500 MPa）和形变量（小于0.5%～800%）等介于橡胶和纤维之间，T_m 和 T_g 可在很大范围内变化。根据合成树脂在加工时的行为，又可以将其分为热塑性树脂和热固性树脂。热塑性树脂是线型或支化的高分子，可以重复地受热变软、冷却变硬，在受热受压时保持其化学本性，如聚乙烯、聚丙烯、聚苯乙烯和聚氯乙烯等。热固性树脂是在加工过程中从线型结构变成体型结构，具有不熔、不溶的特性，由于具有紧密连接的交联结构，热固性树脂能比大多数热塑性树脂承受较高的温度，具有较大的尺寸稳定性。热固性树脂一般不能单独作为结构材料使用，而必须采用玻璃纤维或碳纤维来进行增强而成为复合材料后使用。热固性树脂有氨基树脂、酚醛树脂、不饱和聚酯、环氧树脂、氰酸酯树脂、聚酰亚胺树脂、有机硅树脂等。热塑性树脂按照其性能和应用情况还可以分为通用塑料、工程塑料和高性能工程塑料。通用塑料如聚乙烯、聚丙烯、聚氯乙烯、聚苯乙烯、聚氨酯、酚醛塑料等，它们用途广、产量大、价格相对低廉；工程塑料如聚碳酸酯、聚甲醛、聚酰胺、ABS树脂、聚苯醚、聚酯等，其机械强度和耐热性较好，适用于工程结构、机械部件和化工设备等工业用途；高性能工程塑料如热致液晶聚合物、聚苯硫醚、聚醚砜、聚砜、聚醚醚酮、聚醚酰亚胺、聚酰胺亚胺等，其特点是机械强度高，抗蠕变，连续使用温度在150℃以上，热变形温度高，耐磨损，对大多数溶剂稳定。塑料具有许多独特的优良性能，如密度小、比强度高，可代替木材、水泥、砖瓦等大量用作建筑材料；耐化学腐蚀性优良，可制作化工设备；电绝缘性和隔热性好，可用以制造电子元器体，其性能可与陶瓷、橡胶相比；摩擦因数小，耐磨性好，有消声减振作用，可部分代替金属制造轴承和齿轮，并可在无润滑的条件下高速运转；易于加工和着色，采用不同的原料和加工方法，可制得坚韧、刚硬、柔软、轻盈、透明的各种制品，广泛用于日常生活、包装材料、农用薄膜、管材和零配件。纤维增强树脂塑料还是航空航天等国防工业关键材料。

2）橡胶或弹性体　是一种具有弹性的高分子材料，其分子链具有较高柔性；分子间次价力小，约8.4 kJ/mol；玻璃化温度低，约－110～－40℃；橡胶应在其玻璃化温度以上使用。橡胶的分子链可以交联（硫化），交联后的橡胶受外力作用发生变形时具有迅速复原的能力。橡胶具有良好的物理力学性能和化学稳定性。按来源，橡胶分为天然橡胶和合成橡胶。目前合成橡胶的产量和消耗量远大于天然橡胶。首先产量最大的是丁苯橡胶，约占合成橡胶的50%；其次是顺丁橡胶，约占15%；此外还有耐老化的乙丙橡胶、耐油的丁腈橡胶、耐燃的氯丁橡胶和气密性好的丁基橡胶等。近年来还发展了由高分子链间的作用力进行物理交联的热塑性弹性体，如SBS橡胶（苯乙烯－丁二烯－苯乙烯三嵌段共聚物），它可用塑料加工方法模塑成各种形状的橡胶制品。橡胶广泛应用于制造轮胎、胶管、胶带、电缆等制品。

3）纤维　呈纤维状、可制成纺织品或用作复合材料增强体的高分子材料。分为天然纤维和合成纤维两大类。蚕丝、棉、麻等是天然纤维；以天然高分子或合成高分子为原料，经过纺丝和后处理制得的纤维称合成纤维，如聚丙烯腈纤维（腈纶）、聚丙烯纤维（丙纶）、聚对苯二甲酸乙二酯纤维（涤纶）、聚酰胺纤维（尼龙66、尼龙6等）、聚芳酰胺纤维（芳纶）和超高模量聚乙烯纤维等。纤维的基本特征是模量高、强度高、结晶度高、内聚能高，并能通过取向获得沿纤维轴向的最佳性能。制造合成纤维的成纤高聚物应有一定的稳定性，在熔融时不分解或能在溶剂中溶解而形成浓溶液，具有良好的成纤行为。成纤高聚物具有线性结构，并具有适宜的分子量，分子量分布窄、支链少。合成纤维既用于衣着，也用于工程，按其工程用途可分为分离用纤维、超吸水纤维、生物降解纤维、光导纤维、离子交换纤维等。

4）涂料　是能涂敷于底材表面并能形成坚韧连续的涂膜的液体或固体高分子材料。涂料由成膜物质（即基材）、颜料、填料、涂料助剂，功能组分及分散介质（有机溶剂或水）所组成。涂料可分为油脂涂料、天然树脂涂料和合成树脂涂料。聚合物涂料已经从金属防腐涂料发展到红外与电磁波隐身涂料、微波屏蔽涂料、防核辐射和防核污染涂料、高温抗氧化涂料、高温隔热涂料、温控涂料等。

5）胶黏剂　又称黏合剂，是能使物体的表面与另一物体表面结合在一起的材料。以高分子化合物为主体的胶黏材料称为高分子胶黏剂，按其来源可分为天然胶黏剂和合成高分子胶黏剂。按化学组成可分为热固性树脂胶黏剂、热塑性树脂胶黏剂和橡胶型胶黏剂。合成胶黏剂一般是多组分的混合物，除聚合物外，尚需加入固化剂、增韧剂、无机填料和溶剂等。用于非承力结构部件之间结合的称为非结构胶黏剂，而必须具有足够高的强度、韧性、疲劳性能和对环境作用具有稳定性的胶黏剂则称为结构胶黏剂。此外，有些胶黏剂还必须具有导热、导电、折射率匹配和真空密封等作用。

6）功能高分子材料　具有特殊的化学反应性、或具有特殊的物理、生物、药物和其他用途的高分子材料的统称。如催化材料能有选择性地结合特殊物质（高分子吸附剂）；捕获或输运电荷或能量的导电高分子和压电高分子材料；能将光转换成电荷载体或将电荷载体转换成光的光敏高分子材料和发光高分子材料；以及能将药物输送到特定组织细胞的控释和靶向高分子药物体系等。

7）主要高分子材料的名称与缩语

名称	缩语	名称	缩语
聚乙烯	PE	聚对苯二甲酸乙二酯	PET
聚丙烯	PP	聚对苯二甲酸丁二酯	PBT
聚苯乙烯	PS	聚砜	PSF
聚氯乙烯	PVC	聚醚醚酮	PEEK
聚丙烯腈	PAN	聚氨酯	PU
聚乙烯醇	PVA	聚酰亚胺	PI
聚四氟乙烯	PTFE	天然橡胶	NR
聚酰胺	PA	聚丁二烯橡胶	BR
聚碳酸酯	PC	丁苯橡胶	SBR
聚甲醛	POM	异戊二烯橡胶	IR

聚苯醚	PPO	乙丙橡胶	EPR
聚苯硫醚	PPS	氯丁橡胶	CR
聚醚砜	PES	丁腈橡胶	NBR
丁基橡胶	IIR		

4.2 产业现状

新中国成立以后，特别是改革开放20余年来，我国的高分子化工材料从小到大、从无到有，形成了一个相对完整的工业体系。到20世纪末，我国三大合成材料（合成树脂、合成纤维和合成橡胶）的产量均已进入世界前五位。其中合成纤维产量已跃居世界首位，达到549.27万吨（1999年），2003年更急剧增长到1 069.20万吨；合成树脂是石油化工最重要的下游产品，是三大合成材料中产量最大的品种，1999年我国合成树脂及塑料产量为851万吨，其中五大通用树脂（聚氯乙烯、聚乙烯、聚丙烯、聚苯乙烯和丙烯腈-丁二烯-苯乙烯ABS）产量达802.62万吨，2003年合成树脂总产量已高达1 593.80万吨；合成橡胶1999年我国生产了76.11万吨，而2003年则已产127.20万吨。我国三大合成材料近几年来的产量见表1.3-5（根据有关行业协会的统计）。

表1.3-5 我国三大合成高分子材料的近年产量

万吨

品种 \ 年代	1999	2001	2002	2003
合成树脂及塑料	851	1 209.90	1 366.46	1 593.80
合成橡胶	76.11	104.68	116.76	127.20
合成纤维	549.27	757.96	915.16	1 069.20

除三大合成高分子材料外，我国的涂料、胶黏剂、催化材料、微电子化工材料和新型化工材料也有迅猛的发展。

4.3 发展趋势

国民经济增长和社会发展对化工材料的需求十分强劲，国内市场潜力很大。据有关行业协会预测，到2010年我国三大合成高分子材料的产量将以略高于我国GDP的速度继续增长，合成树脂的产量可能达到3 000万吨以上，合成橡胶的产量可能达到200万吨以上，而合成纤维的产量可能达到1 500万吨以上。今后的发展趋势是要进一步调整产品结构，解决“瓶颈”产品问题，增加产品牌号，提高产品质量，提高国内市场占有率，充分消化吸收引进技术，努力创建具有中国特色的名牌品种，打入国际市场。

5 建筑材料

5.1 分类概况

建筑材料是基本建设和建筑业的物质基础。市政建设、水利水电建设、港口、铁路、机场、农田水利建设和国防基础建设等都是由大量的水泥、混凝土及其制品、石灰、砖瓦、玻璃、建筑卫生陶瓷和新型墙体材料等建筑材料构筑而成的。建筑材料的主要用途见图1.3-7。建筑材料在这里包括：①水泥及水泥制品；②平板玻璃；③建筑卫生陶瓷；④新型建筑材料；⑤无机非金属材料；⑥非金属矿物材料。

1）水泥及水泥制品　我国已研制成功并投入生产的水泥品种迄今已有60余种，其中大量生产的有20余种，是世界上水泥品种最多的国家之一。按我国GB/T 4131—1997标准，水泥分为通用水泥、专用水泥和特性水泥三大类。通用水泥包括硅酸盐水泥、普通硅酸盐水泥、矿渣硅酸盐水泥、火山灰硅酸盐水泥、粉煤灰硅酸盐水泥和复合硅酸盐水泥6个品种。专用水泥和特性水泥也合称为特种水泥，按水泥

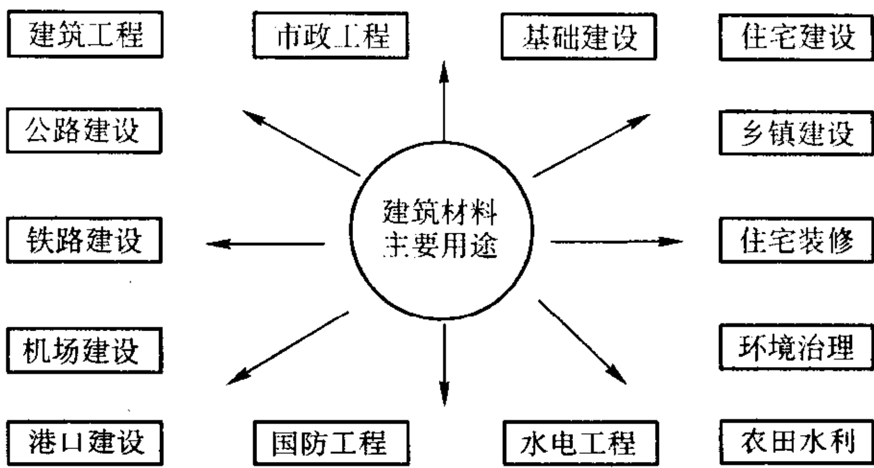


图1.3-7 建筑材料的主要用途

主要矿物所属体系可分为硅酸盐、铝酸盐、硫铝酸盐、铁铝酸盐、氟铝酸盐和其他6个体系；特种水泥按其功能和用途可以分为快硬早强水泥、低水化热水泥、膨胀和自应力水泥、油井水泥、耐高温水泥、装饰水泥和其他水泥七大类。

2）平板玻璃（建筑玻璃）　传统建筑玻璃与新型建筑玻璃之间并无明确的界限，通常用于民用建筑的2 mm和3 mm厚的普通平板玻璃称为传统建筑玻璃；而具有隔音、隔热、透光、多色、异型等特殊功能的玻璃则称为新型建筑玻璃。

3）建筑卫生陶瓷　建筑卫生陶瓷包括建筑陶瓷（釉面砖、彩釉砖、瓷质砖、锦砖、劈离砖、广场砖、琉璃制品等）和卫生陶瓷（坐便器、蹲便器、水箱、洗面器、小便器、净身器、洗涤槽等）两大类。建筑陶瓷是房屋建筑内外墙面、地面的重要装饰材料，卫生陶瓷是现代房屋卫生间中不可缺少的主要部体。建筑卫生陶瓷产品具有无毒、无味、防火、防老化、耐腐蚀、耐热、易于清洁等特性。

4）新型建筑材料　新型建筑材料是现代建筑工程的重要物质基础，主要包括新型墙体材料、新型防水密封材料、新型保温隔热材料、新型建筑装饰装修材料等。

5）无机非金属材料　无机非金属材料工业是我国四大材料工业之一（钢铁、有色金属、有机非金属材料和无机非金属材料）。严格地讲，它已经超出了建筑材料的狭义范畴。无机非金属材料主要包括玻璃纤维和特种纤维、玻璃钢/复合材料、特种玻璃、加工玻璃、石英玻璃、特种陶瓷、人工晶体、特种密封材料和特种胶凝材料八大类产品，在国防建设和国民经济建设中有着广泛的应用，见图1.3-8。

6）非金属矿物材料　是指非金属矿物经深加工处理后产出的一类新型矿物材料。非金属矿物材料由于矿物品种很多，目前已发现的矿种有300多个，已被利用的有100多个。非金属矿物材料的用途，主要取决于它的矿物特性、化学组成、结构构造特征以及物理、化学特性等，如沸石因其特有的孔径结构（分子筛）常常是石油工业、化学工业不可缺少的矿物原料；硅藻土因其具有特殊的多孔形态又是固液分离、乙炔气提纯以及水体净化等优良的助滤材料；高岭土、滑石、方解石（重钙）等是塑料、橡胶、油漆、造纸、涂料等工业不可缺少的功能性填料。非金属矿物材料的主要应用领域见图1.3-9。

5.2 产业现状

建国50年来，尤其是改革开放20多年来，中国建材工业从小到大，从无到有，目前已经形成了比较完整的建材工业体系，中国主要建材产品水泥、平板玻璃、建筑卫生陶瓷等的产量近年来一直居世界第一位。表1.3-6是中国主要建材产品近年来的产量增长情况。图1.3-10是1998年中国建材工业主要产品占世界总产量的比重。

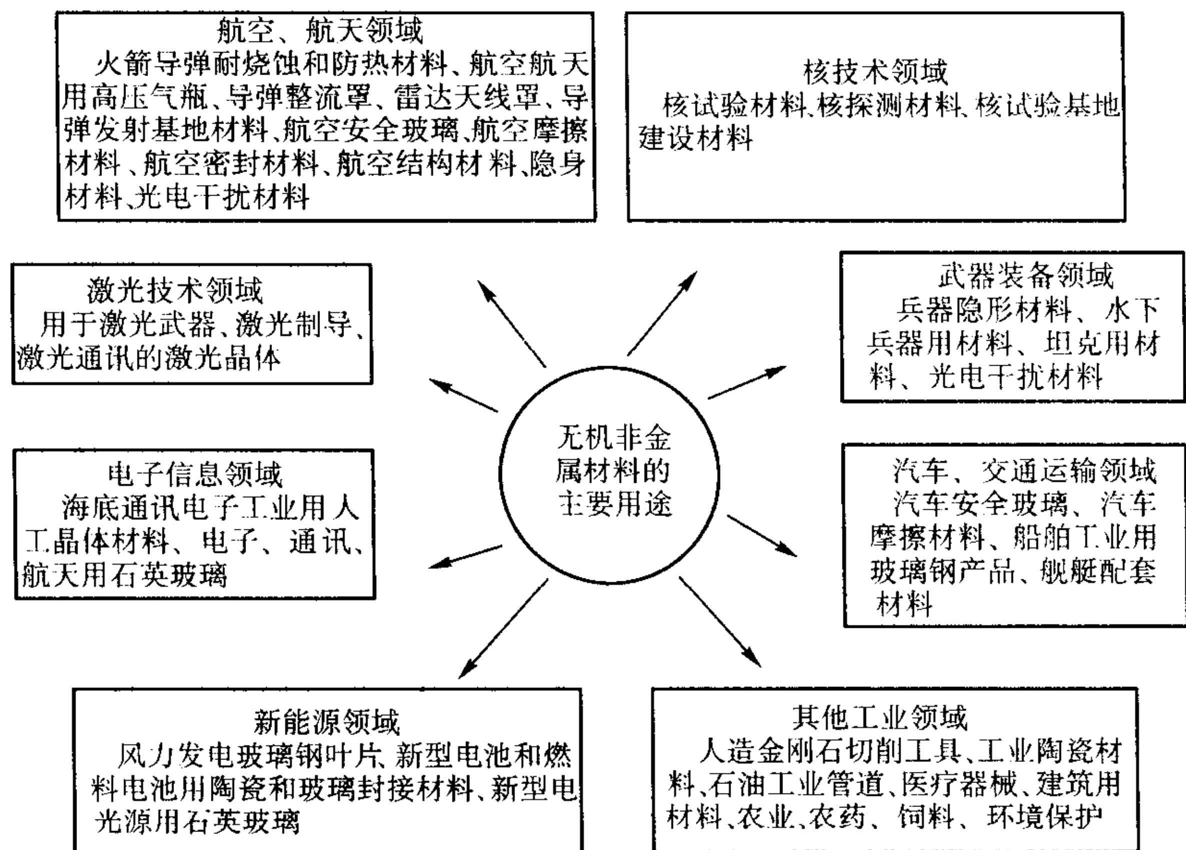


图 1.3-8　无机非金属材料的主要应用领域

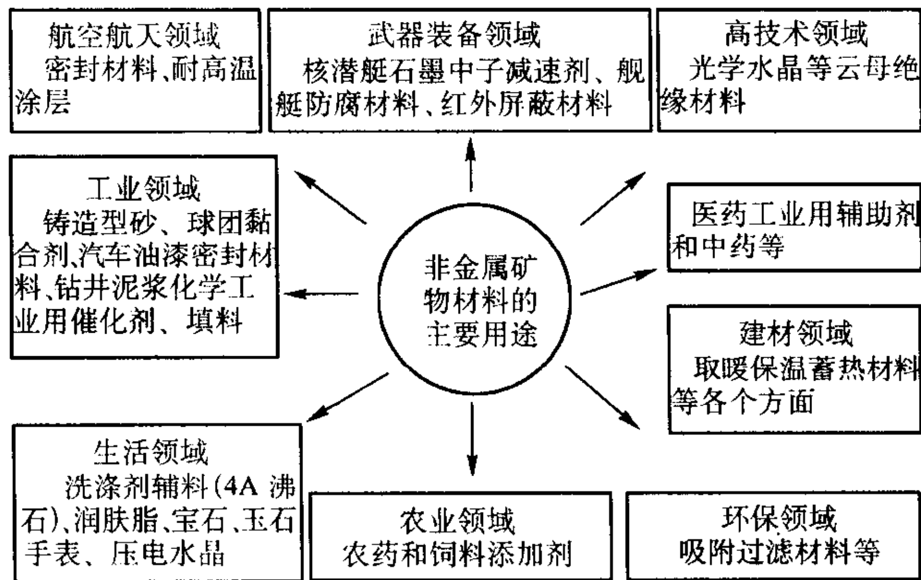


图 1.3-9　非金属矿物材料的主要应用领域

5.3　发展趋势

中国是世界建材生产大国，但还不是强国，主要差距是劳动生产率低（见表 1.3-7）、能耗高（见表 1.3-8）以及高技术含量低。

在水泥和混凝土方面要研究开发大型和新型生产技术和装备，重点是建设日产 5 000 t 以上的超大型新型干法水泥生产线的烧成工艺及技术装备，必须用现代高新技术来改造传统的占水泥产量 80% 以上的立窑与小回转窑；要优化硅酸盐水泥熟料矿物组成，开发低钙水泥新品种；要对水泥生产原燃料资源合理利用与开发，研究水泥生态生产技术；研究开发高性能混凝土、智能型混凝土、混凝土外加剂和先进水泥基材料。

表 1.3-6　我国建材主要产品增长情况

主要建材产品	年份							
	1980	1985	1990	1995	2000	2001	2002	2003
水泥/万吨	7 986	14 595	20 971	47 591	59 700	66 100	72 500	86 300
平板玻璃/万箱	2 466	4 942	8 067	15 723	19 101	20 964	22 800	25 200
卫生陶瓷/万件	292	775	1 783	5 448	5 500	6 000	6 500	7 000
建筑陶瓷/万平方米	1 178	5 641	17 920	158 410	190 000	200 000	210 000	250 000
墙体材料（折合亿块标准砖）	1 537	2 942	4 485	7 691				9 500

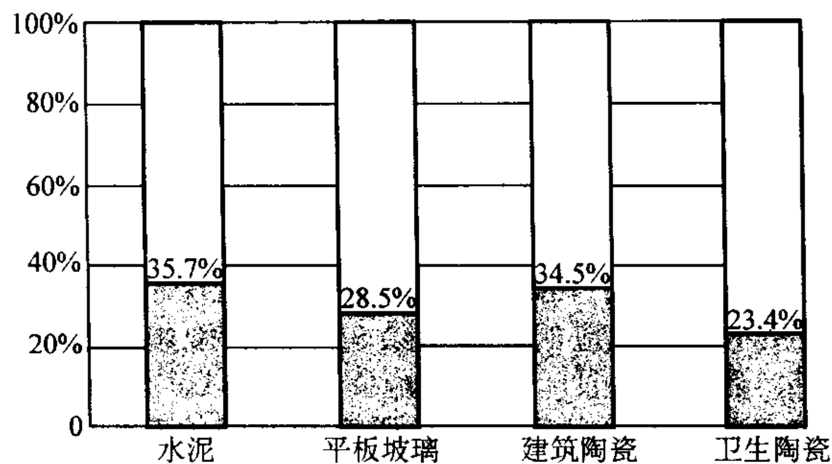


图 1.3-10　1998 年中国建材工业主要产品占世界总产量的比重

在平板玻璃方面，要进一步提高和完善“中国洛阳浮法”玻璃生产技术，积极发展加工玻璃产品和生产技术与装备，研究开发富氧和全氧燃烧技术。

在建筑卫生陶瓷方面，要发展装饰技术，提高产品的质量和档次，积极发展节能、节水和生态环保型产品和工艺技

表 1.3-7　中国建材主要产品的劳动生产率与国际水平的比较

产　品	单　位	中　国	国外先进水平	中国与国外比较
水泥	t/（人·a）	200	3 000	差距 15 倍
平板玻璃	重量箱/（人·a）	700 ~ 1 200	8 000 ~ 10 000	差距 5 ~ 12 倍
建筑陶瓷	m^2/（人·a）	4 000 ~ 6 000	10 000 ~ 50 000	差距 2.5 ~ 10 倍
卫生陶瓷	t/（人·a）	3 ~ 17	120 ~ 180	差距 10 ~ 40 倍
砖	万块/（人·a）	一般≤10 最高 30	200 ~ 400	差距 20 ~ 40 倍
玻璃纤维（坩埚法）	t/（人·a）	2 ~ 4		差距 15 ~ 110 倍
玻璃纤维（池窑）	t/（人·a）	20	30 ~ 220	差距 1.5 ~ 11 倍

表 1.3-8 中国主要建材产品能源消耗与世界水平的比较

项目	单位	中国	国外先进水平	中国与国外比较
水泥熟料烧成能耗	kg 标煤/t 熟料	175.3	117	高 50%
平板玻璃综合能耗	kg 标煤/重量箱	33.6	20	高 68%
建筑陶瓷综合热耗	kg 标煤/m^2	2.5 ~ 15	0.77 ~ 6.42	平均高 150%
卫生陶瓷综合热耗	kg 标煤/t	400 ~ 1 800	238 ~ 467	平均高 200%

术，研究开发新型成型技术，如高、中压注浆成型技术等，可显著提高卫生陶瓷产品质量和档次。

在新型建筑材料方面，迈入新世纪以来，新型建筑材料应按照中国新世纪建筑发展的总体战略目标和技术政策，以提高人民生活质量为目的，适应新世纪建筑结构体系的发展，满足建筑节能的发展需求，重点发展节能、节土、利废、轻质高强与多功能的新型墙体和屋面材料、高效绝热材料、中高档防水密封材料以及高档新型装饰装修材料等。

在无机非金属材料方面，要深入开展材料结构与成分、合成与加工、性能、工程技术、应用技术直至产业化全过程的研究。无机非金属材料的品种、性能和质量应满足军工配套及高技术发展的需要，要加快一批无机非金属新材料的产业化进程，使其成为建材工业的新的经济增长点，提高国际市场竞争力。

在非金属矿物新材料方面，要加强矿物性质基础和深加工技术研究，研究开发非金属矿物新材料，为高性能摩擦材料、高性能密封材料、高性能石墨功能材料、环境保护及废物处置用非金属矿物材料以及农牧业用非金属矿物材料及医药与保健用非金属矿物材料等。

6 电子信息材料

6.1 分类概况

20 世纪以来，信息技术是靠电子学技术和微电子学技术发展起来的，如通信是从长波到微波，存储是从磁芯到半导体集成，运算是从电子管发展到以大规模集成电路为基础的电子计算机等。所以，目前论及信息技术都称为电子信息技术，其特征是信息的载体为电子。所以电子技术，特别是微电子技术，仍然是当前信息技术的主要支撑技术。但随着高容量和高速度信息的发展，已显示出电子学（electronics）和微电子学（microelectronics）具有局限性，信息的载体必然由电子发展到速度要快得多的光子，目前信息的探测、传输、存储、显示、运算和处理已由光子和电子共同参加来完成，从而产生了光电子学（optoelectronics）。继光电子学后，光子学（photonics）技术正在兴起。从电子学到光电子学和光子学是跨世纪的发展（见图 1.3-11）。从材料的角度来说，微电子材料是最重要的信息材料，光电子材料是发展最快的信息材料，而光子材料是最有前途的信息材料。

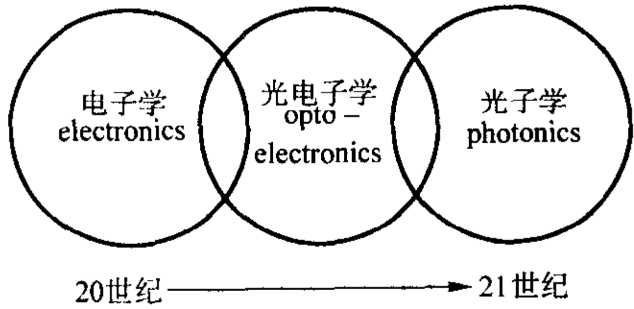


图 1.3-11 从电子学到光子学的发展

电子信息材料是指在微电子技术，光电子技术和新型电子元器件基础产品领域中所用的材料，主要包括微电子材料、光电子材料、磁性材料、电子陶瓷材料、绿色电池材料等。这些基础材料支撑着通信、计算机、信息家电与网络技术等现代信息产业的发展。在所有电子信息材料中，最基础也是最重要的材料是半导体材料，半导体材料按其组成可分为元素半导体和化合物半导体两大类。化合物半导体包括有机半导体。在元素半导体与化合物半导体之间还可形成组分连续可变和性能可控的固溶体。典型的半导体材料以共价结合为主。从结构上讲，元素半导体和Ⅲ - Ⅴ族化合物半导体材料分别具有金刚石和闪锌矿结构，而其他的半导体材料一般为立方或六方结构。此外，还有非晶和微晶半导体以及人工设计并制造出的半导体超晶格材料。

1）元素半导体　现已被确认具有半导体性质的元素有 Si、Ge、B、Se、I 以及 C、As、Sb、Al、S 的某种同素异构体。元素半导体 Si 和 Ge（特别是 Si），是研究得最多、应用最为广泛的半导体材料。

2）无机化合物半导体　已查明的无机化合物半导体材料约 600 多种。许多化合物，如Ⅲ - Ⅴ、Ⅱ - Ⅵ、Ⅳ - Ⅵ族化合物和多元化合物等，都具有半导体性质，如 GaAs、GaN、ZnS、SiC、Bi_2Te_3、ZnO、Cu_2O 等。

3）固溶半导体　指具有半导体性质的固态溶体。在固溶体中各组元的原子排列成点阵（结构有序），各种原子随机地分布在格点上（组分无序）。这类材料的特点是其性质随成分的不同而变化。因此可得到性质更为多样的半导体材料以满足各种实际需求。现已知存在大量的固溶体材料。其中，有元素之间形成的固溶体如 GeSi，Ⅲ - Ⅴ族固溶体如 AlGaAs，Ⅱ - Ⅵ族固溶体如 HgCdTe，Ⅳ - Ⅵ族固溶体如 PbSeTe，以及其他一些固溶体如 AlGaInP、GaInAsP、GaInAsSb 等。

此外，一些半导体可以与磁性半导体形成固溶体，例如 HgMnTe 称为稀磁半导体。

4）半导体微结构材料　用外延方法周期性地生长极薄的，成分不同的固溶体或掺杂不同的半导体，就可以使半导体内的势具有一个周期性的变化，这种一维周期性结构称为超晶格。1970 年后，进行了大量的研究工作，开发了多种超晶格材料，如 AlGaAs - GaAs、SiGe - Si 等。

5）微晶和非晶半导体　晶粒细小或原子排列短程有序、长程无序的半导体材料。

6）有机半导体　很多有机化合物具有半导体性质。它们包括分子晶体、电荷转移复合物以及高分子材料，如萘、蒽、聚丙烯腈、酞菁和一些芳香族化合物等。用有机半导体材料制备发光、MOS 以及存储器件具有广阔的应用前景。

6.2 产业现状

在工业上已获得应用或有重要意义的半导体材料有 Si、Ge 等元素半导体，GaAs、GaP、InP、CdTe 等化合物半导体和 GaAsP、AlGaAs、HgCdTe、GaInAsP、GaInAsSb、SiGe 等固溶半导体。批量供应的外延片除 Si、GaAs 和 GaP 的同质外延片外，还有一些Ⅲ - Ⅴ族固溶体（如 GaAsP）、Ⅱ - Ⅵ族固溶体（如 HgCdTe）等。正在研究开发中的有Ⅳ - Ⅵ族、Ⅱ - Ⅵ族的量子阱超晶格材料，以及金刚石、SiC、ZnSe 等薄膜材料。非晶硅薄膜材料已大批量生产。图 1.3-12 表明，2001 年，约 210 亿美元的半导体材料支撑了约 8 790 亿美元的电子设备市场，2004 年有约 280 亿美元的材料支撑着约 9 900 亿美元的电子设备市场。以半导体材料为基础的信息产业已成为世界上最庞大的产业之一。

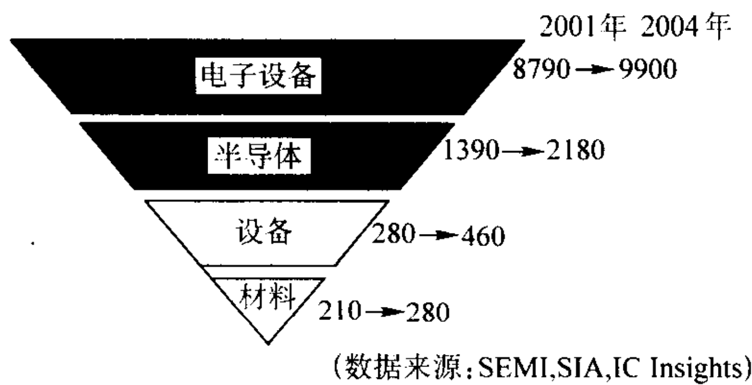


图 1.3-12 世界材料、设备、半导体及电子设备市场
(单位:亿美元)

近 10 年来，随着计算机、通信、信息家电及数字化技术的快速发展，带动了微电子、光电子和元器件的发展，同时也为我国电子基础材料的发展创造了良好的机遇，半导体材料、光电子材料、覆铜板材料、压电晶体等快速增长，出口不断增加。磁性材料、电子陶瓷材料等稳步增长，目前我国电子信息材料整体上取得较大进步，建成了相应的研发中心和生产基地，逐步在国际上占有一定地位。根据有关行业协会统计和预测，我国主要电子信息材料产业规模与市场测算见表 1.3-9。

表 1.3-9 主要电子信息材料产业规模与市场预测

材料类别	单位	目前生产能力	2004 年需求测算
多晶硅	t/年	100	约 1 900
单晶硅	t/年	1 000（太阳能单晶占约 60%）	约 1 500
其中，硅抛光片(4 in、5 in)	万片/年	约 1 200	约 1 300
硅抛光片(6 in、8 in)	万片/年	约 500	约 300
硅外延片(4 in、5 in)	万片/年	300	约 400
硅外延片(6 in、8 in)	万片/年		约 100(含少量 12 in)
4 in、6 in 砷化镓单晶片	万片/年	10	10
高纯试剂	t/年	5 000	5 000
光刻胶	t/年	150	约 200
其中，紫外正胶、负胶	t/年	约 140	150
G 线、I 线光刻胶	t/年	10	50
深紫外光刻胶(248 nm)	t/年	0.001	0.005
深紫外光刻胶(193 nm)	t/年		

续表 1.3-9

材料类别	单位	目前生产能力	2004 年需求测算
环氧模塑料	t/年	15 000	30 000
键合金丝	kg/年	4 000	约 4 000
引线框架铜带	t/年	5 000	20 000
Nd·YAG 激光晶体	万元/年	6 000	1 000 ~ 2 000
液晶材料	t/年	TN·STN 30	约 50
	t/年	TFT 用液晶材料进口	
LCD - ITO 导电玻璃	万片/年	14 in × 14 in 大于 500	约 400
光纤材料	万公里/年	3 480	约 1 200
光纤预制棒	万公里/年	1 900	约 1 300
磁性材料			
其中，高性能永磁铁氧体	t/年	20 000	约 40 000
高性能软磁铁氧体	t/年	约 10 000	约 30 000
高性能稀土永磁	t/年	4 000	约 6 000
电子陶瓷材料			
其中，热敏元件瓷料	t/年	2 000	约 3 500
微波介质陶瓷	t/年	500	约 1 000
MLCC 瓷料	t/年	600	约 3 000
圆片容器等各种瓷料	t/年	2 000	约 5 000
压电晶体材料			
其中，人造石英晶体	t/年	1 760	约 2 000
铌酸锂、钽酸锂晶体材料	t/年	约 60	约 60
覆铜板材	t/年	210 000	约 300 000
铜箔	t/年	15 000	约 30 000

6.3 发展趋势

我国电子信息材料虽然已取得了长足的进步，但从总体上看仍然存在产品档次不高、品种规格不全、规模效益较差、关键产品依赖进口等问题。据有关行业协会的分析，到 20 世纪末我国主要电子信息材料国内外技术水平比较见表 1.3-10。21 世纪初国内外电子信息材料的发展趋势如下。

表 1.3-10 主要电子信息材料国内外技术水平比较（20 世纪末）

材料名称	主要用途	国外水平	国内水平
单晶硅	集成电路、半导体器件、太阳能电池	研制水平为 ϕ400 mm，主流产品为 ϕ200 mm	已具备 ϕ200 mm 小批量生产能力，但目前只能作陪片使用，实用化水平为 ϕ100 ~ 150 mm
GaAs 单晶	微波功率、低噪声器件，微波毫米波单片集成电路（MIMIC）和移动通信手机	研制水平为 ϕ150 ~ 200 mm，ϕ100 mm 和 ϕ150 mm Si - GaAs 单晶已分别达到商品化和批量生产	已研制出 ϕ100 mm Si - GaAs 单晶样品，实用化水平为 ϕ50 ~ 75 mm，抛光晶片达不到“开盒即用”
InP 单晶	3 mm 或更高频的微波毫米波器件，1.35 μm 和 1.55 μm 长波长光纤通信器件及光电集成电路（OEIC）	实验室研制出 ϕ100 mm 样品，ϕ75 mm 晶片达商品化，ϕ50 ~ 75 mm 抛光片已达到外延“开盒即用”	已研制出 ϕ50 ~ 75 mm 单晶，实用化水平为 ϕ50 mm 抛光片，但未达到外延“开盒即用”

续表 1.3-10

材料名称	主要用途	国外水平	国内水平
液晶材料	TN－LCD，STN－LCD，TFT－LCD在计算机、电视机、摄像机、仪器仪表等领域有重要应用	苯基环己烷类、环己基环己烷类、嘧啶类、手征性液晶的光学各向异性、介电各向异性等性能优异，黏度低，达到实用化要求并已实现规模化生产	只能提供用于TN－LCD的酯类、联苯类液晶、STN和TFT显示用液晶处于研究阶段
磁性材料：稀土永磁材料、永磁铁氧体、软磁铁氧体材料 磁记录材料：金属磁粉	磁盘驱动器电机、音响器件、开关电源、电子镇流器、显示器偏转线圈、微波磁控管、行波管、返波管、磁盘、磁带等；在移动通信、计算机、汽车工业和家用电器等领域有重要应用	烧结NdFeB已开始大量生产最大磁能积 $BH_{max}=413\ kJ/m^3$ 的永磁合金；国外已大量生产 FB_4，FB_5，FB_6 产品，BH_{max} 达 416 kJ/m^3；PC44级材料1992年就开始大批量生产；颗粒超细均匀，性能稳定，工艺性好，规模生产	大量生产的是最大磁能积 $BH_{max}=199\sim278\ kJ/m^3$ 的中低档NdFeB材料，只有少数先进工厂开始批量生产 $BH_{max}=334\sim358\ kJ/m^3$ NdFeB材料，只有部分工厂能生产 FB_4 系列产品；大量生产的是中低档产品、相当于PC44级的材料尚待开发和生产；实验室制出样品，颗粒分布不匀，工艺性不好，不能规模生产
电子陶瓷材料	陶瓷电容器、热敏电阻器、压敏电阻器、压电频率元件及微波器件的主结构材料，在移动通信、计算机、汽车电子、家电等领域有广泛、重要的应用	瓷料超细、均匀、性能稳定、规模生产，产量近3万吨，陶瓷元器件片式化率达70%以上	瓷料一致性、稳定性差，生产规模小，国内总产量仅为1 590 t；热敏电阻器、压敏电阻器、压电频率元件尚未片式化，陶瓷电容器片式化率仅为39%；高性能微波陶瓷与器件尚未商品化
绿色电池材料：负极材料MCMB	移动通信手机、笔记本电脑电池	比容量的研制水平和实用化水平分别为330 mA·h/g和320 mA·h/g	比容量的研制水平和实用化水平分别为300 mA·h/g和260mA·h/g

6.3.1 半导体材料

1) Si、GaAs、InP等半导体单晶材料向着大尺寸、高均匀性、高完整性方向发展，如国外硅晶片以 $\phi200$ mm为主流产品，$\phi300$ mm硅片开始上市，GaAs$\phi100$ mm已进入大批量生产阶段，并且正在向 $\phi150$ mm生产线过渡，同时对单晶电阻率均匀性、金属杂质含量、微缺陷、位错密度、晶片平整度、表面洁净度等提出了更加苛刻的要求。晶片大尺寸化也是光刻线条变细、集成度提高、提高成品率、降低成本的必然趋势。

2) 在以Si、GaAs为代表的第一代、第二代半导体材料继续发展的同时，第三代半导体材料——宽禁带半导体材料SiC、GaN、ZnSe、金刚石材料发展迅速，成为今后半导体材料的重要发展方向。另外，采用SiGe/Si、SOI等新型硅基材料使得原有硅集成电路性能有很大提高，也是今后的重要发展方向。

3) 半导体外延材料继经典的同质结、异质结之后，又向着量子阱、量子线、量子点方向发展。分子束外延(MBE)和金属有机化合物化学气相外延(MOCVD)技术的应用更为广泛。

4) 高纯化学试剂纯度要求更高（金属杂质在1~0.1 ng/g），0.5 μm以上的杂质颗粒必须控制在5个/mL以下，管道输送到用户。

5) 特种电子气体纯度在6 N以上，金属杂质在pg/g级，并且开发替代有毒气体的新品种电子气体。

6) 塑封料表面贴装用环氧模塑料。引线框架材料有Cu－Fe－P（KFC，CI9400），Cu－NiSi，Cu－Cr－Zr。

6.3.2 光电子材料

1) 激光晶体材料向着大尺寸、高功率应用、LD泵浦、宽带可调谐以及新波长、多功能应用方向发展。

2) 红外探测器材料大面积高均匀性HgCdTe外延薄膜及大尺寸ZnCdTe衬底材料在2010年前仍是红外探测器所用的主要材料。

3) 液晶材料研究发展超扭曲向列型（STN）和薄膜晶体管型（TFT）显示器所用混合液晶，提高性能，降低成本。

4) 高亮度发光二极管材料继规模生产发红、橙、黄光的GaAs基、GaP基外延材料之后，拓宽发光波段，开发蓝光的GaN、ZnSe基外延材料成为研究热点。

5) 光纤材料总体发展趋势向着不断扩展通信容量，降低损耗，增加传输距离，降低色散、提高带宽、抑制非线性效应，实现密集波分复用以及高灵敏度传感方向发展。光纤预制棒的生产制造上单一工艺（MOCVD，PCVD，OVD和VAD）向着混合工艺方向发展，预制棒尺寸（单棒拉丝长度）不断增大。

6.3.3 新型电子元器件用材料

1) 磁性材料　从总体上说，永磁材料正在向着高磁能积、高矫顽力、高剩磁方向发展，如国外生产NdFeB永磁合金最大磁能积已达416 kT·A/m。软磁材料正在向着高饱和磁通密度、高磁导率、低磁损耗、低矫顽力、高截止频率方向发展。国外正在开发的纳米微晶软磁合金高斯磁导率达100 000，饱和磁感应强度达1.3 T。磁记录总的发展趋势是向小型化、高密度、低噪声方向发展，因而对磁粉的要求是颗粒尺寸由微米向亚微米、纳米方向发展，颗粒分布均匀。

2) 电子陶瓷材料　小型化、片式化是电子元件发展的主要方向；多功能化、复合化是电子元件增值的主要体现；元件制造商将面向计算机、通信设备和数字化装置调整产品结构。世界各著名的大公司加大对电子陶瓷材料的开发力度。例如，日本TDK和京都陶瓷公司的研究开发费分别占销售额的5%和3%，达到2.93亿美元和2.3亿美元；美国AMP公司开发费为5.79亿美元，占销售额的10.6%。这些公司全面推进片式化、小型化元件用材料的规模生产。

3) 绿色电池材料　绿色电池材料正向着无汞、无污染、高比能量、长寿命、小型化和轻型化绿色电池方向发展，主要是开发镍氢电池和锂离子电池所用的正、负极材料——MH合金，$Ni(OH)_2$，$LiCoO_2$，$LiMnO_4$，MCMB等。

4) 信息传感材料　除光纤材料外，半导体材料、功能陶瓷材料和高分子传感材料都有一定基础，目前正围绕汽

车、医疗、信息家电、通信广播、办公和工业生产自动化等提供信息传感材料及其传感器。

7 航空航天材料

7.1 历程与成就

1）航空工业发展历程与成就 我国的航空工业是在全国解放后于20世纪50年代开始建立的，它是和新中国一起成长、壮大起来的。中国的航空工业建立以来，从修理飞机到制造飞机，从仿制前苏联航空产品到自行设计研制产品，40余年来共生产交付了各种类型飞机14 000余架，各种类型发动机50 000余台；生产定型的机载设备5 000余种。

在民用飞机方面，支线客机Y-7-200A和Y-12均已获得适航证；干线飞机通过国际合作转包生产零部件奠定了较为坚实的可持续发展的技术基础，其中MD－90干线飞机国产化率达到70%。目前已正式立项研制70座以上的新型涡扇支线客机ARJ21系列以及50座以下的涡扇支线客机ERJ145系列。航空发动机方面也已从仿制走向了自行设计研制，目前已能自行设计和研制包括涡扇、涡喷、涡桨和涡轴在内的各型发动机。

2）航天工业发展历程与成就 1956年中国开始筹建航天工业，并命名为国防部第五研究院，以后陆续演变为第七机械工业部、航天工业部、航空航天工业部、中国航天工业总公司及中国航天科技集团公司和中国航天科工集团公司。我国航天工业坚持走自力更生、独立自主的发展道路。40余年来（统计到2000年1月为止），我国已研制出11种长征系列火箭并成功地进行了60次发射，包括近地轨道（LEO）、地球同步转移轨道（GTO）和太阳同步轨道（SSO）。从1970年4月24日我国成功地发射了第一颗人造卫星——东方红一号算起，到2000年1月为止，我国已把56颗卫星送上蓝天，其中外星23颗，内含2颗铱星模拟星。这些卫星中包括通信卫星、气象卫星、科学实验卫星及国际商业卫星等，其中返回式卫星共发射了17颗，成功回收了16颗，回收成功率达94%，目前世界上能掌握卫星回收技术的只有俄罗斯、美国和中国三个国家。1999年11月20日我国成功地发射了“神舟”号飞船并成功地回收了返回舱，为载人航天发射奠定了基础。2003年11月15日中国成功地发射了第一艘载人航天飞船“神舟5号”，成为世界上第三个能独立掌握载人航天技术的国家。

新中国的航空航天工业在原来极其薄弱的基础上，发展成今天门类比较齐全、专业基本配套、具有相当规模和技术基础的高科技工业体系。中国航空航天工业在为国民经济建设、国防建设和推动科学技术发展等方面所取得的丰硕成果令世人瞩目，已经成为世界航空航天产业家族中的重要一员。

3）航空航天材料的重要作用 航空航天材料包括金属材料、无机非金属材料、高分子材料和先进复合材料四大类，按使用功能又可分为结构材料和功能材料两大类。航空航天材料既是研制生产航空航天产品的物质保障，又是推动航空航天产品更新换代的技术基础。

航空航天产品受使用条件和环境的制约，对材料提出了严格的要求。对结构材料而言，其中最关键的要求是轻质高强和高温耐蚀。从这一点上可以说，航空航天产业把结构材料的能力提高到了它的极限水平。飞行器的设计准则已经从原始的静强度设计发展到今天的损伤容限设计，设计选材时的重要决定因素是寿命期成本、强度重量比、疲劳寿命、断裂韧性、生存力及可靠性等。对于航天产品来说，还要考虑材料更高的比刚度和比强度、低的热膨胀系数，耐超高温和超低温能力，以及在空间环境中的耐久性。如为了提高航天器燃料燃烧时的效率和保证重返大气层时的防护，需要有耐高温防热材料。为了保存低温燃料如液氢、液氧，需要有耐低温材料和绝热材料。空间用高效能源也提出了许多关键性的材料问题。环境问题还包括外层空间的高真空状态、宇宙射线辐照和低地球轨道上原子氧的影响等问题。航空航天飞行器在超高温、超低温、高真空、高应力、强腐蚀等极端条件下工作，除了依靠优化的结构设计之外，还有赖于材料所具有的优异特性和功能。除结构材料外，功能材料在航空航天产品的发展中同样具有重要的作用，如微电子和光电子材料、传感器敏感元件材料、功能陶瓷材料、光纤材料、信息显示与存储材料、隐身材料，以及智能材料等。由此可见，航空航天材料在航空航天产品发展中的极其重要的地位和作用。

7.2 航空材料

1）飞机机体材料 20世纪90年代国际上最先进的第四代战斗机以美国的F22为代表，最先进的民用飞机以波音公司的B777为代表。机体结构用材料的主要特点是大量采用高比强度和高比模量的轻质、高强、高模材料，从而提高飞机的结构效率，降低飞机结构重量系数。具体来说，就是树脂基复合材料和钛合金用量的增加，传统铝合金和钢材的用量相应减少。以F22战斗机为例，树脂基复合材料的用量已达到整机结构重量的24%，而钛合金的用量则达到整机结构重量的41%。与此同时，铝合金的用量下降为占整枫结构重量的15%，而且主要是高纯、高强、高韧先进铝合金，钢的用量则下降为只占整机结构重量的5%。先进民用飞机以B777为例，树脂基复合材料的用量已占整机结构重量的11%，而钛合金的用量则占整机结构重量的7%。与此同时，铝合金的用量降为占整机结构重量的70%，而且大量采用了高纯、高强、高韧先进铝合金，钢的用量下降为只占整机结构重量的11%。国外先进军用飞机机体结构用材料的对比见表1.3-11。

表1.3-11 国外军用飞机结构材料用量对比

（质量分数） %

飞机型号	设计年代	钛合金	复合材料	铝合金	结构钢
F14	1969	24	1	39	17
幻影2000	1969	23	12	—	—
Cy27	1969	18	1	60	8
F15	1972	27	2	36	6
F18	1978	13	12	49	17
AV8B	1982	9	26	44	8
F117	1983	25	10	20	5
B2	1988	26	50	19	6
A12	1989	20	—	20	15
YF22	1989	24	23	35	5
F22	1989	41	24	11	5

目前，我国机体结构材料的研制、应用与国外的主要差距是：钛合金和树脂基复合材料等先进结构材料应用基础薄弱、材料用量偏低。机体结构最基本的用材铝合金，因品种规格或性能质量不稳定等方面的问题而不得不部分从国外进口。先进树脂基复合材料原材料基础薄弱。碳纤维等重要原材料尚不能完全立足国内。

21世纪头10年，我国军用飞机要完成先进战斗机的研

和服役。民用飞机要继续扩大国际合作和“三来加工”，逐步进行支线客机和大型飞机的研制。在机体结构材料方面，要重点抓好树脂基复合材料、钛合金、先进铝合金、超高强度钢和隐身材料等的研究和应用，同时要竭力提高许多基础航空材料的质量控制和管理水平。军用飞机和民用飞机对材料的需求目标和发展重点分别见表 1.3-12 和表 1.3-13。

表 1.3-12 先进军用飞机材料的发展需求目标和重点

特 点	对材料要求	重点发展的材料
1）具有超音速机动和超音速巡航能力 2）具备一定的隐身能力 3）超视距攻击能力和夜战能力 4）高可靠性、可维修性及高耐久性	1）大量采用轻质、高比强、高比模材料 2）某些部位的材料需要具有对电磁波和红外隐身特性 3）大量采用损伤容限型材料 4）材料环境适应性高 5）高压液压系统要求高	1）树脂基复合材料 2）铝锂合金 3）各类钛合金 4）隐身材料 5）新型超高强度钢 6）高性能透明材料 7）新型功能材料

表 1.3-13 先进民用飞机材料发展的需求目标和重点

特 点	对材料要求	重点发展的材料与技术
1）安全性、可靠性要求高 2）采用损伤容限设计 3）耐久性要求高 4）舒适性 5）经济性和竞争力	1）重要材料必须经过适航认证 2）要求材料具有轻质/高强/高温/耐蚀等特性 3）高可靠性要求材料性能实验数据要按一定数量来满足民机设计数据的统计要求 4）耐腐蚀性和环境适应性高 5）消声阻尼减振的要求高 6）舱内材料要求阻燃、无毒、低烟雾密度、高热释放速率 7）低成本材料与技术	1）适航条件与技术 2）损伤容限型铝合金 3）损伤容限型钛合金 4）超高强度钢 5）内装饰材料 6）减振阻尼材料（金属、非金属） 7）材料损伤容限数据测试及方法研究 8）低成本材料与技术 9）先进复合材料

2）航空发动机材料 我国航空发动机材料是依据不同时期发动机的仿制、改进改型和自行研制的需求而研制、生产、应用发展起来的，已初步建立中国航空发动机材料体系。目前正按照需求开展发动机用新材料的预研，型号发动机用材料的研制，定型发动机用材料的生产和改进，在计划指导下进行，基本可适应国内发动机在研型号的需要。

研制与应用现状 我国研制的发动机有活塞式、涡轴喷气式、涡轮喷气式、涡轮风扇式、涡轮螺桨式等大、中、小型发动机共数十种。航空发动机的性能水平很大程度上依赖于高温材料的性能水平，如发动机推重比的提高有赖于涡轮前进口温度的提高，而涡轮前进口温度的提高又有赖于涡轮转子部件设计结构的改进和材料的更新。我国在研发动机用材中，高温合金用量约占 60%，钛合金占 25%，其他材料占 15%。

高温合金主要用于涡轮叶片、盘、环和机匣、燃烧室、加力燃烧室及尾喷口等零件。我国已研制变形合金超过 50 种牌号，其中有燃烧室用的 GH1140，GH1015 和 GH99 等；涡轮盘用 GH2036，GH2132，GH2135，GH4133，GH698，GH4169 和 FGH95 等；环形件用 GH903，GH907 和 GH909 等。已研制铸造高温合金超过 35 种牌号和精铸用型芯、型壳和涂料等工艺材料；并研制成功 DZ4 和 DZ22 定向柱晶和 DD3 单晶等空心、无余量精铸导向叶片和涡轮叶片等。值得指出的是，金属间化合物 Ni_3Al 在 1 100℃、100 h 持久强度居国际先进水平。

钛合金主要用于风扇、压气机盘、鼓筒、叶片、环和机匣等，高温钛合金用于高压压气机部分。我国已研制的钛合金超过 37 种牌号，其中变形合金有 TC4，TC11，Ti－1023，Ti－153 和 Ti－55 等；铸造钛合金有 ZT－3，ZT－4，ZT－5 和 XTA－7 等；已研制成功 600℃（Ti－60）和 700℃（TD－2）Ti_3Al 合金。现正探索 TiAl 合金。

根据先进发动机发展的要求，据有关方面预测，在实现航空发动机高推重比的道路上，50%以上的技术要依靠新材料和与之相应的新结构、新工艺来支撑。由于燃烧室温度大幅度提高，出口温度可达 2 200～2 500 K，且由于设计与结构上的原因，可用于冷却的空气很少，据称，美国综合高性能发动机技术计划 IHPTET 的实现，70%～80%要靠材料的改进。而且，21 世纪先进航空发动机性能对材料的要求将不满足于传统材料的渐进式提高，要求开辟新的材料系统及工艺领域。英国罗·罗公司对 21 世纪发动机材料可能发生的变化进行了预估，按此预测，传统的铝合金及结构钢，在发动机中的用量会进一步减少，高温合金、钛合金等特种金属材料到 21 世纪初叶也会有所降低，代之而来的将是陶瓷基和金属基复合材料。美国一些发动机公司对下一代发动机用材提出的设想，也认为未来先进发动机中占主导地位的将是各种耐高温基体的复合材料。表 1.3-14 是 IHPTET 提出的推重比15～20 发动机主要部件的用材设想。

表 1.3-14 推重比 15～20 发动机主要部件用材料

部 件	主 要 特 性	材 料
风扇	后掠空心风扇叶片，3 级变 1 级，减重 50%	钛合金＋聚合物基复合材料
压气机	鼓筒式叶环转子，减重 70%	704～982℃钛基复合材料
燃烧室	变几何结构，减小出口温度分布系数	陶瓷基复合材料

续表 1.3-14

部　件	主 要 特 性	材　料
涡轮	整体叶盘结构，减重30%，2 270～2 470 K 超气冷涡轮叶片。F119温度为1 997 K	陶瓷基复合材料，减重80%
加力燃烧室	单位推力比F100高70%～80%	1 204℃陶瓷火焰稳定器－喷嘴环 1 538℃陶瓷加力燃烧室－喷嘴
尾喷管	全方位矢量喷管	982℃TiAl复合材料 ＞1 538℃陶瓷、C－C复合材料
飞机特点	$H=21\ 000$ m，$M_a=3\sim4$，作战半径1 850 km，隐身，载弹1 t	

3）机载设备材料　机载设备是保证飞机正常工作及完成各项飞行和作战任务的机上各系统及设备的总称，它包括飞行保障设备、辅助动力装置设备、电子设备和武器设备四大类。机载设备目前正朝着综合化、小型化、模块化、智能化的方向发展，成本费用也不断增加。一架先进军用飞机的机载设备费用已占到整架飞机费用的30%～40%。第四代战斗机上机载设备的特点是航空电子发展迅速，已由联合式发展到综合式，所有电子设备均是统一设计、综合集成，而机电系统则朝着公共管理系统的方向发展，即采用公用计算机来管理控制电气、液压、燃油、环控、生命保障、辅助动力、刹车、起落架、综合显示、监视等系统。机载设备中的关键材料主要是各种微电子、光电子、传感器等光、声、电、磁、热的高功能及多功能材料。机载设备材料的发展目标要求及重点见表1.3-15。

表 1.3-15　机载设备对材料发展的需求

特　点	对材料要求	重点发展的材料
1）超视距攻击能力 2）近距格斗能力 3）精确性高 4）灵敏反应 5）抗干扰能力强	1）缺陷密度极低 2）针对不同用途对其物理性能（光、声、电、磁、热）要求高 3）加工、成形、联结、涂覆等技术不能对材料物理性能和装备功能产生有害影响	1）高灵敏度红外探测材料 2）高透过率红外头罩材料 3）电致磁致伸缩陶瓷材料 4）双脉冲点火发动机舱隔板材料 5）激光倍频材料 6）高强度激光材料 7）双模制导头罩材料 8）零膨胀微晶玻璃材料 9）极高反射率镀膜材料及技术

7.3　航天材料

（1）导弹、运载火箭（含火箭发动机）材料

表1.3-16列出了我国几种液体地地导弹弹体和运载火箭箭体用主要结构材料。液体导弹弹体和运载火箭箭体所用结构材料主要为铝合金，此外还有少量的钛合金和结构钢等材料。在此领域国内外新材料研发的重点如下。

1）运载火箭　根据发展新一代大型运载火箭的发展要求，对新材料研制的要求见表1.3-17。其主要要求是：开展新型高强轻质箭体结构材料的研制。开展新型高强轻质液体推进剂贮箱结构材料的研制。开展大推力液氢－液氧火箭发动机关键结构材料的研制。开展大推力液氧－煤油火箭发动机关键结构材料的研制。

表 1.3-16　我国几种液体地地导弹弹体和运载火箭箭体用主要结构材料情况

名称	推进剂贮箱	尾段、箱间段、级间段	增压气瓶
近程	5A03O（LF3M）	2A12（LY12）	25CrMnSi
中近程	5A03O（LF3M） 5A06O（LF6M）		
中程	5A06O（LF6M） 5A06H×4（$LF6Y_2$）		TC4
中远程	5A06O（LF6M） 5A06H×4（$LF6Y_2$）	2A12（LY12） 7A09（LC9）	
洲际	2A14（LD10）	2A12（LY12） 7A04（LC4） 7A09（LC9） 30CrMnSi	

2）战略导弹及弹头　为了提高整个武器系统的生存能力、突防能力和综合作战能力，必须解决第二代固体洲际导弹小型化、轻质化、高性能和具有全天候作战能力等问题。为此，对今后新材料的研究，提出了多方面的需求。总的要求可大致归纳如下：实现弹头结构小型化、轻质化，减轻弹头结构质量。实现弹体结构轻质化，减轻弹体结构质量。提高导弹第三子级固体火箭发动机质量比。实现弹上仪器设备小型化，减轻仪器设备结构质量。对导弹的地面设备（如发射筒、发射车等）所用材料都提出了相应的要求。重点研发的材料见表1.3-18。

（2）航天器材料

从1970年我国发射第一颗人造卫星以来，已发射了包括返回式卫星、资源卫星、气象卫星、通讯卫星在内的各种卫星50余颗。

航天器长期在高真空状态下工作，对材料有特殊要求，除高比模量和高比强度外，还要求耐空间环境，耐电子、质子辐照，耐氧原子，耐冷热交变。

返回式卫星和通讯卫星所用材料有所不同，返回式卫星属低轨道卫星（230～400 km），有两个舱段，即仪器舱和返回舱，返回舱要求密封，内部承力结构主要采用轻金属结构材料，外部为防热结构。通讯卫星属高轨道卫星（36 000 km），80%以上采用复合材料。

通讯卫星主要包括天线结构、太阳电池阵和卫星本体结构。表1.3-19列出了我国发射的通讯卫星上用的主要材料。

表 1.3-17 新一代运载火箭对新材料的需求

应用部位	材 料	技 术 要 求
箭体结构	1）高强轻质铝合金 2）高性能碳－环氧复合材料	比常规铝合金减轻结构质量
推进剂贮箱	1）高强可焊 Al－Li 合金（2195 和 1460 合金） 2）高性能碳－环氧复合材料	使液氢和液氧贮箱比用常规铝合金减轻结构质量
液氢－液氧火箭发动机	1）电铸镍锰合金材料及电铸工艺技术 2）GH4169 合金材料及精铸工艺技术 3）新型高温合金材料 4）低温（－253℃）钛合金材料及成形技术 5）高强钛合金薄壁管材 6）Ti_3Al 及以其为基的复合材料与成形技术	抗拉强度比电铸镍提高 满足泵壳体及涡轮壳体成形要求 性能分别与 Incoloy903 和 Mar－M246 相当，满足高压系统零组件要求 满足液氢泵诱导轮成形要求 减轻发动机机架结构质量 比镍基高温合金涡轮盘减轻结构质量
液氧－煤油火箭发动机	1）新型不锈钢材料 2）新型铸造不锈钢材料及工艺技术	性能分别与 BHC－25、BHC－16 和 0Cr16Ni6 合金相当，满足导管、涡轮泵轴杆、低温紧固件等要求 性能分别与 BHJI－1 和 BHJI－6 相当，满足涡轮泵壳、液氧泵叶轮要求

表 1.3-18 固体洲际导弹对新材料的要求

应用部位	材 料	技术要求
弹头结构	1）先进碳－碳复合材料 2）高性能布带斜缠碳/酚醛复合材料 3）新型陶瓷基复合材料 4）高强轻质 Al－Li 合金材料 5）高性能抗核爆 X 射线防护材料 6）高性能红外、雷达隐身材料 7）多功能诱饵材料	实现弹头小型化、轻质化、高性能、全天候、强突防，减轻结构质量
弹体结构	1）高性能碳－环氧复合材料 2）高性能碳－双马来酰亚胺复合材料 3）高性能碳－聚酰亚胺复合材料 4）高强轻质金属结构材料（Al－Li 合金、B－Al 金属基复合材料等）	实现弹体结构轻质化，减轻结构质量
固体火箭发动机壳体	1）新型芳纶－环氧复合材料 2）高强中模碳－环氧复合材料 3）四向碳－碳喉衬材料和工艺技术 4）碳－碳喷管材料和工艺技术	提高发动机质量比
仪器框架	1）高性能碳－环氧复合材料 2）高强轻质铝合金材料	弹上设备小型化、轻质化，减轻结构质量
地面设备	1）碳－环氧复合材料 2）高性能金属结构材料	地面设备轻质化，减轻地面设备（如发射筒等）结构质量

表 1.3-19 我国通讯卫星采用的主要结构材料

结．构	材 料
天线反射器	碳纤维复合材料
太阳电池阵结构	高模碳纤维、铝蜂窝夹心
卫星本体结构	
承力筒	高模碳纤维
蜂窝夹层板结构	铝蜂窝
桁架结构	高模纤维
其他部件	
返回舱	5A06（LF6）铝镁合金
仪器舱	2A12T4（LY12CZ）高强铝合金
相机支架	ZM5 铸镁合金
气瓶、球底、支架	TC4
表面张力箱	TB2

需重点发展的航天器材料包括超高模碳纤维复合材料、超高模高强碳纤维复合材料，结构胶粘剂，防原子氧保护层，卫星活动部件的表面超硬、耐磨、润滑膜层，铝合金和铝－锂合金，金属基复合材料，空间站的柔性太阳电池帆板，高可靠、长寿命的密封圈。

铝－锂合金密度（2.47～2.49 g/cm^3）比常用铝合金低10%，而模量却高 10%，采用铝－锂合金取代常规的高强铝合金，能使结构件质量减小 10%～20%，刚度提高 15%～20%，被认为是 21 世纪航空航天器的主要结构材料。材料研制部门应研究满足卫星需求的中强 Al－Li 合金（如俄罗斯的 1420，1421）、高强 Al－Li 合金（如美国的 2090，8091；俄罗斯的 1450，1460）和高损伤容限 Al－Li 合金（如美国的 9T8096）。应用部门需解决铝－锂合金阳极化、焊接工艺、成型工艺和复合工艺等技术问题。

7.4 航空航天关键功能材料

在现代航空航天工程上，功能材料用于制造各种各样的传感器、换能器和信息处理器。它们是飞机、火箭、导弹、卫星、航天器以至星际航行的制导、控制、环境控制、能源供给、电气系统、电子系统、仪表、通信、武器火力控制系统以及生命保障系统中不可缺少的重要材料，在航空航天上占有非常重要的地位，已成为现代航空航天工程先进性的决定因素之一。

为适应航空航天技术发展的需要，需研发的功能材料包括微电子器件材料锗－硅（GeSi）材料、微电子射频元件材料铁氧体和稀土永磁材料、光电子器件量子阱材料、光电子光学晶体材料、传感器用功能陶瓷材料、信息显示发光材料、信息显示液晶材料、智能结构传感与驱动元件材料等。

1）功能材料的发展重点　碲镉汞，金刚石厚膜材料，ZnS－ZnSe 叠层材料，大尺寸蓝宝石材料，大直径硅单晶，砷化镓、磷化铟单晶，氟化钙，增透膜材料，Nd: YAG，极小（零）膨胀材料，液晶材料，大功率激光器材料（如化学氧碘激光器材料），固体激光器材料等。

2）必须着重发展的先进功能材料　锑化铟，PtSi－Si，量子阱材料，硅－锗材料，GaP，BP，GeC，$CaLa_2S_4$，磷化硼，锗化硅增透膜，保偏光纤，稀土永磁材料，功能陶瓷材料，毫米波铁氧体材料，加固抗辐射封装材料，超导材料（高温及低温），HgTeMn，HgZnTe，InAsSbBi，InSbBi 等。

3）在急需的传统材料中尚未完全过关的材料　蓝宝石（Al_2O_3），尖晶石（$MgAl_2O_4$），压电陶瓷，电致伸缩材料，减磨、耐磨用铸造双金属板，高强韧铸造铝硅合金，碳－碳复合刹车材料，密封橡胶，阻尼减振橡胶，不燃液压油，敷铜印制板，透波材料，特种陶瓷，隔热材料等。

4）需及时启动研发的关键功能材料　这类材料是研究发展是先进航空航天器所需的关键的功能材料，也可以说这类功能材料代表着一个国家的科学技术先进性。例如，智能材料在自适应机翼形状的控制、减振、降噪、飞机故障的监测等方面的应用。用智能材料制造的主动控制自适应旋翼及相关控制系统比现在的铰接式旋翼轻 20%～30%，还可降低飞行阻力及功耗 15%，直升机的机动性可提高 15%～30%。又如微电子技术所需的能够显著提高性能并降低成本的各种新型半导体材料，如高功率工作和热控制性能优于砷化镓和硅的金刚石、SiC 和 InP 等先进材料。还有发展红外探测器装置的大密度、低价格、高性能的红外阵列技术和长波红外传感器所需的红外材料（Ⅲ～Ⅴ族化合物量子阱材料，锗化硅），以及生物传感器材料等。

8 先进陶瓷材料

先进陶瓷材料一般是指以精制高纯人工合成的无机化合物为原料，采用精密控制工艺烧结而制得的高性能陶瓷，又称新型陶瓷、精细陶瓷。先进陶瓷是相对于用天然无机物烧结制得的传统陶瓷而得名。先进陶瓷按其使用性能可以分为结构陶瓷、功能陶瓷和生物陶瓷三大类。

8.1 结构陶瓷

结构陶瓷又称工程陶瓷，具有耐高温、耐磨、耐腐蚀、耐冲刷、抗氧化等一系列性能，以发挥其机械、热、化学等性能为主而用于各种结构部件，如要求耐腐蚀、耐磨损的部件：各种机械密封件、轴承、刀具、球阀、缸套等。结构陶瓷可以承受金属材料和高分子材料难以胜任的严酷工作环境，在能源、航空航天、机械、冶金和化工等工业中有重要的应用前景。结构陶瓷大致分为氧化物陶瓷和非氧化物陶瓷两大类，典型的结构陶瓷成分类别见表 1.3-20。

表 1.3-20 典型的结构陶瓷成分类别

种类		材料
氧化物系		Al_2O_3、MgO、ZrO_2、SiO_2、UO_2、ThO_2、BeO 等
非氧化物系	碳化物	SiC、TiC、B_4C、WC、UC、ZrC 等
	氮化物	Si_3N_4、AlN、BN、TiN、ZrN 等
	硼化物	ZrB_2、WB、TiB_2、LaB_6 等
	硅化物	$MoSi_2$ 等
	氟化物	CaF_2、$BaFe_2$、MgF_2 等
	硫化物	ZnS、TiS_2、$M_xMo_6S_8$（M＝Pb、Cu、Gd）等
	碳和石墨	C

（1）氧化物陶瓷

氧化物陶瓷是发展比较早的高温结构陶瓷材料，一般是指熔点高于二氧化硅晶体熔点（1 728℃）的各种氧化物陶瓷或复合氧化物陶瓷。氧化物陶瓷是典型的离子晶体，阳离子或阴离子由较强的离子键结合。多晶材料的氧化物陶瓷，受到晶体本身晶粒大小、气孔大小和分布、晶界物质的性质和分布等微观结构的影响，常温强度较高而高温强度较小。

氧化物陶瓷在高温下化学稳定性好，特别是具有优异的抗氧化性能。在各种气氛和较高的温度下，大多数氧化物陶瓷是稳定的，具有良好的耐腐蚀性能，但在 HF、HCl 和 H_2S 气氛中显示出不稳定性。在较高温度下，氧化物陶瓷与高熔点金属和碳接触时，容易被还原成低价氧化物而挥发。多数氧化物陶瓷在高温下具有较好的电绝缘性，可以用作高温电绝缘材料。

氧化物陶瓷主要有氧化铝陶瓷（Al_2O_3），氧化锆陶瓷（ZrO_2），莫来石陶瓷（$3Al_2O_3\cdot 2SiO_2$）和氧化锆增韧氧化铝陶瓷（ZTA）等。普通氧化铝陶瓷的典型性能见表 1.3-21。几种主要的氧化锆陶瓷材料体系的组分特点见表 1.3-22。几种典型氧化锆增韧陶瓷的性能数据见表 1.3-23。氧化锆陶瓷的应用领域和相关的应用场合见表 1.3-24。

表 1.3-21 普通氧化铝陶瓷的典型性能

性能	Al_2O_3 含量（质量分数）/%			
	99～96.5	94.5～96.5	88～94.5	80～86
密度/$g\cdot cm^{-3}$	3.75～3.8	3.7～3.9	3.4～3.7	3.3～3.4
硬度/GPa	12.8～15	12～15.6	9.7～12	—
弹性模量/GPa	300～380	300	250～300	200～240
抗弯强度/MPa	230～350	310～330	250～330	200～300
线胀系数（200～1 200℃）/$10^{-6}K^{-1}$	8～8.1	7.6～8	7～7.6	—
室温热导率/$W\cdot m^{-1}\cdot K^{-1}$	24～26	20～24	15～20	—
烧成温度/℃	—	1 520～1 600	1 440～1 600	—

（2）非氧化物陶瓷

非氧化物陶瓷是由金属和非金属的碳化物、氮化物、硅化物、硼化物、氟化物、硫化物等不含氧的化合物制造的陶瓷材料的统称。非氧化物高温陶瓷主要为周期表中Ⅱ～Ⅲ主族、Ⅲ～Ⅶ副族、第Ⅷ族、镧系、锕系等元素与 B、C、N、P、S、Si、F 等元素的化合物以及这些非金属元素之间的互化物。

表 1.3-22　几类主要的氧化锆陶瓷材料体系的组分特点

中文全称	英文全称与缩写	材料体系的组分特点
部分稳定氧化锆陶瓷	partially stabilized zirconia，PSZ	通常以立方相氧化锆为主相，其中存在有四方相氧化锆析出相作为增强剂
氧化钙部分稳定氧化锆陶瓷	Ca - PSZ	CaO 为稳定剂，含量一般在 7.5% ~ 8.7%（摩尔分数）之间
氧化镁部分稳定氧化锆陶瓷	Mg - PSZ	MgO 为稳定剂，含量一般在 8.5% ~ 10%（摩尔分数）［约 2.8% ~ 3.5%（质量分数）］之间
相变增韧陶瓷	transformation - toughened ceramics，TTC	在陶瓷体系中引入特定结构的氧化锆颗粒来提高力学性能
相变增韧氧化锆陶瓷	transformation - toughened zirconia，TTZ	含 PSZ 和 TZP 相的氧化锆陶瓷
四方氧化锆多晶体	tetragonal zirconia polycrystals，TZP	由在室温下稳定存在的四方相氧化锆晶粒构成，常见的稳定剂是 Y_2O_3 和 CeO_2
氧化钇稳定的四方多晶氧化锆增韧陶瓷	Y - TZP	Y_2O_3 为稳定剂，含量一般在 2% ~ 3%（摩尔分数）［约 7% ~ 8%（质量分数）］之间
氧化锆增韧陶瓷	zirconia toughened ceramics，ZTC	在陶瓷体系中引入特定的氧化锆来提高其力学性能
氧化锆增韧氧化铝陶瓷	zirconia toughened alumina，ZTA	在氧化铝基体中引入特定的氧化锆来提高其力学性能

表 1.3-23　几种典型氧化锆增韧陶瓷的性能数据

性　能	Y - TZP	Ce - TZP	ZTA	Mg - PSZ	3Y - TZP ($20\% Al_2O_3$)
密度/$g \cdot cm^{-3}$	6.05	6.15	4.15	5.75	5.51
硬度（HV_{30}）	1 350	900	1 600	1 020	1 470
抗弯强度/MPa	800 ~ 1 000	400 ~ 700	500 ~ 600	400 ~ 800	2 400
弹性模量/GPa	205	215	380	205	260
泊松比	0.3			0.23	
断裂韧度 /$MPa \cdot m^{1/2}$	9.5	15 ~ 20	4 ~ 5	8 ~ 15	6 ~ 8
线胀系数 /$10^{-6} \cdot K^{-1}$	10	8	8	10	9.4
室温热导率 /$W \cdot m^{-1} \cdot K^{-1}$	2 ~ 3.3	2	23	1 ~ 2	3

表 1.3-24　氧化锆陶瓷的应用领域和相关的应用场合

应用领域	应用场合
机械、纺织工业	特种陶瓷切削刀具、标准量具、机械密封件、热机部件（气门座、气缸套）、冲压模具、各类耐磨件
冶金、钢铁和石化工业	连铸浇口和喷嘴、陶瓷阀和口环、有色金属拉丝模具、高性能耐火材料、耐火纤维和保温材料、高温发热体、热障涂层
生物医用和环保产业	各类人工植入体（如人工关节等）、多孔催化剂载体、三元催化剂
电信产业	光纤连接器用氧化锆插芯和套筒
电子功能材料	固体电解质、氧敏感探头、高温电极材料、各类高温承烧板
日常生活	体育器械、日常切割刀具、耐磨部件（如表带等）

1）碳化物是一种最耐高温材料，以通式 Me_xC_y 表示，可以分为金属碳化物和非金属碳化物两种。碳化物高温结构陶瓷通常是指 SiC、TiC、B_4C、WC、Cr_3C_2、ZrC 等以及它们的复合材料。碳化物陶瓷的主要特性是具有高熔点，如 TiC 的熔点高达 3 460℃，ZrC 的熔点则达 3 540℃。另外，许多碳化物陶瓷均具有很高的硬度，B_4C 是仅次于金刚石和立方氮化硼的超硬材料。碳化物陶瓷还具有良好的导电性和导热性，化学稳定性也比较优良。大多数碳化物陶瓷在常温下不与酸反应，最稳定的碳化物陶瓷甚至不受硝酸和氢氟酸混合酸的腐蚀。

碳化硅陶瓷的物理性质见表 1.3-25，其力学性能见表 1.3-26，其主要应用领域见表 1.3-27。

表 1.3-25　碳化硅陶瓷的物理性质

相对分子质量	40.07
颜色	纯度 < 99.0%，黑色；纯度 < 99.5%，深绿色；纯度 > 99.7%，浅绿色
密度/$g \cdot cm^{-3}$	3.21
分解温度/℃	2 300
熔点/℃	2 700（汽化）
比热容/$J \cdot kg^{-1} \cdot K^{-1}$	640
热导率/$W \cdot m^{-1} \cdot K^{-1}$	350（单晶）
线胀系数/K^{-1}	2.2×10^{-6}（室温） 4.0×10^{-6}（1 000℃）
介电常数	9.6 ~ 10.0（静态） 6.5 ~ 6.7（高频）
弹性模量/GPa	470
泊松比	0.21（CVD SiC）
抗氧化性能	800℃以上，有轻微缓慢的氧化

2）氮化物陶瓷也是一类高硬高强陶瓷材料，典型的如氮化硅陶瓷具有很高的高温强度，优良的抗热振性能，较高的硬度和耐磨性，抗氧化性强；立方氮化硼的硬度与金刚石相当，部分性能如热稳定性优于人造金刚石；氮化钛硬度高、

表 1.3-26 碳化硅陶瓷的力学性能

性能	制备工艺			
	CVD	反应烧结	常压烧结	热压烧结
理论密度/g·cm^{-3}	3.20	2.9~3.1	3.2	3.15~3.20
烧结密度/%	>99.99	>80	>95	>98.0
室温抗弯强度/MPa	375	150~450	400~600（固相烧结） 750~900（液相烧结）	600~1 000
断裂韧度/MPa·m$^{1/2}$	3.1~3.5	2.5~4.5	3.5~4.5（固相烧结） 8.0~10.0（液相烧结）	4.1~5.2
Weibull 模数	12	12	15	12~18

性能		制备工艺			
		CVD	反应烧结	常压烧结	热压烧结
弹性模量/GPa		440	300~393	410~430	420~450
泊松比		0.17（0.14~0.21）			
硬度	HV/kgf·mm^{-2}	2 550~2 850		2 600	2 500~2 700
	HRA	93~95	90~93	92~95	
线胀系数（室温~1 000℃）/10^{-6}·K^{-1}		2.2(室温) 4.0(1 000℃)	4.5~5.0	4.2~4.5	4.5
热导率/W·m^{-1}·K^{-1}		200~300	70~125	90~125	110~180
热震因子		157	—	180	164

表 1.3-27 先进碳化硅陶瓷材料的主要应用领域

应用领域	应用环境	用途	性能特点
石油工业	高温、高压（液）研磨性物质	喷嘴，轴承，密封，阀片	耐磨，导热
微电子工业	大功率散热	封装材料，基片	高热导，高绝缘
化学工业	强酸（HNO_3，H_2SO_4，HCl）、强碱（NaOH）、高温氧化	密封、轴承、轴套、泵部件、热交换器、气化管道、热电偶保护管	耐磨损，耐腐蚀，气密性，高温耐腐蚀性
机械	研磨、滑动、旋转	内衬，泵零件，喷砂嘴，轴承，阀	耐磨损，耐腐蚀，高强度，高硬度，低摩擦
汽车、发电	燃烧（发动机、发电机）	燃烧室部件、涡轮增压器	高温强度好，摩擦性好，耐高温腐蚀
航空航天	燃烧，高温、大流量气流冲刷	涡轮叶片，燃气轮机动、静叶片，火箭喷嘴	高温性能好，低惯性负荷，耐热冲击
空间光学	振动，空间环境	空间光学系统（CCD 相机），反射镜	低密度，高比刚度，高热导
激光	高温、大功率	反射镜	高比刚度，高热导，高温稳定
喷砂器	高温研削	喷嘴	耐磨损
造纸业	纸浆、纸浆废液（50%NaOH）	密封，套管，轴承衬垫	耐腐蚀，耐磨损，低摩擦
钢铁工业	高温气体、金属液体	热电偶保护管，辐射管，热交换器，燃烧嘴	耐高温，耐腐蚀，气密性
矿业	研削	管道内衬，泵部件	耐磨损
原子能	含硼高温水	密封，轴套	耐放射性
其他	高温金属加工	拉丝、成型模具	耐磨损，耐腐蚀

化学稳定性好，还具有较高的导电性和超导性。氮化铝陶瓷导热性好，电绝缘性能高，是一种优良的低介绝缘基板材料。将某些金属氧化物或者氮化物固溶到氮化硅晶格，可以形成赛隆（Sialon）陶瓷，这种新型的结构陶瓷具有较低的热膨胀系数、较高的耐腐蚀性及优良的抗氧化性、很高的常温和高温强度、很强的耐磨性、良好的热稳定性和不高的密度。

氮化硅陶瓷的典型性能见表 1.3-28。氮化铝及其他基板材料的力学性能与热性能见表 1.3-29。不同类型的热压氮化硼陶瓷性能见表 1.3-30。

表 1.3-28 Si_3N_4 陶瓷的典型性能

性能	烧结方法					
	反应烧结	无压烧结	反应烧结重烧结	热压烧结	气氛压力烧结	热等静压烧结
密度/g·cm^{-3}	2.6~2.8	3.2~3.4	3.2~3.4	3.2~3.3	3.2~3.3	3.2~3.3
烧结密度/%	80~85	95~98	95~97	99.5~99.9	99~99.5	99.5~99.9
抗弯强度 σ_{RT}/MPa	200~300	850~950	700~850	1 000~1 300	900~1 000	950~1 100
硬度（HRA）	83~85	90~91	90~91	92~93	91~92	92~93
断裂韧度 K_{IC}/MPa·m$^{1/2}$	3~4	5~6	5~6	6~7	6~7	6~7
Weibull 模数	15~20	10~18	10~18	15~20	12~18	15~20
弹性模量 E/GPa	160~200	280~300	280~300	300~320	290~320	300~320
线胀系数 α/10^{-6}·K^{-1}	2.5~3.0	2.8~3.2	2.8~3.2	3.0~3.5	2.8~3.2	2.8~3.4
热导率 λ/W·m^{-1}·K^{-1}	17	20~25	20~25	25~30	23~28	25~30
扩热冲击温差 ΔT_c/℃	300~400	700~900	700~900	800~1 000	700~900	800~1 000

表 1.3-29 AlN及其他基板材料的力学性能与热性能

性能	AlN	BeO	Al_2O_3	Si
室温热导率/$W\cdot m^{-1}\cdot K^{-1}$	270(320)	260(370)	20	120
线胀系数/$10^{6}\cdot K^{-1}$	4.4	8	7.2	3.2
体积电阻率/$\Omega\cdot cm$	10^{14}	10^{14}	10^{14}	$10^{-3}\sim10^{3}$
介电常数（1 MHz）	8.9	7	9.4	12
介质损耗(1 MHz)/10^{-4}	8	3	4	N/A
介电强度/$kV\cdot mm^{-1}$	15	10	15	10
密度/$kg\cdot m^{-3}$	3 250	2 900	3 890	
Vickers 硬度/GPa	12	12	23～27	
抗弯强度/MPa	340～490	245	304～314	
透光度（λ = 6 μm，t = 0.5 mm）/%	48	不透光	不透光	

表 1.3-30 不同类型的热压BN陶瓷的性能

性能指标	高热导型 BN	高强型 BN	高纯型 BN
BN含量/%	95	97	98.5
密度/$g\cdot cm^{-3}$	2.1	2.1	1.9～2.0
硬度（莫氏）	2	2	2
抗弯强度/MPa	63	160	40.3
抗压强度/MPa	101	250	61
线胀系数	2.7（//）	7.93（//）	2.59（//）
(25～700℃)/$10^{-6}\cdot K^{-1}$	0.3（⊥）	0.59（⊥）	0.71（⊥）
热导率/$W\cdot m^{-1}\cdot K^{-1}$	56.94（⊥）	28.89（⊥）	35.17（⊥）
电阻率/$\Omega\cdot cm$	3.5×10^{13}	4×10^{13}	1.6×10^{14}

注：（//）表示平行于热压方向；（⊥）表示垂直于热压方向。

3）硼化物的结构比碳化物、氮化物复杂得多，硼原子之间以共价键构成链状、网状和架状坚固结构，它与金属的结合以离子键为主，具有高熔点、高硬度、难挥发等性能，导电、导热性好，但高温抗蚀性、抗氧化性较差。硅化物中的硅原子之间同样具有牢固的共价键结构，金属原子处于硅原子的结构单元之间，或者硅原子的结构单元同金属原子的结构单元分开。硅化物一般熔点、硬度都较低，力学性能较差，在高温下容易产生塑性变形。

结构陶瓷的弱点在于脆性、均匀性差，可靠性低。近一二十年来围绕这些关键性问题的研究取得了很大进展。在对结构陶瓷的显微结构、缺陷形成和微裂纹扩展机理进行深入研究的基础上，人们建立了相变增韧、弥散强化、纤维增韧、复相增韧、表面强化、原位生长强化增韧等多种有效的强化、增韧技术；用化学法制备高纯、超细、均质粉体技术的出现；注射成型、高温等静压、微波烧结、自蔓延烧结等新工艺、新技术的发展以及有关相平衡、反应动力学、胶体化学、表面科学、烧结机理等基础研究的新成就，推动了结构陶瓷的合成与制备技术的新发展，结构陶瓷的强度和韧性有了大幅度的提高，脆性得到改善，某些结构陶瓷的韧性已经接近铸铁水平。这些进步将使结构陶瓷得到更广泛的应用。

8.2 功能陶瓷

功能陶瓷（functional ceramics）是指检测、转换、耦合、传输及存储电、磁、声、光、力、热、化学和生物等信息的介质材料。由于这类材料的组成可控和性能多样，应用十分广泛。功能陶瓷主要包括铁电、压电、介电、热释电、半导、导电、超导和磁性等陶瓷，是电子信息、集成电路、计算机、通信广播、自动控制、航空航天、海洋超声、激光技术、精密仪器、机械工业、汽车、能源、核技术和医学生物等近代高新技术领域的关键材料。功能陶瓷是一个涉及材料科学、物理学、化学、电子学等多个学科的交叉领域。随着现代电子信息技术、新能源以及军用技术的发展，功能陶瓷的战略地位日益得到重视，世界各国均投入大量人力、物力、财力研究开发各类功能陶瓷材料。

功能陶瓷种类繁多，包括作为表面组装技术（SMT）重要基础的高性能片式元件及其材料、压电驱动器与超声微马达、复合与复相功能陶瓷器件、软化学与功能陶瓷薄膜、半导体陶瓷与传感器、微波介质陶瓷、电子封装用陶瓷基片材料等。功能陶瓷的主要分类见表1.3-31。

表 1.3-31 功能陶瓷的分类

类别		成分	应用
电功能陶瓷	绝缘陶瓷	Al_2O_3、BeO、MgO、AlN、Si_3N_4	集成电路基板、封装陶瓷、高频绝缘瓷
	介电陶瓷	TiO_2、$La_2Ti_2O_7$、$Ba_2Ti_9O_{20}$	陶瓷电容器、微波陶瓷
	铁电陶瓷	$BaTiO_3$、$SrTiO_3$	陶瓷电容器
	压电陶瓷	PZT、PT、LNM、(PbBa) $NaNb_5O_{15}$	超声换能器、谐振器、滤波器、压电点火器、压电电动器、微位移器等
	半导体陶瓷	NTC（Mn、Co、Ni、Fe、$LaCrO_3$、$ZrO_2-Y_2O_3$、SiC）	温度传感器、温度补偿等
		PTC（Ba－Sr－Pb）TiO_3	温度补偿和自控加热元件等
		CTR（V_2O_5）	热传感元件、火灾传感器等
		压敏电阻 ZnO	浪涌电流吸收器、噪声消除器、避雷器等
		SiC 发热体	电炉、小型电热器等
		半导性 $BaTiO_3$、$SrTiO_3$	晶界层电容器
	快离子导体陶瓷	$B-Al_2O_3$、ZrO_2	钠硫电池固体电解质、氧传感器陶瓷等

续表 1.3-31

类别		成分	应用
磁功能陶瓷	软磁铁氧体	Mn－Zn、Cu－Zn、Cu－Zn－Mg、Ni－Zn	记录磁头、温度传感器、电视机、收录机、通讯机、磁芯、电波吸收体
	硬磁铁氧体	Ba、Sr 铁氧化物	铁氧体磁石
	记忆用铁氧体	Li、Mn、Ni、Mg、Zn 与铁形成的尖晶石型铁氧体	计算机磁芯等
光功能陶瓷		透明 Al_2O_3 陶瓷	高压钠灯
		透明 MgO 陶瓷	照明或特殊灯管、红外输出窗材料
		透明 $Y_2O_3-Th_2O_3$ 陶瓷	激光元件
		PLZT 透明铁电陶瓷	光存储元件、视频显示和存储系统等
生物及化学功能陶瓷	湿敏陶瓷	$MgCr_2O_4-TiO_2$、$TiO_2-V_2O_5$、Fe_2O_3、$ZnO-Cr_2O_3$（$LiZnVO_4$）	工业湿度检测、烹饪控制元件等
	气敏陶瓷	SnO_2、$\alpha-Fe_2O_3$、TiO_2、ZrO_2、$CoO-MgO$、ZnO、WO_3	汽车传感器、锅炉燃烧控制、气体泄漏报警、各类气体探测等
	载体用陶瓷	堇青石瓷、Al_2O_3 瓷、$SiO_2-Al_2O_3$ 瓷等	汽车尾气催化载体、化学工业用催化载体、酵素固定载体等
	催化用陶瓷	沸石、过渡金属氧化物	接触分解反应催化、排气净化催化等
	生物陶瓷	Al_2O_3、氢氧磷灰石	人造牙齿、关节骨等

8.3 生物陶瓷

生物陶瓷指用作生物医学材料的陶瓷材料。可以是单晶的，也可以是多晶的；可以是致密的，也可以是多孔的；可以是单相的，也可以是多相的。主要用于人体骨骼－肌肉系统的修复和替换，也用于心血管系统的修复以及药物运达和缓释载体。

1892年硫酸钙材料已用于充填骨缺损，但是生物陶瓷的发展主要是在20世纪60年代以后。1963年和1964年，多晶氧化铝陶瓷分别临床用于骨矫形和牙种植。1969年热解碳涂层临床用于人工心瓣膜成功，此后各种类型的碳素涂层发展成为心血管系统修复的重要材料。1969年美国人 L. 亨奇（Hench）发明了生物玻璃，并发现它植入生物体内后可以在界面上与硬、软组织形成键接。1971年羟基磷灰石陶瓷应用成功，从此开始了生物活性陶瓷发展的新纪元。1971年又开始对可吸收生物陶瓷进行广泛的临床应用研究。生物陶瓷复合材料的发展主要在20世纪70年代以后，大量工作集中于80年代。目前生物陶瓷复合材料虽没有多少品种达到临床应用阶段，但它已成为生物陶瓷研究中最活跃的领域。进入21世纪以来，生物陶瓷研究的一个重要特点是与生物技术相结合，在生物陶瓷中引入活体细胞与生长因子，以及赋予生物陶瓷以药理作用。

根据生物陶瓷植入生物体内引起的组织－材料反应，生物陶瓷可分为4种。①近于惰性生物陶瓷：暴露于生物环境中几乎不发生化学变化，引起的组织反应是材料周围形成厚度不同的包囊性纤维膜。典型的有氧化铝生物陶瓷、氧化锆生物陶瓷，玻璃碳、低温各相同性（LTI）碳和超低温各向同性（ULTI）碳等。②表面生物活性陶瓷：可与组织在界面上化学键性结合。有磷酸钙基生物陶瓷、生物活性玻璃陶瓷等。③可吸收生物陶瓷：可在生物环境作用下被降解和吸收，并随之为周围新生的组织所代替。有 β－磷酸三钙、硫酸钙生物陶瓷等。④生物陶瓷复合材料：所引起的组织－材料反应决定于其组成和结构。生物陶瓷的物理性能见表 1.3-32。

表 1.3-32 生物陶瓷的物理性能

材料 \ 性能	密度 /g·cm^{-3}	弹性模量 /GPa	抗弯强度 /MPa	抗压强度 /MPa	断裂韧度 K_{IC} /MPa·m$^{1/2}$	硬度 HV
多晶 Al_2O_3	>3.93	380	550	4 500	5~6	2 300
ZrO_2	5.6~6.12	200	1 200	>4 000	15	1 300
LTI 碳	1.5~2.2	17~28	275~620			
生物玻璃 45S5			101~152	507		
A-W 生物活性玻璃陶瓷	3.04	117	180	922	2.0	
羟基磷灰石生物活性陶瓷	2.97~3.13	36~105	100~200	517~932	1.0	640
皮质骨	1.6~2.1	7~25	50~150		2~12	340
牙釉质		80		390		240

8.4 先进陶瓷产业及发展趋势

由于先进陶瓷拥有某些特殊性能，它在许多高新技术产业领域中，如电子信息、航空航天、汽车、能源、生物和国防等，有着广泛的应用。世界主要国家和地区的先进陶瓷销售额见表 1.3-33。美国的先进陶瓷市场见表 1.3-34。

表1.3-33　世界主要国家先进（特种）陶瓷销售额

亿美元

国别	1985年	1995年	2000年	年增长率/%
美国	13	55	87	9.5
日本	28	90	175	5.1
欧洲	6	11	18	10
合计	47	156	280	

表1.3-34　美国先进（特种）陶瓷市场①（1995～2000年）

亿美元

品种	1995年	所占比例/%	2000年	所占比例/%	年平均增长率/%
结构陶瓷	5.0	9.1	8.0	9.2	9.9
电子陶瓷	42.15	76.2	65.73	75.4	9.3
环境/工业	2.4	4.3	4.0	4.6	10.8
陶瓷涂层	5.75	10.4	9.4	10.8	10.2
合　计	55.30	100	87.13	100	9.5

① Ceramic Industry，October 1996。

据有关行业协会统计，我国先进陶瓷（特种陶瓷）的产值1995年约80亿人民币，2000年以来超过了100亿人民币，产品结构中电子陶瓷约占70%，结构陶瓷约为30%。

综合国内外先进陶瓷的发展趋势主要表现在以电子陶瓷为主的功能陶瓷发展迅速，电子陶瓷要充分满足电子元器件关于小型化、多功能化和集成化的发展要求，要加强先进陶瓷材料微观结构和性能设计及工艺控制技术的研究。在结构陶瓷方面，加强陶瓷复合、改性和复相陶瓷研究，如陶瓷的自增韧、纳米及亚微米复合技术、梯度复合技术等以改善陶瓷的脆性和可靠性。能源领域用固体电解质燃料电池用陶瓷材料、环境友好陶瓷材料和生物活性医用陶瓷材料等新兴特种陶瓷材料将获得更快的发展。

9　先进复合材料

复合材料指根据应用的需要进行设计，把两种或两种以上的有机高聚物材料、或无机非金属材料、或金属材料组合在一起，使之互补性能优势，从而制成的一类新型材料。复合材料一般由基体组元与增强体或功能体组元所组成，属于多相材料范畴。复合材料按其性能高低的层次分为常用复合材料和先进复合材料。常用复合材料包括普通玻璃纤维、合成与天然纤维等增强低性能高聚物（树脂）的复合材料。先进复合材料则以碳、芳纶、陶瓷纤维、晶须等高性能增强体与耐高温高聚物、金属、陶瓷和碳（石墨）等基体所构成的复合材料。复合材料按用途可以分为结构复合材料和功能复合材料两大类。

9.1　结构复合材料

结构复合材料主要用作承力结构和次承力结构等主要要求质量轻、强度和刚度高、耐温、耐蚀的材料等。结构复合材料基本上由增强体和基体组成，通常按所用的基体材料进行分类，如聚合物（树脂）基复合材料、金属基复合材料、陶瓷基复合材料、玻璃基复合材料、碳基复合材料和水泥基复合材料等，如图1.3-13所示。部分常用高性能增强体的性能见表1.3-35。结构复合材料与其他材料的比强度、比模量的对比见图1.3-14。

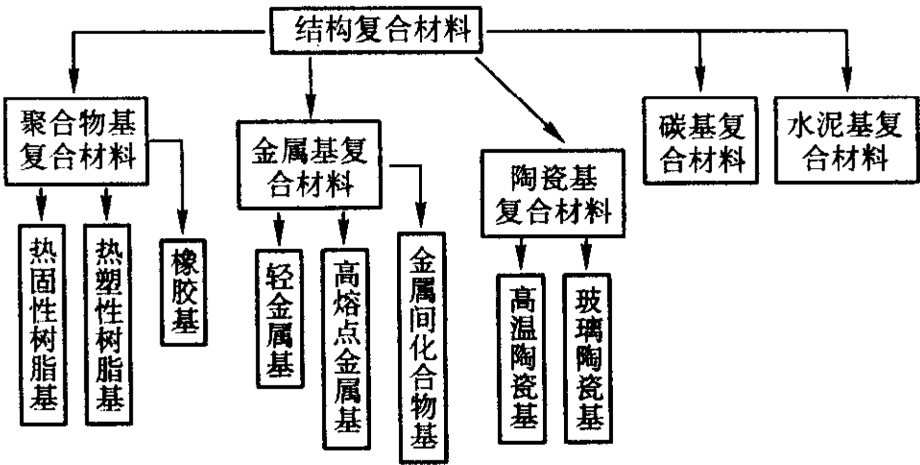


图1.3-13　结构复合材料按不同基体的分类

表1.3-35　部分常用高性能增强体及新品种的性能

类别	型　号	直径/μm	密度/$g\cdot cm^{-3}$	抗拉强度/GPa	弹性模量/GPa
碳与石墨纤维	T300（PAN基碳纤维，高强度型）	6～7	1.76	3.5	230
	M40（PAN基石墨纤维，高模量型）	6～7	1.81	2.7	390
	T1000（PAN基碳纤维，超高强度型）	6～7	1.72	7.2	220
	M60J（PAN基石墨纤维，高强高模型）	6～7	1.94	3.8	590
	P120（沥青基石墨纤维，超高模量型）	7	2.10	2.1	810
有机高分子纤维	Kevlar 49（聚芳酰胺纤维）	12	1.54	3.9	120
	Kevlar 149（聚芳酰胺纤维）	12	1.54	3.1	146
	Ekond　（聚芳酯纤维）	10	1.6	3.9	138
无机非金属纤维	Borsic（硼纤维，钨芯CVD法）	147	3.32	3.24	400
	SCS－6（碳化硅纤维，碳芯CVD法）	142	3.44	3.4	365
	Nicalon（碳化硅纤维，先驱体法）	10～15	2.55	3.0	200
	Tyranno（含钛碳化硅纤维，先驱体法）	7	2.4	3.0	220
	EP（三氧化二铝纤维，先驱体法）	20	3.85	1.37	382
晶须	SiC（β晶形）	0.1～1	3.19	>7	>700

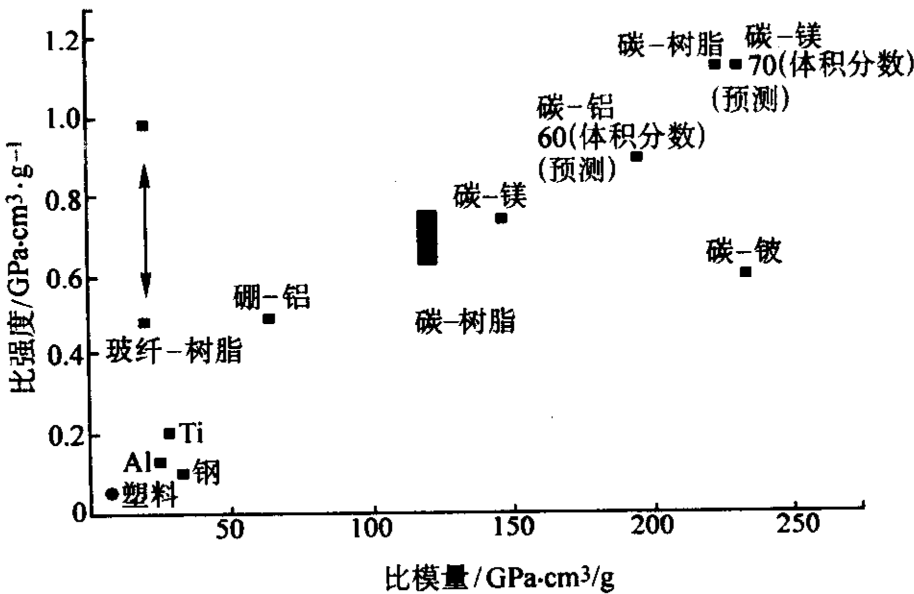


图1.3-14　复合材料与其他材料的比强度与比模量的对比图

9.1.1 聚合物基复合材料

是目前复合材料的主要品种，其产量远远超过其他种类的复合材料。主要有热固性聚合物（树脂）基复合材料和热塑性聚合物（树脂）基复合材料两大类。几种聚合物基复合材料的比强度和比模量及其与传统材料的比较见表 1.3-36。

表 1.3-36 几种复合材料的比强度和比模量及其与传统材料的比较

材料名称	密度 $/g\cdot cm^{-3}$	抗拉强度/MPa	弹性模量/GPa	比强度 $/MPa\cdot cm^3\cdot g^{-1}$	比模量 $/GPa\cdot cm^3\cdot g^{-1}$
铝合金	2.78	393	72.0	141.3	25.9
结构用钢	7.85	413	206	52.6	26.3
钛合金	4.52	712	116.7	157.5	25.8
玻璃纤维－聚酯复合材料	2.0	1 245	48.2	623	24.1
高强度碳纤维－环氧复合材料	1.45	1 471	137.3	1 014	94.7
高模量碳纤维－环氧复合材料	1.60	1 049	235	656	146.9
Kevlar 纤维－环氧复合材料	1.40	1 373	78.4	981	56.0

热固性树脂是由某些低分子的合成树脂在加热、固化剂或紫外光等作用下，发生交联反应并经过凝胶化阶段和固化阶段形成不熔、不熔固体，因此必须在原材料胶凝化之前成型，否则就无法加工。热固性树脂早期有酚醛树脂，随后有不饱和聚酯树脂和环氧树脂，近来又发展了性能更好的双马树脂和聚酰亚胺树脂。这些树脂适合于各类增强体。它们可以湿法成形（即浸渍后立即加工成形），但通常都先制成预浸料（包括预浸丝、布、带、片状和块状模塑料等），使浸入增强体的树脂处于半胶凝化阶段，在低温保护条件下限制固化反应的发展，并在一定时期内进行加工。所用的加工方法有手工辅设法、模压法、缠绕法、拉挤法、热压罐法、真空袋法和近期发展的树脂传递模塑法（RTM）和增强式反应注射成形法（RRIM）等。各种热固性树脂的固化反应机理各不相同，一般固化条件有常温固化、中温固化（120℃左右）和高温固化（170℃以上）。目前正在发展一种树脂体系（包括固化剂、促进剂等助剂）可以低温成形，然后在脱离模具的自由状态下加热后固化定型。

热塑性树脂基复合材料发展较晚，产量规模远不及热固性树脂基复合材料，它具有一些特出的优点，如断裂韧性好、吸湿性低和可以重复加工成型等，有很好的发展潜力。可作复合材料基体的热塑性树脂很多，包括各种通用塑料（如聚丙烯、聚氯乙烯等）、工程塑料（如尼龙、聚碳酸酯等）以及特种耐高温的聚合物（如聚醚醚酮、聚醚砜和杂环类聚合物）。热塑性树脂基复合材料必须先将树脂基体与增强体制成连续的片（布）状、带状、和粒状预浸料，才能加工成各种形状的复合材料构件。常用的预浸方法有薄膜法、溶液法、熔融法、粉末法、纤维法、造粒法等。

9.1.2 金属基复合材料

与树脂基复合材料相比，金属基复合材料具有耐高温、环境稳定性好、导热、导电等优点，迄今为止由于工艺复杂、成本较高还没有形成大规模产业应用，但具有很好的发展潜力和应用前景。金属基复合材料的基体有铝、镁、铜、钛、钢、高温合金和金属间化合物等各类金属及合金，而增强体则有连续纤维、短纤维、晶须和颗粒等。按基体分类金属基复合材料有铝基复合材料、镁基复合材料、铜基复合材料、钛基复合材料和金属间化合物基复合材料等。按增强体分类则有连续纤维增强金属基复合材料、短纤维增强金属基复合材料、晶须增强金属基复合材料、颗粒增强金属基复合材料、混杂增强金属基复合材料和骨架增强金属基复合材料等。

连续纤维增强金属基复合材料典型品种的基本力学性能见表 1.3-37。短纤维及晶须增强铝基复合材料典型品种的基本力学性能见表 1.3-38。颗粒增强铝基复合材料的基本力学性能见表 1.3-39。

表 1.3-37 连续纤维增强金属基复合材料典型品种的基本力学性能（纵向）

复合体系	纤维体积分数/%	材料密度 $/g\cdot cm^{-3}$	抗拉强度/MPa	弹性模量/GPa
硼纤维－铝	45～50	约 2.6	1 250～1 550	200～230
碳纤维－铝	50	2.4	760	140
高模石墨纤维－镁	38	1.8	510	300
碳化硅纤维－铝	45～50	约 3.0	1 250～1 600	210～240
碳化硅纤维－钛	35	3.9	1 210	260

表 1.3-38 短纤维及晶须增强铝基复合材料典型品种的基本力学性能

复合体系	增强体体积分数/%	抗拉强度/MPa	屈服强度/MPa	弹性模量/GPa	伸长率/%
$SiC_{(w)}$－6061Al	15	461	382	97	1.5
$SiC_{(w)}$－2024Al	17	539	461	107	1.3
$SiC_{(w)}$－2009Al	15	634(纵向)	483(纵向)	106(纵向)	6.4(纵向)
$SiC_{(w)}$－2009Al	15	552(横向)	400(横向)	98(横向)	8.4(横向)
氧化铝短纤维－铝	20	380	—	108	—

表 1.3-39 颗粒增强铝基复合材料的基本力学性能

复合体系	制备工艺	增强体体积分数/%	抗拉强度/MPa	屈服强度/MPa	弹性模量/GPa	伸长率/%
$SiC_{(p)}$－6061Al	粉末冶金法	20	496	414	103	5.5
$SiC_{(p)}$－2124Al	粉末冶金法	20	552	400	103	7.0
$SiC_{(p)}$－2009Al	粉末冶金法	20	572	421	109	5.3
$SiC_{(p)}$－7090Al	粉末冶金法	20	724	655	103	2.5
$SiC_{(p)}$－6061Al	粉末冶金法	40	441	—	123	0.7
$SiC_{(p)}$－A356	搅拌铸造法	20	350	320	98	0.5
$SiC_{(p)}$－A359	无压浸渗	30	382	—	125	0.4
$Al_2O_{3(p)}$－Al－10Mg	无压浸渗	51	531	—	161	0.63
$SiC_{(p)}$－Al－Si－Mg	无压浸渗	70	330（弯曲强度）	—	235	—

金属基复合材料的复合工艺有：

1）固相复合工艺包括：①粉末冶金工艺；②热压扩散结合工艺。

2）液相复合工艺包括：①搅拌复合工艺；②浸渗复合工艺：如挤压铸造、气压浸渗和无压浸渗等；③喷射共沉积；④熔体原位复合工艺：如放热弥散法（XD™）、接触反

应法、直接氧化法、气-液反应法（VSL法）和反应喷射沉积法等。

3）固-液两相区复合工艺包括：①流变铸造工艺；②固-液两相区热压复合工艺。

金属基复合材料的主要应用领域为航空航天工业，特别是航天及空间技术领域。经过近40年的发展，其应用领域已逐步扩展到交通运输业，如汽车工业，以及电子元器件等领域。

9.1.3　陶瓷基复合材料

这是以陶瓷、玻璃或玻璃陶瓷为基体，以某种结构形式引入颗粒、晶片、晶须或纤维等形状的增强体材料，通过适当复合工艺，从而改善或调整原基体材料的某种性能所获得的复合材料。按基体材料的不同陶瓷基复合材料可以分为：①玻璃、玻璃陶瓷基复合材料，其使用温度一般低于1 000℃；②高温陶瓷基复合材料，其使用温度一般可达1 000～1 400℃或更高。按照增强体的不同，陶瓷基复合材料可以分为：①长纤维补强陶瓷基复合材料；②短纤维、晶片、晶须陶瓷基复合材料；③颗粒弥散强化陶瓷基复合材料；④结构复合陶瓷基复合材料（如梯度功能复合、层状复合等）。

Nicalon-SiC增强玻璃基复合材料的力学性能见表1.3-40。几种SiC晶须增强LAS玻璃基复合材料的力学性能见表1.3-41。

表1.3-40　Nicalon-SiC增强玻璃基复合材料的力学性能

基体牌号	基体材料		复合材料		
	抗拉强度/MPa	弹性模量/GPa	抗拉强度/MPa	断裂应变/%	弹性模量/GPa
1723玻璃	100	60	734	1.2	128
LAS-Ⅱ	100～150	100	641～686	0.78～1.08	124～140
CAS-Ⅱ	—	97	450	0.95	124

表1.3-41　几种SiC晶须增强LAS玻璃基复合材料的力学性能

基体材料	晶须牌号	抗弯强度/MPa			室温断裂韧度/$MPa\cdot m^{1/2}$
		室温	800℃	1 000℃	
LAS		103	—	<69	0.85
LAS	AC-1	393	400	290	4.44
LAS	TK-1	338	496	248	4.20
LAS	TK-2	372	358	310	4.06

陶瓷基复合材料所用主要纤维的性能见表1.3-42。常见的一些非金属纤维补强陶瓷基复合材料的工艺与性能见表1.3-43。部分颗粒补强陶瓷基复合材料的性能见表1.3-44。一些层状陶瓷复合材料的力学性能见表1.3-45。

表1.3-42　陶瓷基复合材料所用主要纤维的性能

纤维		ρ/$g\cdot cm^{-3}$	σ/GPa	E/GPa	直径/μm	最高使用温度/℃
Al_2O_3	Fibre FP	3.9	1.38	380	21	1 315
	PRD 166	4.2	2.07	380	21	1 400
	Sumitomo	3.9	1.45	190	17	1 250
莫来石	Nextel 400	3.1	2.7	186	12	1 427
	Nextel 312	2.7	1.55	150	12	1 205
β-SiC	Nicalon	2.55	2.62	193	10	1 205
SiTiCO	Tyranno	2.5	2.76	193	10	1 300
Si_3N_4	TNSN	2.5	3.3	296	10	1 205
SiC单丝	SCS-6	3.05	3.45	410	140	1 300
	Sigma	3.4	3.45	410	100	1 260
纯石英玻璃	Astroquartz	2.2	3.45	69	9	980
石墨	T300	1.8	2.76	276	10	>1 650
	T40R	1.8	3.45	276	10	>1 650

表1.3-43　常见的一些非金属纤维-陶瓷基复合材料的工艺与性能

材料	工艺	抗弯强度/MPa	断裂韧度/$MPa\cdot m^{1/2}$
$SiC_{(f)}$-玻璃陶瓷		850	17
$SiC_{(f)}$-SiC	浸渍+反应烧结	800	
$SiC_{(f)}$-SiC	前驱陶瓷聚合物浸渍+热解	300（1 000℃） 250（1 300℃）	
$SiC_{(f)}$-SiC泡沫	纤维缠绕泡沫	4.8	
$SiC_{(f)}$-锆英石	1 610℃热压	700	
$SiC_{(f)}$-ZrO_2		200	25
BN涂层$SiC_{(f)}$-$ZrTiO_4$	热压	950（室温） 700（800℃） 400（1 200℃）	20 18.5 7.5
$SiC_{(f)}$-Al_2O_3	金属直接氧化法	461（室温） 488（1 200℃）	27.8 23.3
$SiC_{(f)}$-Si_3N_4	浸渍+反应烧结	750	
$C_{(f)}$-Si_3N_4		690（室温） 532（1 200℃）	28.1 41.8
SiC涂层$C_{(f)}$-Si_3N_4	金属直接氧化	392	18.5
$C_{(f)}$-SiO_2	定向缠绕+热压	152（室温） 103（1 200℃）	
$C_{(f)}$-莫来石		610（室温） 882（1 200℃）	18 18.2
$C_{(f)}$-SiC	浸渍+热解	400	15
$BN_{(f)}$-塞隆陶瓷	热压	600	5.5～6
5%$BN_{(f)}$-MgO	热压	130	
15%$BN_{(f)}$-MgO	热压	190	

表1.3-44　部分颗粒补强陶瓷基复合材料的性能

复合材料	抗弯强度/MPa		断裂韧度/$MPa\cdot m^{1/2}$
	室温	1 400℃	
$SiC_{(p)}$-TiC	580.0		6.5
$SiC_{(p)}$-TiB_2	490.0		8.9
$SiC_{(p)}$-Si_3N_4	930.0		7.0
Al_2O_3-$TiC_{(p)}$	940.0		4.0
Al_2O_3-$SiC_{(p)}$	1 000.0		

续表 1.3-44

复合材料	抗弯强度/MPa		断裂韧度/$MPa\cdot m^{1/2}$
	室温	1 400℃	
(Y)TZP[①]	1 402.4	49.0	
(Y)TZP - 40%(质量分数)Al_2O_3	1 802.5	201.0	
(Y)TZP - 20% Al_2O_3,30%(质量分数)TiC	1 375.9	527.6	

① (Y) TZP为以 Y_2O_3 作稳定剂的四方氧化锆多晶体陶瓷。

表 1.3-45 一些层状陶瓷复合材料的力学性能

材料体系		断裂韧度/$MPa\cdot m^{1/2}$	断裂功/$J\cdot m^{-2}$	抗弯强度/MPa
基体层	界面层			
SiC	石墨	17.7	6 152	633
Si_3N_4	BN + 12%(质量分数)Al_2O_3		6 500	437
β - Si_3N_4	BN + 10%(体积分数)Si_3N_4		4 500	530
Si_3N_4	BN纤维		4 700	280
Si_3N_4 - $SiC_{(w)}$	BN - Al_2O_3	20.4		651
SiC	BN - Al_2O_3 - SiC	20.9		677

陶瓷基复合材料以其高强度、高模量、高硬度、低密度、耐高温、耐腐蚀等特出优点，已在机械、电子、航空航天、医学工程等领域得到了很好的应用。随着陶瓷基复合材料在设计和制备技术上的进一步发展，其可靠性和耐久性获得进一步的提高，作为高温结构材料的应用将会不断扩大。

9.1.4 碳 - 碳复合材料

这是以碳纤维或石墨纤维为增强体，碳或石墨为基体的一类复合材料。它保留了碳、石墨材料的诸多优点，如密度低、高温强度高、热膨胀系数低、热导率高、抗热冲击性能好等，同时又克服了碳、石墨材料断裂韧性差的缺点，还可灵活设计结构、尺寸及性能，改变各向异性。主要用于航空航天领域，如航天飞行器鼻锥、飞机刹车系统、生物医学以及化学化工等领域。

碳 - 碳复合材料的生产流程图见图 1.3-15。碳 - 碳复合材料用碳纤维的典型性能见表 1.3-46。用于飞机刹车盘的碳 - 碳 复合材料性能见表 1.3-47。

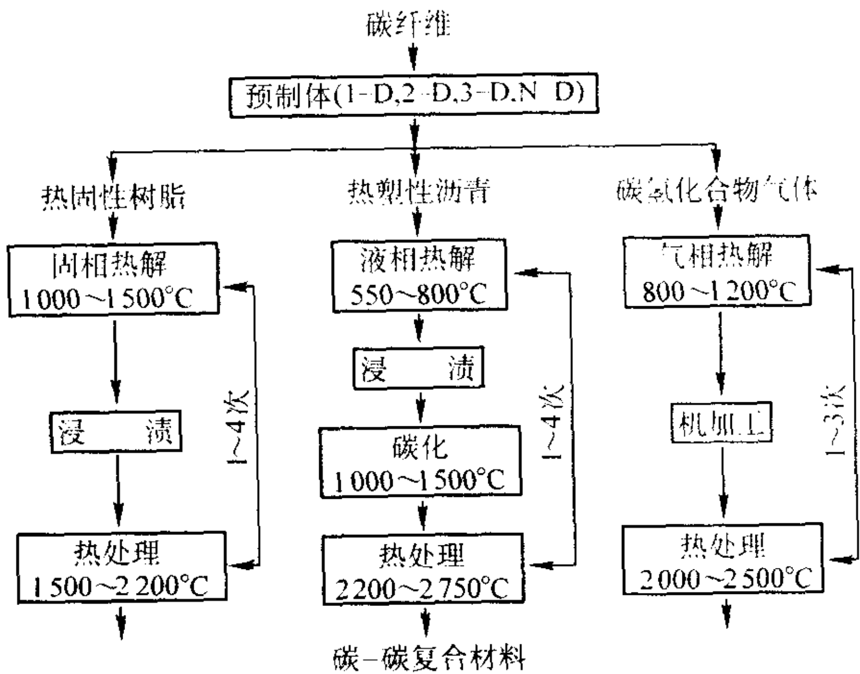


图 1.3-15 碳 - 碳复合材料的生产流程图

表 1.3-46 碳 - 碳复合材料用碳纤维的典型性能

性能	黏胶基		PAN基		沥青基
	低模量 - 低强度	高模量 - 高强度	高强度	高模量	
抗拉强度/GPa	0.62	2.2	3.1	2.4	2.1
拉伸弹性模量/GPa	41	393	230	390	380
密度/$g\cdot cm^{-3}$	1.53	1.66	1.73	1.81	2.0
纤维直径/μm	8.5	6.5	7	6.5	10
断裂伸长率/%	1.5	0.6	1.3	0.6	0.5
弹性回弹	100	100	100	100	100
碳含量/%	98.8	99.9	92	99	99
表面积/$m^2\cdot g^{-1}$	<4	1	1	1	1
热导率/$W\cdot m^{-1}\cdot K^{-1}$	38	122	2.1	70	100

表 1.3-47 用于飞机刹车盘的碳 - 碳复合材料的性能

性能	数值	性能	数值
密度/$g\cdot cm^{-3}$	1.72 ~ 1.78	抗压强度/MPa	120 ~ 140
开孔孔隙/%	6 ~ 10	层间抗剪强度/MPa	7.5 ~ 11.5
石墨化度/%	75 ~ 90	比热容/$J\cdot kg^{-1}$	1.2 ~ 1.5
抗弯强度/MPa	60 ~ 90	热导率/$W\cdot m^{-1}\cdot K^{-1}$	45 ~ 110

9.2 功能复合材料

功能复合材料指具有除力学性能以外的其他物理和化学性能的复合材料。按物理性质分有下列多种：电学方面有导电、超导、变阻、绝缘、离子导电、半导、吸收电磁波、屏蔽、压电、正温变系数 (PTC)、防静电等功能复合材料；磁学方面有永磁、软磁、磁致伸缩、压磁等功能复合材料；光学方面有透光、选择吸收、变频、抗激光等功能复合材料；热学方面有导热、绝热、低热膨胀系数、耐烧蚀等功能复合材料；声学方面有声纳发射和吸收、吸音等功能复合材料；其他还有阻尼、摩擦、抗磨损、自润滑、含能、装甲等功能复合材料；在化学方面则有化学分离，阻燃和化学反应等功能复合材料。预计未来功能复合材料的应用和品种会有很大的发展。主要功能复合材料的应用和分类见表 1.3-48。

表 1.3-48 主要功能复合材料的分类及应用

性质	复合材料品种	应用范围举例	组成举例
电学	压电复合材料	声纳、水听器生物成像传感器	PZT粉纤维（棒）- 聚合物，环氧，橡胶
	导电复合材料	弱电开关，抗腐蚀电极，厚膜电阻线路	导电炭黑 - 硅橡胶，金属粉 - 环氧（Ag，Ni，Cu）
	半导电复合材料	防静电地板、涂料	碳纤维，炭黑 - 环氧，橡胶
	电磁屏蔽复合材料	电子设备用屏蔽、板涂料	金属纤维，碳纤维，金属炭粉 - 环氧聚酯
	雷达隐身复合材料	军用飞机及导电蒙皮	炭黑 - 聚乙烯
	温控导电（PTC）复合材料	自控恒温发热体	

续表 1.3-48

性质	复合材料品种	应用范围举例	组成举例
磁学	永磁复合材料 软磁复合材料		永磁粉 - 橡胶，塑料 软磁铁氧体 - 聚合物
其他	透光复合材料 选择透光复合材料 透X射线复合材料 光致发光 辐射屏蔽复合材料 光致变色复合材料 吸声复合材料（空气） 吸声复合材料（水气） 阻尼复合材料 摩擦复合材料 热良导体复合材料 耐烧蚀复合材料 阻燃复合材料	农用温室顶板 智能窗 医用X光检查设备床板片 荧光显示屏 防射线（X光）板 变色镜，机敏窗 隔音板 消声瓦 机械减振器 轴承刹车片 大规模集成电路封装基板 导弹头部，航天飞机前缘 内装饰隔板	玻璃纤维 - 聚酯 氧化物 - 金属多层膜 - 玻璃 碳纤维 - 环氧 荧光粉 - 透明塑料 氧化锰氧化金色 - 玻璃 中空微球 - 塑料 空腔，多孔材料 - 橡胶，聚氨酯 颗粒，纤维 - 橡胶，塑料 金属板黏性塑料夹层，碳纤维 - 尼龙（轴承） 芳纶，铜丝 - 酚醛（刹车片） 碳化硅颗粒 - 铝 碳 - 碳，碳纤维 - 酚醛 卤化物，水合氧化锂 - 塑料

9.3 先进复合材料发展趋势

复合材料是一种相对年轻的产业，始建于20世纪中叶。所谓传统复合材料主要是指玻璃纤维增强塑料，也称玻璃钢。21世纪初全球玻璃钢产业的年产量规模约500万吨，中国2003年的产量约94万吨（其中热固性60万吨，热塑性22万吨，印刷电路板用12万吨）。就先进复合材料而论，有一定产业规模的主要是碳纤维增强树脂基复合材料。21世纪初，全球碳纤维的产能约31 000 t，年产量约17 000 t（分布情况：北美48%，欧洲28%，亚洲23%，南美1%）。全球碳纤维增强树脂基复合材料的产品结构中用于航空航天工业的约占23%、体育休闲产品约占21%、其他工业和产业约占56%。中国近年来碳纤维的年需求量约3 000 t，主要依靠进口，生产领域主要是体育休闲产品。

先进复合材料的发展原来是以航空航天产业的需求为主的，这方面仍然是一个重要的发展方向，特别是超大型民用客机如A380等的发展需求；随着碳纤维等增强体的品种增加和成本降低，基础建设中复合材料的用量将会逐步增加，如交通运输、桥梁、能源等领域；随着人民生活水平的提高，体育休闲用品中复合材料的用量将会不断增加，这方面制造业的重心正在向中国迁移；由于信息技术的飞速发展，在信息获取，传输、存储和处理等领域对功能复合材料的需求将不断增加，多功能、机敏、智能复合材料将是发展重点；此外关于环保、仿生以及纳米复合等新兴复合材料领域将构成复合材料发展的新生长点。

10 新能源材料

广义的能源包括多种矿物能源（煤、石油、天然气等）、水力、风能、地热、海洋能、核能、太阳能、生物质能等。其中氢能、太阳能、核能等被认为是在21世纪能得到广泛应用的新能源，他们一方面能补充矿物能源等的资源短缺，另一方面他们能解决由于矿物能源燃烧所引起的环境污染问题。表1.3-49是世界能源大会（1992年）关于1990～2020年矿物能源的预测，从中可见原油和天然气资源的需求和形势比较严峻。新能源的发展首先依赖于新能源材料的研发与应用，主要的比较现实的新能源材料有二次电池材料、燃料电池材料、太阳能电池材料及核能材料等，分别介绍如下。

表 1.3-49 1990～2020年矿物能源预测（折合标准油）

亿吨

项目	石油	天然气	煤	合计
探明储量（1991年底）	1 350	1 140	5 200	7 700
估计增加储量	650	1 060	28 800	30 500
合计资源量	2 000	2 200	34 000	38 200
累计需求量（1990～2020年）	1 000	680	850	2 530

10.1 新型二次电池材料

二次电池是指充电反应和放电反应为可逆式、即放电时通过化学反应产生电能、而充电时电能以化学能形式重新储存起来的电池。典型的二次电池体系见表1.3-50，其中铅酸电池和镍镉电池是早已广泛应用的二次电池，由于铅和镉有毒，这类电池所引起的环境污染问题严重，所以发展高比能，无污染的新型二次电池体系（如表中的金属氢化物镍电池Ni/MH和锂离子电池LIB）受到广泛的关注，它们也被称为“绿色电池”。

表 1.3-50 典型的二次电池体系

电池系列	负极活性物质	正极活性物质	电池反应	电性能（理论值）		
				电压/V	比容量/$A\cdot h\cdot kg^{-1}$	比能量/$W\cdot h\cdot kg^{-1}$
铅酸电池	Pb	PbO_2	$Pb + PbO_2 + 2H_2SO_4 = 2PbSO_4 + 2H_2O$	2.1	120	252.00
镉镍电池	Cd	NiOOH	$Cd + 2NiOOH + 2H_2O = 2Ni(OH)_2 + Cd(OH)_2$	1.35	181	244.35
锌镍电池	Zn	NiOOH	$Zn + 2NiOOH + 2H_2O = 2Ni(OH)_2 + Zn(OH)_2$	1.70	189	321.30
金属氢化物镍电池	MH	NiOOH	$MH + 2NiOOH = 2Ni(OH)_2 + M$	1.5	160	240.00
锂高温电池	Li(Al)	FeS	$2Li(Al) + FeS = Li_2S + Fe + 2Al$	1.33	345	458.85

续表 1.3-50

电池系列	负极活性物质	正极活性物质	电池反应	电性能（理论值）		
				电压/V	比容量/A·h·kg⁻¹	比能量/W·h·kg⁻¹
钠硫电池	Na	S	$2Na + 3S = Na_2S_3$	2.1	377	791.70
锂离子电池	LiC_6	CoO_2	$LiC_6 + CoO_2 = C_6 + LiCoO_2$	4.1	170	697.00

新型 Ni－MH 电池。LIB 电池与传统二次电池 Ni－Cd 的主要性能对比见表 1.3-51，由表可见新型二次电池除绿色环保优势外，在比能量和寿命方面都有改进和提高，在世界范围内的产销和应用逐年提高，在许多产品的应用上，目前新型二次电池的用量已经超过 Ni－Cd 电池的用量，见表 1.3-52。

表 1.3-51　Ni－Cd 电池、Ni－MH 电池和 LIB 电池主要性能对比

技术参数	Ni－Cd 电池	Ni－MH 电池	LIB 电池
工作电压/V	1.2	1.2	3.7
质量能量/W·h·kg⁻¹	30～50	50～70	100～150
体积能量/W·h·L⁻¹	150	200	270
充放电寿命/次	500	500	1 000
充电速率	1*C*	1*C*	1*C*

表 1.3-53 所示为电动汽车候选电池性能对比，从中可以看出锂离子电池具有较高的能量，100 A·h 极大容量锂离子电池已在国内外研发应用。

表 1.3-52　世界范围小型二次电池应用情况

百万只

年份	二次电池种类	便携式计算机	移动电话	摄录一体机	通信机械	声像机械	一般电器	其他	总计
1997	Ni-Cd	0	80	89	100	10	200	200	679
	Ni－MH	100	200	15	8	20	30	6	380
	LIB	135	45	19	11	10	—	—	220
	PLB	—	—	—	—	—	—	—	—
1998	Ni－Cd	0	70	74	90	10	185	180	609
	Ni－MH	85	210	20	12	20	23	20	400
	LIB	175	65	26	13	10	—	11	300
	PLB	—	5	—	—	—	—	—	5
1999	Ni－Cd	—	30	50	70	15	165	170	500
	Ni－MH	70	230	20	15	20	35	20	410
	LIB	228	85	30	15	10	2	30	400
	PLB	—	10	—	—	—	—	—	10
2000	Ni－Cd	—	30	50	70	15	165	170	500
	Ni－MH	70	230	20	15	20	35	20	410
	LIB	228	85	30	15	10	2	30	400
	PLB	—	10	—	—	—	—	—	10
2001	Ni－Cd	—	—	30	50	10	100	110	300
	Ni－MH	30	260	15	20	20	35	30	410
	LIB	320	110	40	30	10	10	50	570
	PLB	—	50	—	—	—	—	—	50

表 1.3-53　电动汽车候选电池性能比较

性能	铅蓄电池	Ni－Cd 电池	Ni－MH 电池	Na－S 电池	锂离子电池
质量能量/W·h·kg⁻¹	40 (60)	50 (60)	60 (80)	100 (110)	120 (180)
体积能量/W·h·L⁻¹	70 (100)	140 (180)	200 (250)	100 (100)	290 (360)
质量功率/W·kg⁻¹	150 (200)	170 (180)	170 (180)	150 (160)	56 (180)
循环寿命/次	400～1 000 (1 000 以上)	500 以上 (1 000 以上)	500 (1 000 以上)	350 以下 (1 000 以上)	500 以上 (1 000 以上)

10.1.1　金属氢化物镍（Ni－MH）电池材料

1）正极材料　Ni/MH 电池的容量为正极所限制，改进 $Ni(OH)_2$ 正极材料的性质对提高电池性能有重要意义，包括如何控制 $Ni(OH)_2$ 的形状、化学组成、粒径分布、结构缺陷及表面活性等关键因素，从而进一步提高正极的放电容量及循环稳定性。

2）负极材料（储氢合金）　主要有 AB_5 型混合稀土合金、AB_2 型 Laves 相合金、AB 型钛镍系合金、A_2B 型 Mg－Ni 系合金和钒基固溶体形合金等。其中 AB_5 型混合稀土合金由于具有良好的性能价格比已获广泛应用，对 AB_5 型混合稀土合金的改进工作主要集中在合金成分结构的优化以及表面改性处理等方面。此外对 AB_2 型合金和 A_2B 型合金也均在进一步研发之中。典型储氢电极合金的主要特性见表 1.3-54。

表 1.3-54　典型储氢电极合金的主要特性

合金类型	典型氢化物	合金组成	吸氢质量分数/%	电化学容量/mA·h·g⁻¹	
				理论值	实测值
AB_5 型	$LaNi_5H_6$	$MmNi_a(Mn, Al)_bCo_c$（a = 3.5～4.0，b = 0.3～0.8，$a+b+c=5$）	1.3	348	330
AB_2 型	$Ti_{1.2}Mn_{1.6}H_3$，$ZrMn_2H_3$	$Zr_{1-x}Ti_xNi_a(Mn, V)_b(Co, Fe, Cr)_c$（$a$ = 1.0～1.3，b = 0.5～0.8，c = 0.1～0.2，$a+b+c=2$）	1.8	482	420
AB 型	$TiFeH_2$，$TiCoH_2$	$ZrNi_{1.4}$，TiNi，$Ti_{1-x}Zr_xNi_a$（a = 0.5～1.0）	2.0	536	350
A_2B 型	Mg_2NiH_4	(MgNi)	3.6	965	500
固溶体型	$V_{0.8}Ti_{0.2}H_{0.8}$	$V_{4-x}(Nb, Ta, Ti, Co)_xNi_{0.5}$	3.8	1 018	500

3）电池的再生利用　如何处理大量的经过使用后失效的Ni-MH电池废弃物从而消除对环境的污染并回收稀土元素和镍钴等金属获得再生利用是一项具有重要意义的任务，是今后发展新型二次电池中的一项关键技术。

10.1.2 锂离子二次电池材料

1）碳负极材料　最早使用金属锂作负极，但在充放电过程中由于发生树枝状沉积而引起短路。目前实用化的电池是用碳负极材料，发生锂离子的嵌入或脱嵌而实现充放电。使用的碳材料有硬碳、天然石墨或中间相微珠等。

2）纳米合金材料　为了克服金属锂负极的安全性，曾研究了许多合金体系，最近发现Sn、SnSb和SnAg等金属间化合物可使电池的循环性能明显改善，有望逐步获得应用。

3）正极材料　目前使用的正极材料是$LiCoO_2$。为了降低成本，研究了$LiMn_2O_4$以及双离子传递型聚合物正极材料。锂离子正极活性材料的性能对比见表1.3-55。

4）电解质材料　重点研究发展非水溶剂电解质方面以获得高的电池电压。开发聚合物电解质是锂离子电池的重要方向，它关系到薄型电池的发展。

表1.3-55　锂离子电池正极活性材料性能的比较

材料名称	理论比容量/$mA\cdot h\cdot g^{-1}$	实际比容量/$mA\cdot h\cdot g^{-1}$	密度/$g\cdot cm^{-3}$	价格比	特点
$LiCoO_2$	275	130~140	5.00	3	性能稳定，体积比能量高，放电平台平稳
$LiNiO_2$	274	170~180	4.78	2	高比容量，热稳定性较差，价格较低
$LiMn_2O_4$	148	100~120	4.28	1	低成本，比容量较低，高温循环和存放性能较差。安全性好

10.2 燃料电池材料

燃料电池（FC）是一种在等温下直接将储存在燃料和氧化剂中的化学能高效（50%~70%）且清洁地转化为电能的发电装置。

燃料电池的主要应用方向是汽车、航天航空飞行器、潜艇的动力源或组成区域供电。针对上述不同用途而研发的燃料电池有碱性氢氧电池（AFC）、磷酸型燃料电池（PAFC）、质子交换膜型燃料电池（PEMFC）、熔融碳酸盐型燃料电池（MCFC）和固体氧化物燃料电池（SOFC）等。燃料电池的主要类型、特征与应用见表1.3-56。燃料电池材料的发展主要围绕提高燃料的发电效率、延长电池的工作寿命，降低发电成本等目标来进行。

表1.3-56　燃料电池的类型与特征

类型	电解质	导电离子	工作温度/℃	燃料	氧化剂	技术状态	可能应用领域
碱性（AFC）	KOH	OH^-	50~200	纯氢	纯氧	高度发展，高效	航天，特殊地面应用
质子交换膜（PEMFC）	全氟磺酸膜	H^+	室温~100	氢气，重整氢	空气	高度发展，需降低成本	电动汽车，潜艇推动，可移动动力源
磷酸（PAFC）	H_3PO_4	H^+	100~200	重整气	空气	高度发展，成本高，余热利用价值低	特殊需求，区域性供电
熔融碳酸盐（MCFC）	(Li，K) CO_3	CO_3^{2-}	650~700	净化煤气，天然气，重整气	空气	正在进行现场实验，需延长寿命	区域性供电
固体氧化物（SOFC）	氧化钇稳定的氧化锆	O^{2-}	900~1 000	净化煤气，天然气	空气	电池结构选择，开发廉价制备技术	区域供电，联合循环发电

1）质子交换膜型燃料电池（PEMFC）材料　使用Pt-C或Pt-Ru/C作电催化剂，以提高Pt分散度，并向电极催化层中渗入Nafion树脂，实现电极的立体化以提高铂的利用率，使铂的用量降至原来的10%~5%。此外用金属双极板取代目前使用的无孔石墨板，可以降低厚度，提高电池比功率，降低电池成本。我国组装了35节1 000~1 500 W燃料电池组，其特征性能见表1.3-57。

表1.3-57　kW级PEMFC电池组特征

项目	性能与指标	项目	性能与指标
电池组输出功率/kW	1~1.5	H_2纯度/%	>99.0
单池节数	35	H_2/O_2工作压力/MPa	0.25~0.45/0.30~0.5
电池组输出电流/A	40~69	电池工作温度/℃	室温~100
电极工作电流密度/$mA\cdot cm^{-2}$	300~530	电池启动时间	数秒钟
电池输出电压/V	27~23	电池组能量效率/%	52

2）熔融碳酸盐型燃料电池（MCFC）材料　MCFC的工作温度约650℃，电催化剂以镍为主。有关电池材料的关键问题是在工作过程中阴极材料（NiO）发生溶解，阳极材料（烧结镍）发生蠕变，以及双极板材料发生腐蚀等问题。解决的办法是研制新型阴极材料如$LiCoO_2$等以延长电池寿命，使用Ni-Cr或Ni-Al合金取代烧结镍作阳极材料以改进蠕变性能，以及对双极板导电部分用Ni-Cr-Fe-Al耐热合金包覆、非导电部分用Al包覆以提高其抗腐蚀性能。

3）固体氧化物燃料电池（SOFC）材料　由于SOFC的电解质为固体，无电解液流失问题，燃料的适用范围广、其综合利用率高。通过发展薄氧化钇、稳定化氧化锆（YSZ）膜技术探索新兴中温电解质。对于管式SOFC，正在研发价廉的YSZ膜制备工艺，以降低电池成本。

10.3 太阳电池材料

太阳能是人类最主要的可再生能源。太阳能电池发展的制约因素有：太阳能接受面积问题、能量按时间分布不均匀问题、电池材料的资源问题以及成本问题等。太阳能电池材料的发展主要针对提高转换效率、节约材料消耗、降低成本等问题进行研究。

1）发展材料工艺，提高转换效率　太阳电池用的半导体材料见表1.3-58。材料工艺包括材料提纯工艺、晶体生长工艺、晶片表面处理工艺、薄膜制备工艺、异质结生长工

艺、量子阱制备工艺等。通过上述材料制备工艺的研究改进，使太阳电池的转换效率不断提高。太阳电池材料制备的主要方法见表 1.3-59。单晶硅电池的转换效率已达 23.7%，多晶硅电池达 18.6%，砷化镓基电池可达 30%。

表 1.3-58　太阳电池的半导体材料

材料名称	禁带宽度/eV	禁带性质	迁移率/$cm^2 \cdot V^{-1} \cdot s^{-1}$		晶系	晶格常数/nm	应用实况
			电子	空穴			
晶体 Si	1.12	间接	1 500	450	立方	$a=0.543$	制作的电池占市场份额的 70%～80%
非晶 Si	1.5～2.0		约 1	0.1			制作的电池占市场份额的 10%～20%
Ge	0.66	间接	3 900	1 900	立方	$a=0.5646$	用作空间电池的衬底
GaAs	1.424	直接	8 500	400	立方	$a=0.5653$	已开始用于空间太阳电池
InP	1.35	直接	4 600	150	立方	$a=0.5869$	耐辐射性能优异，处于研究开发阶段
AlSb	1.6	间接	900	400	立方	$a=0.37$	禁带宽度适合太阳电池，但因材料腐蚀，未获应用
CdS	2.42	直接	340	—	立方	$a=0.4136$ $c=0.6176$	构成薄膜电池一极
Cu_2S	1.2	间接	—	30	立方	$a=1.188$ $c=1.349$	与 CdS 构成的太阳电池因出现衰退现象，已被淘汰
CdTe	1.44	直接	700	65	立方	$a=0.6477$	独自制作薄膜电池或与 CdS 结合，构成的太阳电池已商品化
$CuInSe_2$	1.04	直接	300	20	正方	$a=0.5782$ $c=1.162$	与 CdS 构成的太阳电池正进入商品化

表 1.3-59　太阳电池材料制备的主要方法

方法名称	方法特征	应用对象	开发状况
直拉法	①获得单晶；②可使材料进一步提纯；③制作成本较高；④需切片；⑤片为圆形	单晶硅电池	已大批量生产；现多使用 150 mm 单晶
铸锭法	①成本低；②可制作大锭以提高效率，降低成本；③可使材料进一步提纯；④需切片	多晶硅太阳电池	已大批量生产
蹼晶法	①制出叠层单晶体；②不需切片；③宽度小于 10 cm	准单晶硅电池	小批量制作
定边喂模法（EFG）	①制作效率高；②受模具材料污染	多晶硅电池	开始批量生产
涂布烧结法	①方法与设备简单；②厚度与质量难控制	CdS/CdTe 薄膜电池	小批量制作
脉冲电沉积法	①方法与设备简单；②有水溶液造成环境污染	CdS/CdTe 薄膜电池	小批量制作
双源蒸发法	①控制组分良好；②设备与工艺较复杂	$CdS/CuInSe_2$ 电池	技术开发
CVD 法	①设备较简单；②质量参数难以控制	多晶硅薄膜电池	技术开发
辉光放电 CVD 法	①制出能掺杂非晶体；②可大面积制作；③难控制电池的稳定性	非晶硅薄膜电池	技术开发
液相外延法	①设备较简单；②成本较低；③难控制厚度	GsAs 系薄膜电池	小批量制作
MOCVD	①可精确控制组分与厚度；②设备复杂	GaAs 系薄膜电池	小批量制作
分子束外延	①可在原子尺度上控制生长；②设备复杂；③生长慢	量子阱结构电池	试验研究

2）发展薄膜电池，节约材料消耗　多晶硅属间接禁带结构，需较大厚度（一般为 200～300 μm）才能充分吸收太阳能，最近多晶硅薄膜电池的有源层厚度已降到 50 μm。其他薄膜电池如砷化镓电池、碲化镉电池、非晶硅电池等，只需 1～2 μm 的有源层厚度，从而可大幅度降低材料消耗。

3）材料的大规模加工技术　如每炉超过 200 kg 的多晶硅锭生产技术，采用线切割切硅片技术等。

4）与建筑结构相结合的技术　除在屋顶架设太阳电池板之外，将太阳电池做在建筑材料上是值得重视的，如非晶硅太阳电池做在屋顶瓦、半透明或不透明垂直幕墙、窗用玻璃等建筑上。这是材料工艺与设备发展的结果。

进入 21 世纪以来，世界各国均提出大力发展太阳电池的规划，如美国的百万屋顶计划见表 1.3-60。德国于 1999 年开始了 10 万屋顶计划，预计在 6～8 年内完成，每个系统为 5 kW，总容量为 500 MW 的光伏市场。日本于 1997 年已建立 1 600 个屋顶系统，总容量为 37 MW，目前正在不断扩大。我国拥有丰富的太阳能资源，陆地表面每年接受的太阳辐射能为 50×10^{18} kJ，相当于 1 700 亿吨标准煤，每年日照时间大于 2 000 h、辐射总量高于 586 kJ/m^2 的太阳能资源丰富和较丰富的地区占全国总面积的 2/3，而且大部分贫困和无电地区分布在这一范围内，因此我国太阳电池的潜在市场很大，我国太阳电池在农村的潜在市场见表 1.3-61。

表 1.3-60 美国百万屋顶计划的内容与指标

指标 \ 年代	1997	1998	1999	2000	2005	2010
参加城市数	10	25	50	75	200	325
太阳能建筑物/10^3 个	2	8.5	23.5	51	376	1 014
屋顶当量系统/kW	1	1	1	2	3	4
光伏总容量/MW	1	6.5	15	55	270	610
单位成本/美元·W^{-1}	6.5	5.7	4.9	4.3	2.9	2.0
对应的 CO_2 减少/kt	2	11	26	76	313	663
总的 CO_2 的减少/kt	2	13	39	111	1 037	3 510
创立的工作岗位/个	300	1 800	3 800	11 000	40 000	71 500

表 1.3-61 中国光伏发电在农村电气化的潜在市场

省份	已安装的光伏系统/套	已安装的光伏发电功率/kW	无电用户/户	潜在需求/套	2000年以前的市场/MW	2001~2010年光伏市场/MW
西藏	500	80	120 000	110 000 (92%)	2.2	11.0
青海	6 000	100	180 000	150 000 (83%)	3.0	15.0
内蒙古	7 000	120	400 000	200 000 (50%)	4.0	20.0
新疆	7 000	120	490 000	150 000 (30%)	3.0	15.0
甘肃	7 000	50	400 000	240 000 (60%)	4.8	24.0
宁夏	100	18	200 000	120 000 (60%)	2.4	12.0
陕西	1 000	15	300 000	120 000 (40%)	2.4	12.0
其他省份	20 000	350	约2 000万	200 000 (1%)	4.0	20.0
总计	48 600	853	约2 190万	1 290 000 (5.9%)	25.8	129.0

10.4 核能材料

核能材料指各类核能系统主要构件用的材料。核反应堆结构材料可分为堆芯结构材料和堆芯外结构材料。堆芯结构材料有原子能所特有的辐照效应问题，而堆芯外结构材料则无辐照效应问题而与一般结构材料相同。核反应堆可分为裂变反应堆和聚变反应堆两大类，裂变反应堆已大量应用，聚变反应堆仍需数十年的研究发展才能进入商业应用。

10.4.1 裂变反应堆材料

U-235 或 Pu-239 原子核在吸收一个中子后发生裂变，分裂成两个质量大致相同的原子核，并放出 2~3 个中子，并继续产生 U-235 或 Pu-239，原子核裂变链式反应叫核裂变。根据引起燃料核裂变的中子的能量，裂变反应堆可以分为快堆、中能堆和热堆；根据所用燃料的种类可以分为铀堆、钚堆、钍堆和混合堆；根据用于慢化中子的材料，可以分为轻水堆、重水堆、石墨堆和有机介质堆。冷却剂是区分反应堆的一个重要特性，可采用轻水、重水、液态金属、气体、有机介质作冷却剂。根据目的和用途，反应堆有动力堆和生产放射性同位素用堆。动力堆又有固定式的（核电站）和移动式的（船舶用堆、飞机和火箭用堆）。目前应用最广泛的是水冷却（主要是轻水）、加浓铀做燃料的核电站，它被证实技术可靠、经济上有竞争力。水冷堆中应用广泛的是压水堆，典型的压水堆核电厂系统见图 1.3-16。

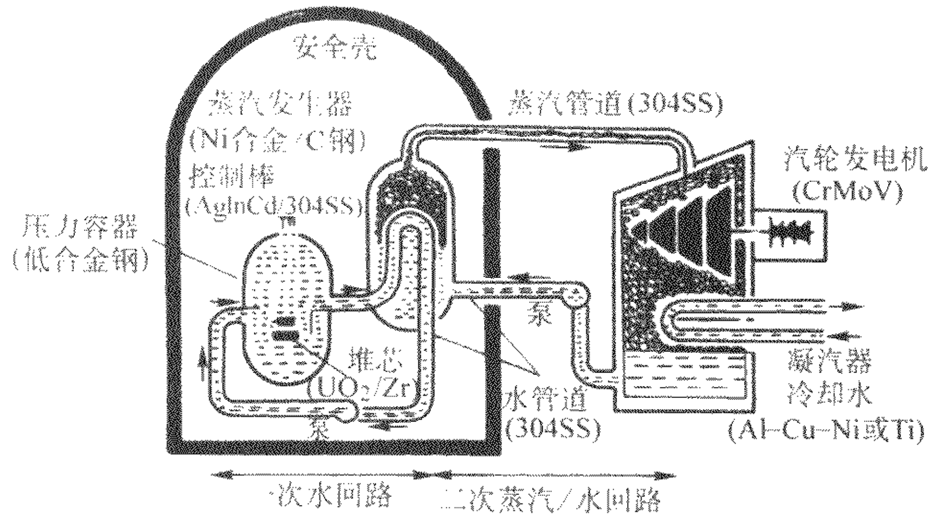


图 1.3-16 压水堆核电厂系统

裂变反应堆材料主要包括：

1）燃料元件用材料 主要是燃料芯体和燃料包壳材料（锆合金等）。

2）慢化剂材料 主要有氢、氘、铍和石墨。

3）控制材料 主要有铪、B_4C、Ag-In-Cd、硼硅酸玻璃等。

4）冷却剂材料 主要有轻水（H_2O）、重水（D_2O）、CO_2、He 等。

5）反射层材料 铍块、石墨块、水等。

6）屏蔽材料 屏蔽 γ 射线选用铁、铅、重混凝土等。屏蔽热中子选用中子吸收材料如硼钢、B_4C-Al 复合材料等。

7）反应堆容器材料 一般选用高强钢，如 A-508 等。

10.4.2 聚变堆材料

两个轻原子核融合形成重原子核叫做核聚变。典型的托卡马克聚变堆示意图见图 1.3-17。聚变堆所用的主要材料如下。

1）聚变核燃料 主要是氘（D）和氚（T）。

2）氚（T）增殖材料 通过锂与中子反应生成氚，这种材料主要有 Al-Li 合金、陶瓷型的 LiO_2、偏铝酸锂（$LiAlO_2$）、偏锆酸锂（Li_2ZrO_3），还有液态锂铅合金、锂铍氟化物熔盐等。

3）中子倍增材料 这是含有能产生（n，2n）和（n+n）

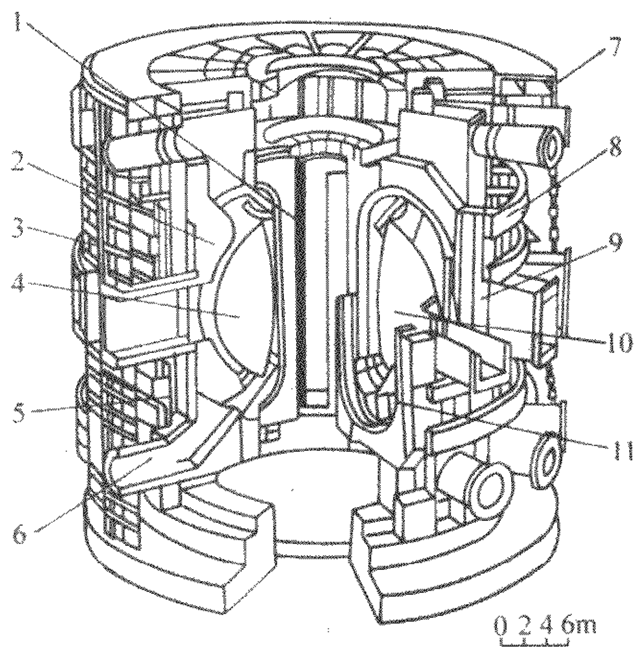


图 1.3-17 托卡马克聚变堆示意图

1—中央螺线管；2—屏蔽/包层；3—活动线圈；4—等离子体；5—真空容器屏蔽；6—等离子体排出；7—低温恒温器；8—轴向场线圈；9—环向场线圈；10—第一壁；11—偏滤板

核反应的核素材料。含有 Be、Pb、Bi 和 Zr 的化合物或合金如 Zr_3Pb_2、PbO 和 Pb－Bi 等都可以作为中子倍增材料

4）第一壁材料　第一壁是托卡马克聚变堆装置包容等离子体区和真空区的部件，又称面向等离子体部件，它与外围的氚增殖区结构紧密相连。第一壁材料主要包括第一壁表面覆盖材料、如铍、石墨、碳化硅，以及碳－碳、碳－碳化硅纤维强化复合材料。第一壁结构材料要在高温、高中子负荷下有适当的工作寿命。目前选择的有奥氏体不锈钢（AISI 316、PCA）、铁素体不锈钢（HT9）、钒（V）、钛（Ti）、铌（Nb）和钼（Mo）等合金。第一壁材料还包括高热流材料，低活化材料等。

除上述材料外，聚变堆还有电绝缘材料、超导磁体芯线、磁体支撑部件、激光窗口材料、辐射屏蔽材料、冷却剂材料等。

11　超导材料

在一定条件下（温度、磁场、电流等）具有超导电性的材料为超导材料。超导材料因其零电阻性、完全抗磁性等优异特性而具有重大的应用前景。超导材料的研究发展分为二个阶段：第一阶段（1911～1986年）是低温超导体和材料发展阶段，从1911年发现汞（Hg）开始，到1980年发现有机超导体，在这些低温超导体中，临界转变温度（T_c）最高的是 Nb_3Ge，为23.2 K；第二阶段为1986年 K.A 米勒（Muler）和 J.G. 贝德诺尔茨（Bednorz）在陶瓷氧化物中发现高临界温度超导体为标志，使超导材料的研究由液氦温度一下跃升至液氮温度，从而开始了高温超导材料的研究阶段，1987年中国赵忠贤、美国朱经武、日本田中昭二各自独立研制的 YBCuO 系超导体的 T_c 达到93 K。超导材料临界温度随年代的变化见图1.3-18。

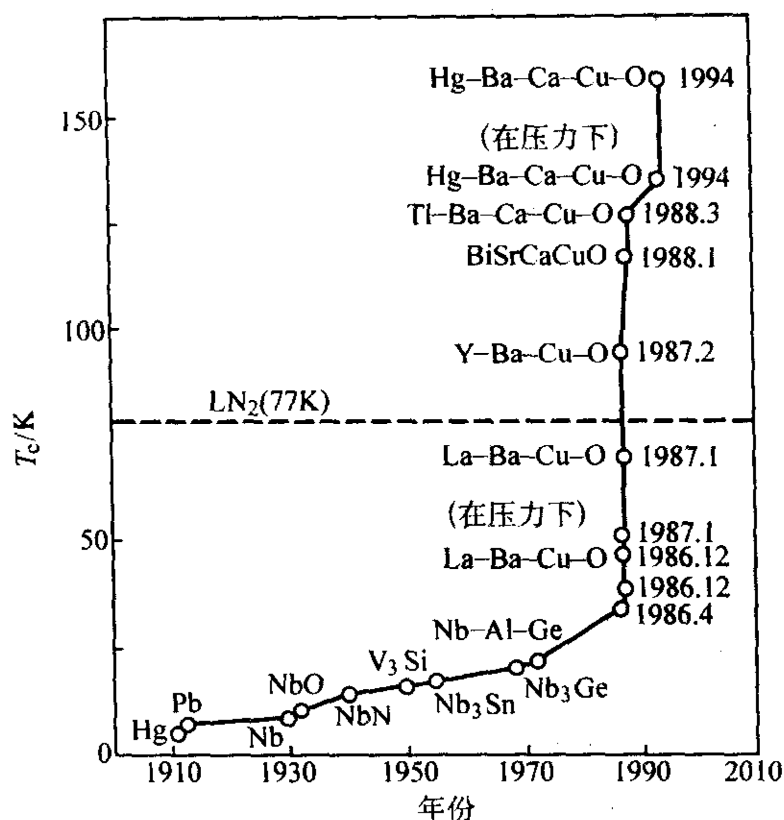


图1.3-18　超导材料临界温度随年代的变化

超导体大致可分为元素、合金和化合物超导体，包括：①金属元素超导体（Hg、Pb、In、Sn、Ga等）；②化合物超导体（AB、AB_2 等）；③合金超导体（NbTi等）；④B1化合物超导体（NbN等）；⑤A15型化合物超导体（Nb_3Sn、V_3Ga 等）；⑥谢弗尔相超导体（$A_xMo_6X_8$）；⑦拉夫斯相超导体（$MgNi_2$、$MgCu_2$ 等）；⑧重电子系统超导体（$CeCu_2Si_2$ 等）；⑨其他超导体（电荷转移有机物、硫族化物和氧化物、钡铅铋钙钛矿氧化物、钡钾铋立方钙钛矿氧化物、富勒烯超导体等）；⑩高温氧化物超导体（YBCO、BSCCO、TBCCO等）。

11.1　实用 NbTi 合金超导材料

NbTi 超导合金从20世纪60年代被发现以来得到了广泛的应用，主要因为它具有高强度、良好加工塑性和良好超导性。NbTi 合金的相图见图1.3-19，超导性能见图1.3-20，NbTi 多芯超导线制造工艺流程见图1.3-21。各种工艺条件下 NbTi 超导线的临界电流密度与磁场的关系曲线见图1.3-22。

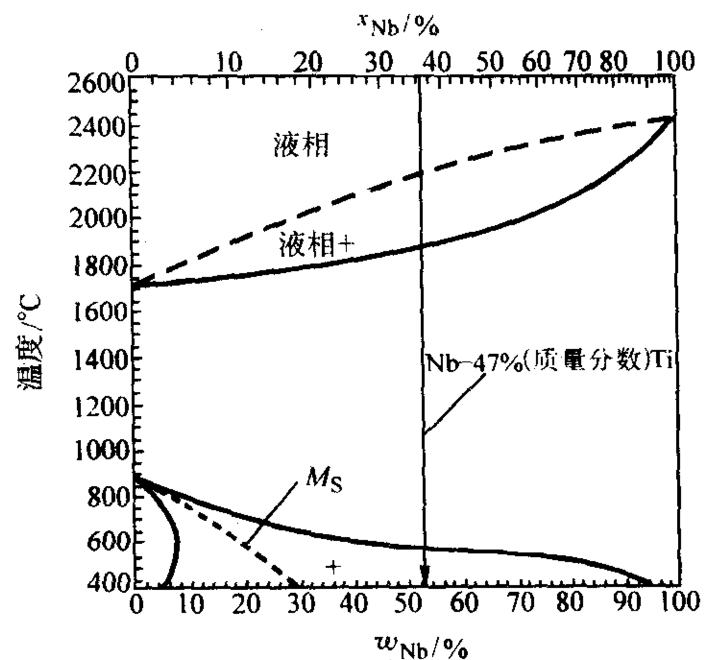


图1.3-19　NbTi 二元合金相图

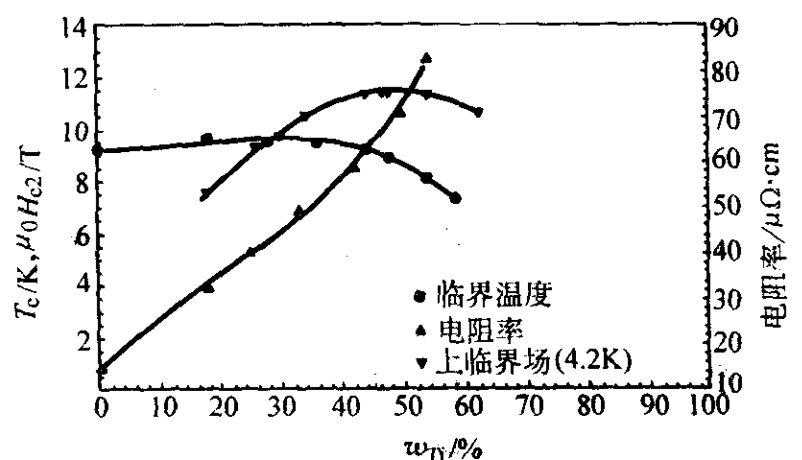


图1.3-20　NbTi 合金在4.2 K下的 T_c，ρ 及上临界场 μ_0H_{c2} 以及293 K下的电阻率 ρ 随合金成分变化的曲线

目前 NbTi 超导材料主要应用于磁共振成像（MRI）、核磁共振（NMR）、实验仪器、超导储能（SMES）、粒子加速器、电力、扫雷、矿石磁分离、磁悬浮列车等，其中 MRI 每年消耗的 NbTi 超导线约1 000 t左右。

11.2　实用 Nb_3Sn 超导材料

在已知超导 A15 化合物中，迄今只有 Nb_3Sn 和 V_3Ga 已发展为商品化导体，主要是因为他们兼具高 T_c、高 H_{c2} 和高 J_c 的结合。Nb_3Sn 是一种具有 A15 晶体结构的铌锡金属间化合物，Nb_3Sn 的晶格常数为0.529 0 nm，密度为8.92 g/cm³，熔点为1 980℃ ± 10℃，4.2 K热导率为0.4 mW/（cm·K），4.2 K的正常电阻率为28 μΩ·cm。Nb_3Sn 具有较高的超导转变温度 T_c（18.3 K）和上临界磁场 H_{c2}（22.5 T/4.2 K），4.2 K/ 10 T磁场下能承载的临界电流密度 J_c 约为 10×10^5 A/cm²。因此 Nb_3Sn 是用于制作10 T以上超导磁体的主要材料，但因其脆性大，且硬度高，因而多年后才发展成为实用超导材料。其制造工艺有气相沉积法、扩散法和青铜法三类，自从20世纪70年代美国 A.R.Kauffman 等人发明青铜法制造多芯线材技术以来，青铜法一直是各种商品化 Nb_3Sn 材料的主要制造工艺。

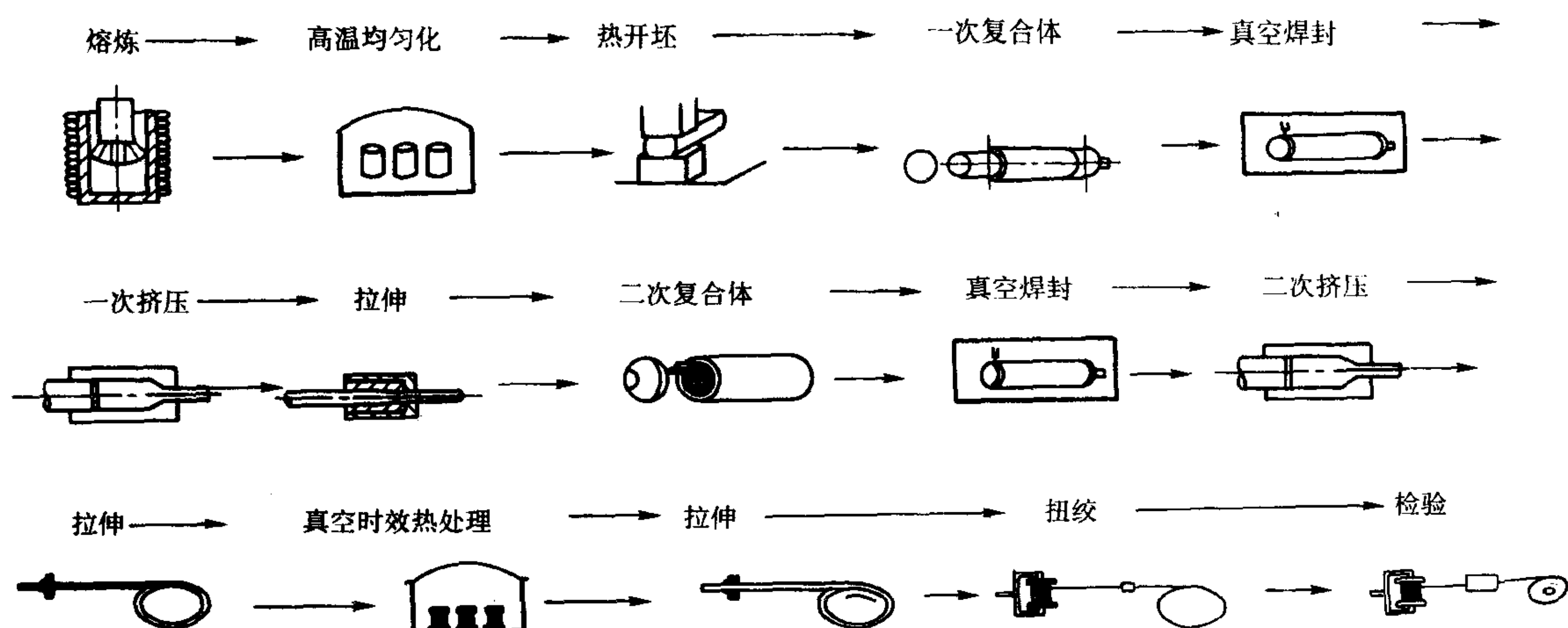


图 1.3-21 NbTi 多芯超导线制造工艺流程

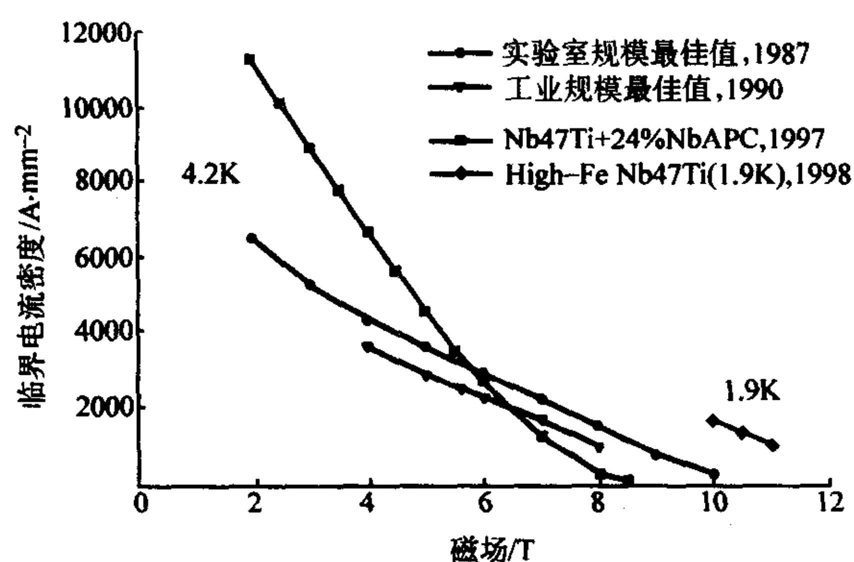


图 1.3-22 各种工艺条件下 NbTi 超导线的临界电流密度与磁场的关系曲线

11.3 高温氧化物超导材料

已发现的高温氧化物超导材料主要可归纳为钇系、铋系和铊系等。它们都是以钙钛矿结构为基的层状铜氧化物，都具有 Cu－O 离子层，结构中的二维铜原子面与高温超导电性直接有关。

1) $YBa_2Cu_3O_{7-x}$系（YBCO） 1987 年初美、中、日三国科学家相继独立地发现了 $T_c=92$ K 的氧化物超导材料 $YBa_2Cu_3O_7$，这是第一个液氮温度超导体，简称 YBCO 或 Y－123。实际上以 YBCO 为代表的高温氧化物超导体包括了 12 种 $RE_1Ba_2Cu_3O_{7-x}$ 氧化物超导体（也可简写为 REBCO 或 RE123），其中 RE 为镧系 La、Nd、Sm、Eu、Gd、Dy、Ho、Er、Tm、Yb 和 Lu 11 种元素，事实上 R^{3+} 稀土离子对 REBCO 的 T_c 影响很小，其超导转变温度均在 90 K 左右。RE123 氧化物超导体为空穴型载流子超导体，超导电性主要发生在 CuO_2 面层，即 CuO_2 面层为超导电层，其他多层构成载流子库层。与传统低温超导相比它们具有三个明显的不同点：①很高的超导转变温度，T_c 均在 90 K 左右；②很短的相干长度，ξ 为 0.15～2.5 nm；③高度的各向异性。YBCO 的下临界磁场μ。H_{c1}为 0.1 ～500 mT，上临界磁场 H_{c2}分别为 55～290 T（4.2 K）和 9～56 T（77 K）。RE123 块材的性能主要是指临界电流密度 J_c、捕获磁场强度和磁悬浮力密度。而影响这些性能的因素主要是微裂纹弱连接和磁通钉扎问题。要在元素掺杂、优化氧缺陷、改善孪晶密度等方面进一步改进 RE123 材料的性能。RE123 氧化物超导体块材的应用背景主要是超导轴承、超导飞轮储能、永磁、超导电机定子、磁滞电机转子、故障限流器、电流引线、弱场下的磁屏蔽等。

2) Bi－Sr－Ca－Cu－O 系（BSCCO） 这是由 Bi、Sr、Ca、Cu 四种或四种以上元素与氧组成的化合物超导体，主要有三个不同的 T_c 超导相：$Bi_2Sr_2CuO_6$（2201）、$Bi_2Sr_2CaCu_2O_8$（2212）、$Bi_2Sr_2Ca_2Cu_3O_{10}$（2223），其 T_c 分别为 10 K、85 K 和 110 K。Bi 系超导体往往同时含有上述三类超导相，纯的高 T_c 2223 相的制备尚有困难。用 Pb 部分取代 Bi 可以增加 2223 相的含量并改善晶粒间的连接。Bi 系超导体具有相干长度短、各向异性较高的特点，表现出较强的二维系统特性。在众多的高温超导材料体系中，Bi 系（Bi－2212 和 Bi－2223）超导材料由于其良好的加工和热处理性能，已被成功地用粉末装管法（PIT）制备成接近实用水平的线（带）材。现已用的 BSCCO/Ag 带或线成功制成了演示线圈，并且测试了它们的超导性能，原则证明了用高温超导 BSCCO/Ag 复合物材料制造超导磁体是可行的。

3) Tl－Ba－Ca－Cu－O 系（TBCCO） 是一种含 Tl（铊）的铜基氧化物超导体。$Tl_2Ba_2CuO_6$ 的 T_c 改为 90 K，$Tl_2BaCaCu_2O_8$ 的 T_c 约为 105 K，$Tl_2Ba_2Ca_2Cu_3O_{10}$ 的 T_c 约为 125 K，$Tl_2Ba_2Ca_3Cu_4O_{12}$的 T_c 约为 113 K。这些超导相的晶体结构与所对应的 Bi 系超导体相似。而 $Tl_2Ba_2Ca_2Cu_3O_{10}$（Tl－2223）具有很高的 T_c（125～128.7 K）。目前已用粉末套管法（PIT）通过拉伸最后轧制复合体成带材，已制成 Tl－1223/Ag 带材。TBCCO 系超导材料由于存在多相性以及高蒸气压等问题，使得单相材料的制备复杂化，而且 Tl 有毒性，造成此类材料研究发展上的限制，有待于在材料体系和制备工艺上的进一步突破。

11.4 MgB_2 超导材料

2001 年 1 月由日本 Jun.Akimitsu 首次发现宣布，MgB_2 的超导转变温度高达 $T_c=39$ K，几乎是 Nb_3Ge（A15 型金属间化合物超导体中 T_c 最高者，23.2 K）的两倍。MgB_2 具有 C32 结构，简单六方晶系二硼化物是各向同性的第二类超导体，其上临界场 H_{c2}（0）＝14～17 T，高于 NbTi 超导体。由于 MgB_2 具有高临界转变温度（属金属间化合物超导体的最高记录）而且结合了化合物的简单性，使其成为当前超导研究中的热点之一，开始了金属间化合物超导体的复兴。初步的研究表明 MgB_2 存在两点不足，一是低温下低的上临界磁场 H_{c2}，至今报道在 4.2 K 下 H_{c2}最高为 15 T 左右，介于 Nb_3Sn 和 NbTi 之间；二是低的不可逆性磁场，液氦下仅有 7 T，明显低于 NbTi 的 10 T 和 Nb_3Sn 的 20 T。MgB_2 的特性参

数与 Nb_3Sn 和 $YBa_2Cu_3O_7$ 的比较见表 1.3-62。由于新发现的 MgB_2 超导体的 T_c 具有远高于低温超导体的 T_c，又不存在高温超导体中难于克服的弱连接问题，又不需要高度织构，且 MgB_2 材料容易加工和成材，因此 MgB_2 材料无论对基础研究还是应用研究都具有十分重要的意义，它预示了磁体和电子应用方面的新一类低成本高性能超导材料的可能性。其临界温度接近 Nb_3Sn 的 2 倍和 NbTi 的 4 倍，使 MgB_2 有可能达到比用这两种超导体更高的磁场和电流密度，并有可能做 25 K 运行的 1.5TMRI 磁体，或指望替代高温超导材料的某些应用，虽然其载流能力目前还远低于 HTS 材料，但将来还有可能提高，其应用前景是诱人的。

表 1.3-62 MgB_2 的特性参数与 Nb_3Sn 和 $YBa_2Cu_3O_7$ 的比较

正常和超导态特性参数	MgB_2	Nb_3Sn	$YBa_2Cu_3O_7$
摩尔质量 $M/g\cdot mol^{-1}$	15.31	99.4	51.5
摩尔体积 $V/cm^3\cdot mol^{-1}$	5.83	11.1	8.0
平均临界温度 T_c/K	36.7~39	18	90
Sommerfeld 常数 $\gamma/mJ\cdot K^{-2}\cdot mol^{-1}$	0.89	13	约 1.5
电子-声子耦合常数			
λ_{ep}（从 $T_c\omega_{in}^{-1}$）	(1.07)	1.8	
电子-声子耦合常数			
λ_{ep}（从 γ/N (0)）	0.58	1.8	约 1.5
平均点阵力常数			
$M\overline{\omega_2^2}/eV\cdot A^{-2}$	10.9	5.3	
电子 Hopfield 参数			
N (0) $\langle I^2\rangle$ $/eV\cdot A^{-2}$	6.3	9.5	
T_c 处归一化比热容跳跃			
$\Delta C/\gamma T_c$	0.82	2.5	2
$T=0$ 处热力学临界场			
$\mu_0 H_c$ (0) /T	0.26	0.52	1.0
凝聚能测量值 $E_c/mJ\cdot cm^{-3}$	27	108	约 400
BCS 凝聚能			
$0.236\gamma T_c^2/mJ\cdot cm^{-3}$	49	90	约 360
超导能隙 2Δ (0) $/k_B T_c$	1.2~4.2	4.8	5
T_c 处上临界场斜率			
$-\mu_0(dH_{c2}/dT)T_c/T\cdot K^{-1}$	0.56	1.6	2.3
$T=0$ 处上临界场 $\mu_0 H_{c2}(0)/T$	14~17	25	150
$T=0$ 处相干长度 $\xi(0)/nm$	4.9~5.2	11.5	1.5
$\kappa=\lambda/\xi$	26~38	3.4	约 100
$T=0$ 处伦敦穿透深度 $\lambda(0)/nm$	140~185	39	约 150
$T=0$ 处下临界场 $\mu_0 H_{c1}(0)/T$	0.018	0.13	约 0.03

11.5 超导材料的应用

超导材料的应用可以分为大规模电力（强电）应用和小规模电子（弱电）应用两大类。从应用领域来看，超导材料可广泛应用于能源、交通、医疗、电子通信、科学仪器、机械工程以及国防工业等。超导材料的应用见表 1.3-63。实用低温超导材料的特性见表 1.3-64。不同高温超导材料对不同应用的适应和预测见表 1.3-65，高温超导材料今后应用领域预测见图 1.3-23。

表 1.3-63 超导电性的应用

已有应用	电力应用	电子应用
医学生物学化学应用	同步发电机	生物应用
化学用核磁共振谱仪	同步电机	无损检测
医用 SQUID 器件	直流机器	地理应用
热核聚变用超导磁体	变压器	重力检测
粒子加速器用超导磁体	电力传输	单磁通量子
超导同步 X 射线源	故障限流器	约瑟夫森电压标准
超导磁分离	飞轮储能	信息处理
高频谐振腔	小型快动作 SMES 系统	被动器件逻辑处理
超导运输系统	超导磁性储能	模-数转换器
超导磁性轴承		电磁波接收
磁屏蔽		热检测系统和天线
		超导外差接收器

表 1.3-64 实用超导材料的特性

超导材料	T_c/K	$\mu_0 H_{c2}$ (4.2 K) /T	J_c (4.2 K) $/10^4\cdot A\cdot cm^{-2}$				特点
			1T	4T	10T	17T	
Nb	9.2	0.3~0.5	0				低损耗，延性，电子学用
Nb-Ti (Ti45%~48%)	10.2	12	100	40	2		高损耗，延性，已商品化
$Nb_{0.75}Zr_{0.25}$	11	9	16				高损耗，韧性好
Nb_3Sn (A15)	18	23	870	200	20	6	低损耗，有商品，脆
V_3Ga (A15)		22		6	6	5	低损耗，有商品，脆
Nb_3Ge (A15)	22	35	200				低损耗，无商品，脆

表 1.3-65 不同高温超导材料对不同应用的适应和预测

材料类型	磁体	电流引线	电缆	故障限流器	电机
Bi-2223 带材	①	②	②	③	③
Bi-2212 带材	②	③			③
Bi-2212 涂层	①	③			
Bi-2212 块材		①			
Re-123 块材	未来	②			①
Re-123 涂层	未来	未来	未来	②	未来
Tl，Hg 系涂层带材	未来		未来		

①已有很大进展；②很可能；③有希望。

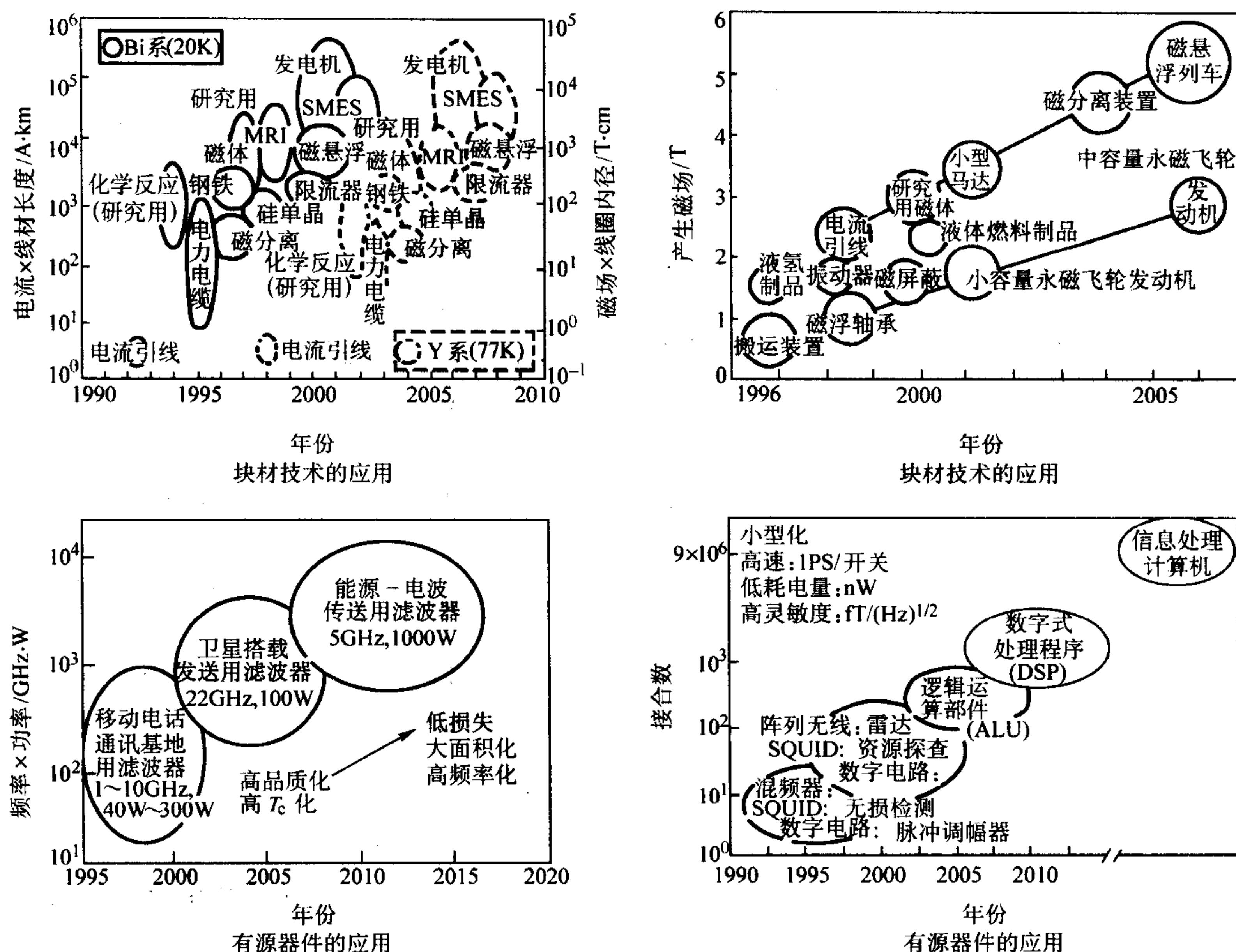


图 1.3-23　高温超导材料应用领域的分析预测

据第五届国际超导体工业顶峰会议（ISISS）估计，世界范围内 2020 年超导产业产值将达（1 500 ~ 2 000）亿美元，其中高温超导占 60% ~ 70%。高温超导在电力能源应用全球市场将达 435 亿美元。据专家估计，高温超导产品进入市场的年份分别为：故障限流器 2004 年，变压器 2005 年，电动机 2006 年，发电机 2011 年。而上述产品能占到市场 50% 的年份：变压器和电缆 2015 年，电动机 2016 年，发电机 2021 年，地下电缆 2013 年。

12　纳米材料

12.1　纳米材料的基本构成和特性

纳米科学技术是 20 世纪 90 年代兴起的高新科技，它的基本涵义是在纳米尺寸（0.1 ~ 100 nm）范围内认识和改造自然，通过直接操作和安排原子、分子来创新物质。

在纳米体系中，电子波函数的相关长度与体系的特征尺寸相当，这时电子不能被看成处在外场中运动的经典粒子，电子的波动性在输运过程中得到充分的展现：纳米体系在维度上的限制，也使得固体中的电子态，元激发和各种相互作用过程表现出与三维体系十分不同的性质，如量子化效应，非定域量子相干，量子涨落与混沌，多体关联效应和非线性效应等等。对这些新奇的物理特性的研究，使得人们必须重新认识和定义现有的物理理论和规律，这必将导致新概念的引入和新规律的建立，如纳米尺度上的能带、费米能级及逸出功将意味着什么？另外，在纳米化学中，对表面的化学过程，如原子簇化合物的研究对吸附质/载体系统的电子性质和对基底表面结构的影响；在纳米生物学中，除了对细胞、膜、蛋白和 DNA 的微观研究外，还要解决人工分子剪裁以及进行分子基因和物种再构；在纳米电子学中，电阻的概念已不是欧姆定律；在纳米力学中，机械性质如弹性模量、弹性系数、摩擦和粗糙概念亦有质的变化。作为纳米科技中的一个重要领域的纳米加工学，也将以崭新的方式进行原子的操纵和纳米尺度的加工以及进行纳米器件的加工和组装，并进一步研究器件的特性及运行机理。

纳米科技主要包括：①纳米物理学；②纳米化学；③纳米材料学；④纳米生物学；⑤纳米电子学；⑥纳米加工学；⑦纳米机械学等。纳米材料是纳米科学技术中富有活力而其研究内容十分丰富的科学分支。在纳米材料研究发展初期，纳米材料是指尺度为 1 ~ 100 nm 的纳米颗粒和由它们构成的纳米薄膜和块体。目前广义地讲，纳米材料是指在三维空间中至少有一维处在纳米尺度范围或由它们作为基本单元构成的材料。如果按维数，纳米材料的基本单元可分为几种。

1）零维　指在空间三维尺度均在纳米尺寸范围，如纳米颗粒、原子团簇等。

2）一维　指在空间有两维处于纳米尺度，如纳米丝、纳米棒、纳米管等。

3）二维　指在三维空间中有一维在纳米尺度，如超薄膜、多层膜、超晶格等。

构成纳米块体、薄膜、多层膜以及纳米结构的基本单元有：团簇、纳米微粒、人造原子、纳米管、纳米棒、纳米线和同轴纳米电缆等。它们的主要特征如下。

1）团簇　原子团簇是指几个至几百个原子聚集体（粒径小于或等于 1 nm），如 Fe_n、Cu_nS_m、C_nH_m（n 和 m 都是整数）和碳簇（C_{60}、C_{70}、富勒烯）等。原子团簇不同于有特定大小和形状的分子，也不同于以弱的结合力结合的松散分子团簇和周期性很强的晶体。原子团簇的形状可以是多种多样的，它们尚未形成规整的晶体，除了惰性气体外，它们是以化学键紧密结合的聚集体。原子团簇可分为一元原子团簇、二元原子团簇、多元原子团簇和原子簇和原子簇化合物。绝大多数原子团簇的结构不清楚，但已知有线状、层

状、管状、洋葱状、骨架状、球状等。原子团簇有许多奇异的特性，如极大的比表面使它具有化学活性和催化活性、光的量子尺寸效应和非线性效应、电导的几何尺寸效应、C_{60}掺杂及其导电性和超导性、碳管和碳葱的超导性等。当前能大量制备并分离的团簇是C_{60}及其他富勒烯。C_{60}的结构示意图见图 1.3-24。

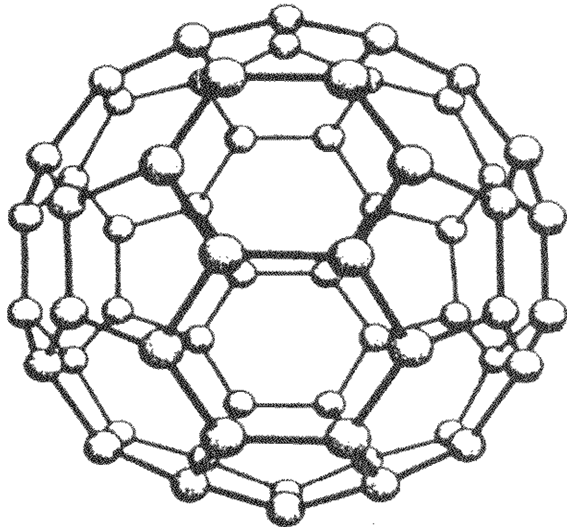


图 1.3-24　C_{60}分子结构

2）纳米微粒　纳米微粒是指颗粒尺寸为纳量级的超细微粒，一般尺寸在 1 ~ 100 nm 之间，大于原子簇、小于通常微粉。纳米微粒（1 ~ 100 nm）具有量子尺寸效应、小尺寸效应、表面效应和宏观量子隧道效应，因而展现出许多特有性质，如在催化、滤光、光吸收、医药、磁介质等方面有广阔应用前景。串成项链状的金纳米粒子见图 1.3-25。

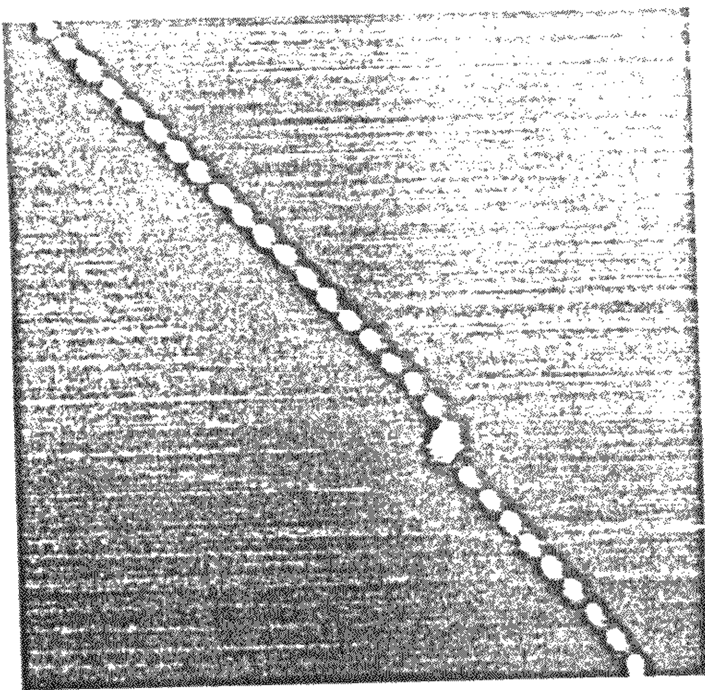


图 1.3-25　美丽的金纳米粒子项链

3）人造原子　也称量子点，是 20 世纪 90 年代提出的新概念。所谓人造原子是由一定数量的实际原子组成的聚集体，它们的尺寸小于 100 nm。人们曾把半导体的量子点称为人造原子。

4）纳米管、纳米棒、纳米丝和同轴纳米电缆　自从 1991 年日本 NEC 公司饭岛（Ijima）等发现碳纳米管以来，目前关于一维纳米材料（纳米管、纳米丝、纳米棒和同轴纳米电缆）的制备研究已有大量报道。因为一维或准一维纳米材料在介观领域和纳米器件研制方面有重要应用前景，如可用作扫描隧道显微镜（STM）的针尖、纳米器件和超大规模集成电路（uLSIC）中的连线、光导纤难、微电子、光电子方面的其他用途等。最小的（4.0）碳管垂直生长在（11，11）碳管上的图像见图 1.3-26。

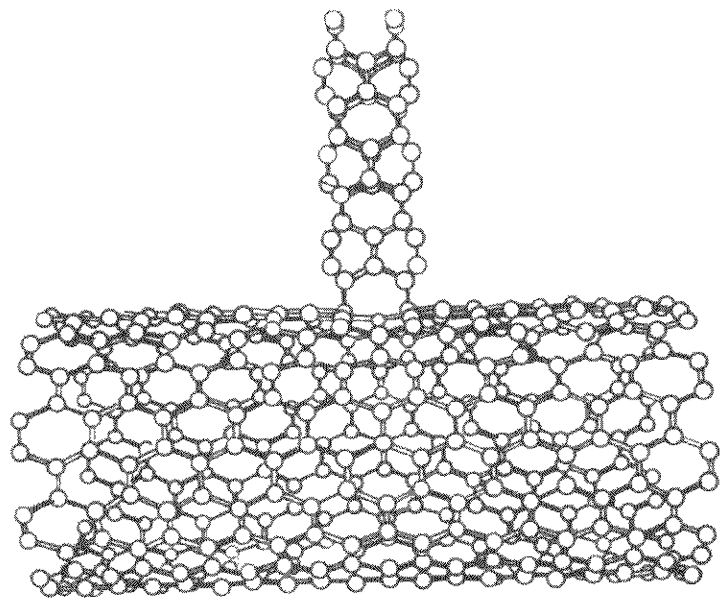


图 1.3-26　最小的（4.0）碳管垂直生长在（11，11）碳管上

纳米材料的基本物理效应和特性如下。

1）小尺寸效应　当超细微粒的尺寸与光波波长、德布罗意波长以及超导态的相干长度或透射深度等物理特征尺寸相当或更小时，晶体周期性的边界条件将被破坏；非晶态纳米微粒的颗粒表面层附近原子密度减小，导致声、光、电磁、热力学等物性呈现新的小尺寸效应。例如，光吸收显著增加并产生吸收峰的等离子共振频移；磁有序态向磁无序态转变；超导相向正常相的转变；声子谱发生改变等。

2）表面效应　纳米微粒尺寸小，表面能高，位于表面的原子占相当大的比例。随着粒径减小，表面原子数迅速增加。这是由于粒径小，表面积急剧变大所致。例如，粒径为 10 nm 时，比表面积为 90 m^2/g，粒径为 5 nm 时，比表面积为 180 m^2/g，粒径下降到 2 nm，比表面积猛增到 450 m^2/g。这样高的比表面，使处于表面的原子数越来越多，同时表面能迅速增加。

由于表面原子数增多，原子配位不足及高的表面能，使这些表面原子具有高的活性，极不稳定，很容易与其他原子结合。例如金属的纳米粒子在空气中会燃烧，无机的纳米粒子暴露在空气中会吸附气体，并与气体进行反应。

3）量子尺寸效应　当粒子尺寸下降到某一值时，金属费米能级附近的电子能级由准连续变为离散能级的现象以及纳米半导体微粒存在不连续的最高被占据分子轨道和最低未被占据的分子轨道能级而使能隙变宽现象均称为量子尺寸效应。

4）宏观量子隧道效应　微观粒子具有贯穿势垒的能力称为隧道效应。近年来，人们发现一些宏观量，例如微颗粒的磁化强度，量子相干器件中的磁通量等亦具有隧道效应，称为宏观的量子隧道效应。

宏观量子隧道效应的研究对基础研究及实用都有着重要意义。它限定了磁带、磁盘进行信息贮存的时间极限。量子尺寸效应、隧道效应将会是未来微电子器件的基础，或者它确立了现存微电子器件进一步微型化的极限。当微电子器件进一步细微化时，必须要考虑上述的量子效应。

12.2　纳米材料研究发展的主要方向和内容

1990 年 7 月，在美国巴尔的摩召开了国际首届纳米科学技术会议。在会上，各国科学家们对纳米科技（主要包括：纳米电子学、纳米机械学、纳米生物学和纳米材料学等）的前沿领域和发展趋势进行了讨论和展望，并决定出版三种杂志："纳米结构材料"、"纳米生物学"和"纳米技术"。1992 年 11 月在墨西哥召开了第一届国际纳米材料会议，后来分别在德国、美国、瑞典等国家召开了多次国际纳米材料会议，就纳米材料的研究方向和内容进行了广泛的交流和研讨，主要包括纳米金属材料、纳米陶瓷材料、有机 - 无机纳米复合材料、纳米传感材料、纳米电子材料、纳米生物医学材料和介孔材料等方面。

纵观纳米材料发展的历史，大致可以划分为三个阶段，

第一阶段（1990年以前）主要是在实验室探索用各种手段制备各种材料的纳米颗粒粉体，合成块体（包括薄膜），研究评估表征的方法，探索纳米材料不同于常规材料的特殊性能。对纳米颗粒和纳米块体材料结构的研究在20世纪80年代末期一度形成热潮。研究的对象一般局限在单一材料和单相材料，国际上通常把这类纳米材料称纳米晶或纳米相（nanocrystalline or nanophase）材料。第二阶段（1994年前）人们关注的热点是如何利用纳米材料已挖掘出来的奇特物理、化学和力学性能，设计纳米复合材料，通常采用纳米微粒与纳米微粒复合（0-0复合），纳米微粒与常规块体复合（0-3复合）及发展复合纳米薄膜（0-2复合），国际上通常把这类材料称为纳米复合材料。这一阶段纳米复合材料的合成及物性的探索一度成为纳米材料研究的主导方向。第三阶段（从1994年到现在）纳米组装体系、人工组装合成的纳米结构的材料体系越来越受到人们的关注（nanostructured assembling system）或者称为纳米尺度的图案材料（patternning mater on the nanometre scale）。它的基本内涵是以纳米颗粒以及它们组成的纳米丝、管为基本单元在一维、二维和三维空间组装排列成具有纳米结构的体系，其中包括纳米阵列体系、介孔组装体系、薄膜嵌镶体系。纳米颗粒、丝、管可以是有序地排列。如果说第一阶段和第二阶段的研究在某种程度上带有一定的随机性，那么这一阶段研究的特点更强调按人们的意愿设计、组装、创造新的体系，更有目的地使该体系具有人们所希望的特性。

纳米材料研究的内涵不断扩大，第一阶段主要集中在纳米颗粒（纳米晶、纳米相、纳米非晶等）以及由它们组成的薄膜与块体，到第三阶段纳米材料研究对象又涉及到纳米丝、纳米管、微孔和介孔材料（包括凝胶和气凝胶），例如气凝胶孔隙率高于90%，孔径大小为纳米级，这就导致孔隙间的材料实际上是纳米尺度的微粒或丝，这种纳米结构为嵌镶、组装纳米微粒提供一个三维空间。纳米管的出现，丰富了纳米材料研究的内涵，为合成组装纳米材料提供了新的机遇。在这方面举一个实例来加以说明，图1.3-27所示为金属纳米线阵列平板显示器结构示意图。

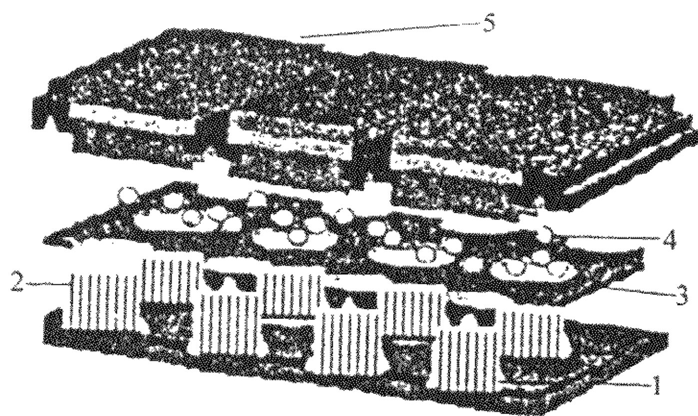


图1.3-27　金属纳米线阵列平板显示器结构示意图
1—金薄膜；2—金纳米线；3—栅极；
4—场发射到空间的电子；5—彩色荧光屏

在20世纪80年代中，曾有相当一些学者认为下一代的器件是分子电子器件，其理论基础是分子电子学，因此纷纷开展了分子电子学的研究。经过几年的工作逐渐认识到，在微电子器件与分子电子器件之间，可能有个过渡时期——纳电子器件，即信息加工的功能元件不是单个分子，而是原子团，它是有限个原子构成的纳米尺度的体系（含$10^2 \sim 10^9$个原子）。这样，人们就从两个方向去研究纳电子器件：以Si，GaAs等为主的无机材料的固体电子器件尺寸小下去，与基于化学有机高分子和生物学技术组装功能材料尺度大起来，两者的交叠区构成21世纪初期新型电子和光电子器件——跨世纪的信息功能器件的研究领域。这个构想被描述成纳米器件的发展路径：自上而下（top down）和自下而上（bottom up），将其画成纳电子器件发展路径示意图（map road），见图1.3-28。从真空管到晶体管再到电子盒的三代电子器件中，纳米电子器件主要是单电子器件，量子效应是其信号加工的主要特征，见表1.3-66。它将构成超高密度集成，是新型个人计算机、高性能计算机和自动器的基础，是未来智能工具的基本组件。因此三代电子器件的发展，具有变革性的进步，给人类社会带来的影响也具有质变的特征，而且是一代比一代对人类和社会产生更大的影响。因此，尽管在这方面还存在许多争议和不确定性，对纳米电子器件和其集成电路的诞生必须给予足够的重视，只有抓住机遇，迎接挑战，才能赶上一个新时代发展的潮流。

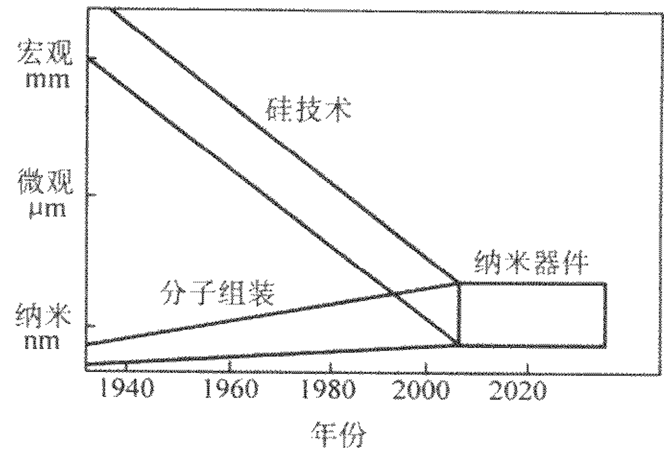


图1.3-28　纳电子器件发展的自上而下和自下而上路径

表1.3-66　三代电子器件的比较

比较项＼器件	真空管	晶体管	电子盒
放大器符号	g; c a	b; e c	g; s d
环境	真空	晶格	隧道隙
器件尺寸	cm	mm	mm
控制区尺寸	mm	μm	nm
自由程	m	0.1 μm	< 10 nm
材料	电真空材料	单晶半导体	有机/无机?
材料纯度	高纯（5~6个9）	半导体纯(8~9个9)	?
工艺	电子管制造工艺	光刻、掺杂技术	组装，自组织生长
理论	真空电子学（宏观）	半导体物理（宏观）	?（局域）
参量	平均值	平均值	非平均值
分布参量作用	很小	小	显著
温度敏感	很小	小	显著
载流子数量	宏观量	宏观量	有限量
信号强度	mA	μA	nA
载流子	电子	电子/空穴	孤子，极化子；电子/空穴
主要件控制	电流	电子/空穴数	隧穿速率
集成度	分立	$< 10^{10}$ bit/cm²	10^{12} bit/cm²
加工模式	模拟	bit	qubit
加工网	分立元件	数字网	类神经网
构造仪器	无线电、雷达	计算机	自动器、信息网

12.3　纳米材料的应用和发展

纳米技术是产业革命主导技术已成为全世界的共识。进入21世纪，各国都把发展纳米技术作为未来政治、经济格局中最富有挑战性的关键技术之一。2003年以来，发达国家对发展纳米技术的战略重新进行调整，特别重视纳米产业的发展，发展中国家也开始重视纳米技术。以市场为牵引，带动纳米技术的发展，已成为当前乃至未来相当长一个时期发展纳米产业的重要策略。德意志银行对2002～2006年纳米技术产品市场做了预测，其中纳米工具将由2002年的300亿美元增加到2006年的580亿美元，纳米器件由200亿美元增加到320亿美元，纳米材料将由250亿美元增加到300亿美元，纳米生物产品将由30亿美元增加到180亿美元，整个纳米技术产品市场到2010年将达2万亿美元。为了在未来市场中争得一席之地，各国都把振兴纳米技术、发展纳米产业提到了议事日程。美国把加快发展纳米技术作为新的战略部署，在2010年希望能占领1万亿美元的市场份额，日本、欧盟也把发展纳米产业放到振兴21世纪国民经济的重要位置，以利在未来市场竞争中获得更多的利益；亚洲的韩国、以色列、新加坡、印度、中国的台湾地区；澳洲的澳大利亚；非洲的南非都相应制定了发展纳米产业的计划，一个新的竞争格局已经形成。

纳米科技已在国际间形成研究开发的热潮，世界各国将发展纳米科技作为国家科技发展战略目标的一部分，纷纷投入巨资用于纳米科技的研究开发。纳米材料是纳米科技的重要组成部分，日益受到各国的重视。各国制定了相应的发展战略和计划，指导和推进纳米科技和纳米材料的发展，形成了政府、研究机构、大学和企业共同研究开发纳米材料和技术的局面。

美国自2000年2月提出“国家纳米技术倡议”（NNI），经国会批准于2000年10月1日执行实施计划以来，美国纳米科技研究开发经费逐年增加，从2001财年的4.22亿美元增至2003财年的7.102亿美元，2003财年分别比2001财年和2002财年增长53.11%和17.50%。国家纳米技术倡议实施计划确定了5个重点发展的战略领域（见表1.3-67）。

表 1.3-67　美国 NNI 在 2000～2001 财年的研究开发经费按领域的分配　亿美元

项　目	2000年	2001年
基础研究	0.87	1.77
重大挑战项目研究	0.71	1.33
杰出中心和研究网络	0.47	0.77
基础研究设施	0.5	0.80
伦理、法律和社会影响以及教育培训	0.15	0.28
合计	2.7	4.95

日本政府早在20世纪80年代初就实施了与纳米技术有关的若干计划。日本政府在第二个“科学技术基本计划”（2001～2006年）中，将纳米技术和材料与生命科学、信息通信、环境保护等作为国家科技重点发展战略的重中之重领域。该计划在2001年投入纳米科技的研究经费达142亿日元，比2000年度增加了88亿日元。其确定的纳米技术与材料重点领域主要包括：纳米物质与材料及其在电子、电磁、光学上的应用；纳米物质与材料及其在结构材料中的应用；纳米信息组件；纳米科技在医疗、生命科学、能源科学及环境科学方面的应用；有关表面和接口控制的物质及材料；纳米计量和标准技术；纳米加工、合成和工程技术；纳米技术的计算、理论和模拟技术；形成安全空间的材料技术等。通产省2001年制定了“纳米材料计划”（NMP），经费每年3 500万美元，为期7年（2001～2007年），由政府部门、政府研究机构、大学和产业界联合研究，旨在为产业界建立集研究开发新的纳米功能材料和教育功能于一体的纳米技术材料研究开发平台。其研究开发重点领域包括：金属、陶瓷、玻璃和聚合物方面的纳米材料；研究纳米氧化物、半导体和金属的单个纳米粒子；反映不同材料接口的纳米涂层；探索研究新型的功能材料，以及与纳米有关的物理表征方面的材料计量研究。通产省2001还制定并实施了“下一代半导体技术开发计划”，开发50～70 nm的下一代半导体处理基础技术，政府每年投资6 000万美元。

欧洲共同体一方面积极创建欧洲新的纳米技术产业，另一方面，力促现有产业部门提高纳米技术能力。欧洲共同体在第6个框架计划（2002～2006年）中，将纳米技术和纳米科学作为7个重点发展的战略领域之一，经费为12亿美元。其确定的重点研究领域。1）纳米技术和纳米科学：了解新现象、掌握新工艺和开发新的研究工具。纳米生物技术，侧重“纳米芯片”的研究，制造材料和部件的纳米工程技术。开发分辨率为0.01 nm的新一代纳米测量和分析仪器。纳米技术在卫生、化学、能源、光学和环境中的应用，侧重研究计算模拟，先进的生产技术，开发能改性的创新材料。2）多功能智能材料：设计和开发已定特性的新结构材料；开发超分子和微观分子工程，重点是新型的高复杂性分子及其复合物的合成、探索和潜在的应用；自修复工程材料；优化材料设计、加工、工具和试验，生产能构造更大结构的新型多功能“智能”材料，考虑生物兼容性与经济效益。3）新型的生产工艺和器件：新生产的概念包含更灵活、集成度更高、更安全和更清洁，这将依赖组织的突破和技术的发展等。研究开发具有创新性、可靠性、智能和有效成本的制造工艺及系统。开发用于清洁和安全生产的新器件和系统。

韩国政府在2002～2006年科学技术发展基本计划中将纳米技术与生物技术、信息技术和航空航天技术等作为国家科技发展的重点战略领域。还制定了一系列发展纳米科技的专项计划。2000年制定的“纳米生物技术发展10年计划”，重点研究开发纳米诊断器件、纳米治疗系统和纳米生物仿生器件（nanobiomimetics）。“2001～2010年太比特纳米器件计划”确定了太比特纳米电子学、自旋电子学、分子电子学和核心技术为研究重点领域。政府投资该计划的经费总计为1.42亿美元，科学技术部还积极鼓励私营企业设立纳米技术专项资金作为匹配经费。“2002年度纳米技术开发行动计划”，预算为2 031亿韩元，比2001年的1 052亿韩元增加了93.1%。旨在开发纳米核心技术，新建国家纳米制造研究中心（250亿韩元），以及信息技术和纳米技术融合中心。到2010年，将使韩国拥有13 000名纳米技术领域的专家并跻身纳米技术领域世界10强之列。

中国台湾顺应国际纳米材料研究的快速发展，1999年制定了“纳米材料尖端研究计划”，“纳米科技研究计划”（2001～2005年）5年预计投入的经费每年达上亿元新台币。台湾计划从2002～2007年，在纳米技术相关领域中投资总额为6亿美元的预算，每年稳中有增，平均每年达1亿美元。

中国通过“国家攻关计划”、“863计划”、“973计划”的实施，纳米材料和纳米技术已取得较为突出的成果，并引起了国际上的关注。例如，在纳电子方面，成功地研制出波导型单电子器件晶体管和对电荷超敏感的库仑计；实现6 nm宽的半导体量子线台面和6 nm宽的线条金属栅，制备出间隔仅为10 nm的多种“纳米电极对”；用GMR效应进行高灵敏度传感器和硬盘磁头原型的研制工作。在纳米器件的构筑

与自组装、超高密度信息存储、纳米分子电子器件等方面也取得了许多有意义和有影响的成果。在材料方面，以纳米碳管为代表的准一维纳米材料及其数组体系、在非水热合成纳米材料方面处于国际领先地位；纳米铜金属的超延展性、块体金属合金、纳米复相陶瓷、巨磁电阻、磁热效应、介孔组装体系的光学特性、纳米生物骨修复材料、二元协同纳米接口材料等领域都处于国际先进水平。在纳米复合材料改造传统材料和产品方面，部分成果已经实现产业化。

中国在纳米技术世界专利申请件数增加较快。据英国Thomson Derwent专利资料，对日本、美国、欧洲、亚洲等世界上约40个主要国家提出申请的专利统计得出，2000~2002年各国纳米专利申请件数，美国居第一位占32%，第2位为日本占21%，中国排名世界第3，占12%。德国（11%）和韩国（8%）分别居第4和第5位。中国自1997~1998年期间发明数量逐渐增加，1999~2000年增长速度加快。

为配合成果产业化，国内目前已建成100多条纳米材料生产线。金属纳米材料的生产能力已达到3~5 t/a；氧化物系列如SiO_2、TiO_2等，生产能力已超过300 t/a，纳米$CaCO_3$生产能力超过1000 t/a，品种包括铁、镍、钴、铜、铝、锌等金属纳米粉，TiO_2、SiO_2、ZrO_2、Al_2O_3、CaO、Fe_2O_3、Sb_2O_3等氧化物粉和SiC、Si_3N_4等陶瓷粉约有30余种，产品质量大都达到国际水平和接近国际水平。这些产品不但满足了国内研究与应用市场的需要，而且有相当部分出口。国家火炬计划支持生产纳米材料的骨干企业有：浙江舟山市陀升光纳米材料开发公司、江苏河海疏浚工程集团公司、长春迪瑞检验制品公司和山东正元纳米材料有限公司等。

我国的纳米材料企业在2000年大约有100家左右，注入资金7.8亿元人民币；到2001年发展到330多家，总投资60多亿元；2002年增至500多家，投资共200多亿元；到2003年底，猛增至800多家，总投资400多亿元；从事纳米材料和技术的公司数量和总投资金额分别见图1.3-29a和图1.3-29b。

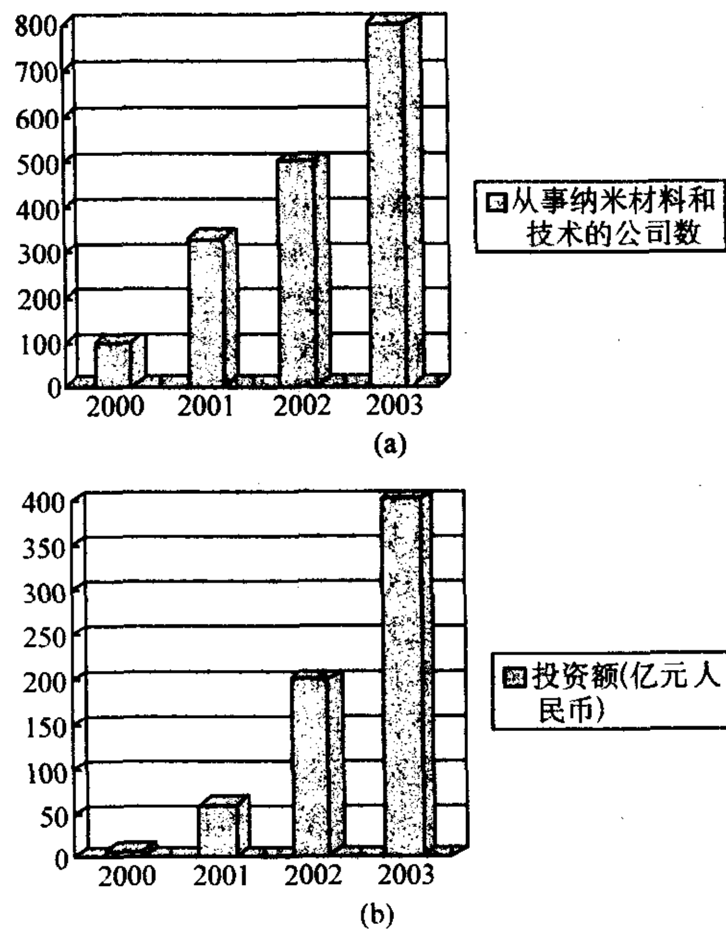


图1.3-29 中国纳米材料企业

以上公司中从事纳米材料制备的约占15%，纳米材料技术应用的约占85%。2002年以来，纳米粉体制备公司呈下降趋势，而纳米材料技术应用的公司大幅度增长，在近700家纳米材料应用技术公司中，应用纳米材料技术来提升传统产业的约占72%、纳米环境保护和水处理方面的产业约占6%、纳米药物产业约占6%、纳米能源产业约占4%、纳米电子产业约占3%、纳米制造业约占3%、交通和农业各占1%、其他行业占5%。在应用纳米材料技术改造传统产业方面效果较好的有精细化工、纺织、建筑、轻工、电力和材料工业。为了便于查找参考，表1.3-68~表1.3-71列出了到2004年止中国纳米材料主要企业。

表1.3-68 纳米粉体材料企业

材料及产品	企业（生产厂家）
纳米钛粉	哈尔滨鑫科纳米科技发展有限公司
金属纳米粉体材料	四平高斯达纳米材料设备有限公司
金属纳米粉体材料	大连星海纳米技术开发应用有限公司
金属纳米粉体材料	许昌银星纳米科技有限公司
金属纳米粉体材料	辽宁盘锦长江龙超微技术发展有限公司
金属纳米粉体材料	金昌凌云纳米材料有限责任公司
金属纳米粉体材料	焦作伴侣纳米材料工程有限公司
各种金属纳米粉	河南焦作纳米材料有限公司
纳米氧化钛	沈阳天润化工有限公司
纳米氧化钛及其他氧化物纳米材料	江苏泰兴纳米材料厂
纳米氧化硅	浙江舟山明日纳米材料有限公司
纳米金属和氧化物	深圳尊业纳米材料有限公司
纳米氧化锌	山西丰海纳米材料有限公司
纳米氧化锆	天津海赛纳米材料有限公司
纳米氧化铁	河北茂源化工有限公司
纳米钛酸盐	陕西泰和纳米材料有限公司
纳米氮化物	安徽合肥开尔纳米材料有限公司
纳米氮化物	齐齐哈尔中超纳米材料有限公司
纳米碳化物	河北华泰纳米有限公司
多种纳米材料	江苏海泰纳米材料有限公司
聚合物纳米复合材料	北京联科纳米材料有限公司

表1.3-69 纳米改性的纺织品及医疗保健企业

材料及产品	企业（生产厂家）
纳米保健纺织品	北京高飞龙有限公司
超双亲纳米材料	北京中商世纪纳米技术有限公司
纳米矿物保健材料	地大纳米技术开发有限公司
纳米抗菌剂	山西榆次开发区创先技术发展有限公司
纳米抗菌剂	山东正元纳米材料工程有限公司
纳米系列抗菌保健品	广州市汇纳科技发展有限公司
纳米无机抗菌剂	浙江丽水金地亚纳米材料有限公司

表1.3-70 纳米复合材料企业

材料及产品	企业（生产厂家）
磁性微粉	马鞍山金科粉体工程有限公司
阻燃增强纳米PET	北京联科纳米材料有限公司
ITO材料	银川
氧化锌压敏电阻	苏州苏净集团

表 1.3-71　特种纳米碳材料企业

材料及产品	企业（生产厂家）
纳米金刚石薄膜	深圳市雷地科技实业有限公司
纳米碳管	深圳纳米港有限公司
纳米碳管	山西路安纳米碳材料有限公司

我国纳米材料的研究、开发应用与国外的差距主要是：资金投入甚少，应用研究滞后于材料技术的发展，原始创新不够，缺少自己的知识产权，开拓国际市场能力差，少量品种还需进口。

在过去的十多年里，不论在全世界或中国，纳米材料技术方面的研究开发和应用已经取得了显著进展，但这确实仅仅是一个开始，无论在基础研究、产品开发和产业化方面还有很长的道路要走。我们必须加强规划、努力奋斗，要在这新的一轮科技和产业革命中占有一席之地。

13　生物医用材料

13.1　生物医用材料的基本特性

生物医用材料（biomedical materials），又称生物材料（biomaterials），用以和生物系统接合，以诊断、治疗或替换机体的组织、器官或增进其功能。它可以是天然的，也可以是人工合成的，或者是它们的结合。生物医用材料不同于药物，它的主要治疗目的不是必须通过新陈代谢或免疫作用等来实现，但可以结合药理作用，甚至起药理活性物质的作用。

与生物系统直接接合是生物医用材料最基本的特征，如直接进入体内的植入材料，以及体外辅助循环装置中人工心肺、肝、肾等与血液直接接触的材料，因此，生物医用材料除应满足一定的理化性质要求外，还必须满足生物学性能要求，即生物相容性要求。这是它区别于其他功能材料的最重要的特征。也正因如此，其使用涉及患者的生命安全，必须按国际或国家规定或认可的标准，严格地进行安全性、可靠性等评价并经国家主管部门认可之后，才能投放市场。

直到 20 世纪 80 年代中期，生物医用材料仍被定义为无生命的材料，随着科学技术的进步，特别是与宿主具有同源性的生物组织衍生材料的成功使用和含有活体细胞或天然组织的杂化人工组织及器官的发展，生物医用材料的定义已不仅仅限定为无生命的材料。

生物医用材料是构成现代医学基础的生物医学工程和生物技术的基础。生物医学工程产业通常指医疗器械产业，包括医用电子仪器、外科器械、治疗和手术消耗品、诊断器械、试剂、生物材料和医用植入体、体外循环装置（通常的人工器官）、药物控释系统等，管理上划属医疗器械范畴，归国家食品药品监督管理局器械司管理，美国则由 FDA（食品药品管理局）管理。

生物医用材料是医疗器械产业的基础，包括一般性材料和高技术材料。一般性材料指制造外科手术器具的材料、绷带、医用棉、一般的输液管等；高技术生物医用材料，即学术界常指的生物医用材料，是直接植入人体或与生理系统直接接触使用的材料，其研究必须与应用目标明确的植入体或装置结合，应用目标不同，同一种材料的制备工艺往往也不同。因此，生物医用材料的产值或市场通常连同其最终制品一起统计。高技术生物医用材料又称生物相容性材料及其终端产品，包括医用高分子、医用金属、生物陶瓷、复合材料、天然材料及医用植入体、人造血管、人工心瓣膜、起搏器、骨修复和替换器械、人工器官、牙科材料、药物控释系统等。仅从材料自身计，2002 年其产值为 10 亿美元，包括终端产品其市场达 500 亿美元，近十年来生物医用材料及其终端产品的世界市场规模发展情况见图 1.3-30。生物医用材料是为人类的健康服务的，涉及到人的生命安全问题，大量的高新技术注入使得生物医用材料具有很高的技术经济附加值，例如建筑材料约为 0.1 ~ 1.2 美元/kg，宇航材料则高达 100 ~ 12 000 美元/kg，而生物医用材料更高达 1 200 ~ 150 000 美元/kg，详见图 1.3-31。

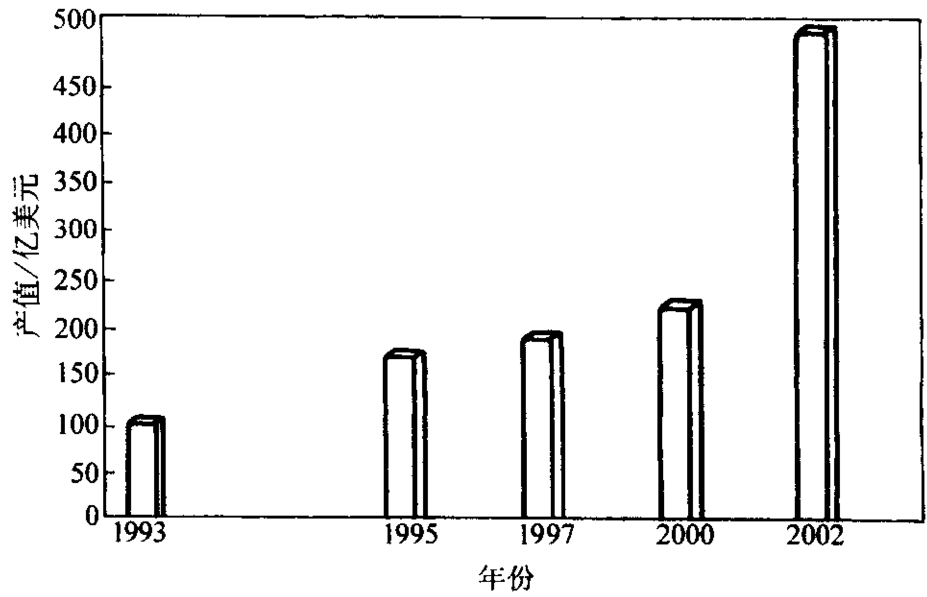


图 1.3-30　高技术生物医用材料及其终端产品世界市场发展情况

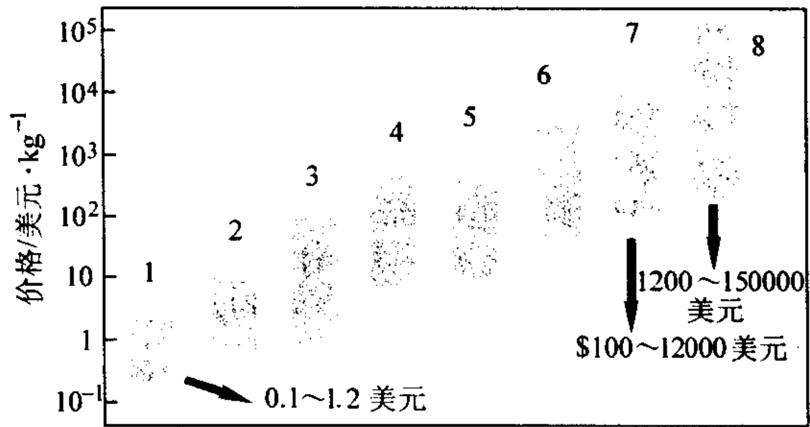


图 1.3-31　高技术生物医用材料与其他材料的价格对比
1—建筑材料；2—包装材料；3—海运材料；
4—汽车材料；5—设备仪表材料；6—体育用品材料；
7—宇航材料；8—生物医用材料

13.2　生物医用材料的分类

生物医用材料有多种分类方法，可以按材料的组成和性质、材料的用途以及材料引起的生物学反应等来分类。

1）按材料组成和性质分类　按材料组成和性质可分为以下几类。

① 医用金属和合金。主要包括三大类：医用不锈钢（316L，317L 等）；医用钴基合金（CoCrMo 铸造合金；CoCrMo、CoCrWNi、CoNiCrMoWFe、MP35N 等锻造合金）；医用钛和钛合金（CP - 纯钛，Ti - 6Al - 4V 合金等）。此外 Ni - Ti 记忆合金亦日益广泛地用于临床。

② 医用高分子。主要包括：医用聚乙烯、聚氨酯、聚硅氧烷、聚丙烯酸酯、聚碳酸酯以及可生物降解高分子（包括合成高分子：聚乳酸、聚乙醇酸、聚乙内酯、聚羟丁酸酯、聚酸酐等及天然高分子：胶原蛋白、纤维蛋白、甲壳素、壳聚糖、纤维素衍生物等）。

③ 生物陶瓷。主要包括强度高、耐磨性好的氧化铝、氧化锆等陶瓷；生物活性玻璃陶瓷；磷酸钙陶瓷以及碳素材料等。

④ 生物衍生材料。经特殊处理的生物组织所形成的一类材料，包括维持组织原有构型，仅消除其免疫排斥反应的较轻微处理过的组织，如经戊二醛处理定型的猪心瓣膜、牛

心包、人颈动脉、脐动脉、冻干的骨片等，以及经拆散原有构型处理的再生胶原、壳聚糖、透明质酸等的粉体、纤维膜、海绵体、凝胶等。

⑤ 生物医用复合材料。由两种或两种以上材料组合而成的生物医用材料，如涂层材料、羟基磷灰石增强聚乙烯、碳纤维增强聚乳酸复合材料等。

上述各类材料中使用量最大的是医用高分子和医用金属材料，两者各占生物医用材料用量的40%以上。

2）按用途分类　按用途分类可分为以下几类。

① 心血管系统植入材料和器械。包括人工心瓣膜、人造血管、血管内支架等。

② 矫形外科植入体。包括人工关节、人工骨、骨固定材料等。

③ 整形和重建外科材料和器械。包括人工乳房、人工皮肤、整形材料、阴茎植入物等。

④ 体外循环辅助装置。包括血液净化、分离材料、气体选择性透过膜等。

⑤ 介入器械。包括各种导管、导丝、封堵器、阻塞剂、缝线和组织黏合剂等。

⑥ 五官修复材料和器械。包括人工晶体、接触镜、人工眼台、喉管支架、耳听骨等。

⑦ 牙科材料。包括牙种植体、复合树脂、义齿等。

⑧ 神经系统修复材料和器械。包括神经导管、脑积水分流器、蜗状植入器等。

⑨ 计划生育材料和器械。包括宫内节育器、避孕套等。

⑩ 药物控释系统。包括药物和生物活性物质控释载体、半渗透聚合物膜等。

⑪ 临床诊断材料及生物传感器。

⑫ 一次性灭菌产品。如一次性注射器、输液器等。

3）按材料在生理环境中的生物化学反应水平分类可包括以下几类。

① 近于惰性的生物材料。在生理环境中能保持稳定，或长期暴露于生理环境中仅发生轻微、甚至不发生化学变化的材料。如合成的非降解高分子材料，医用金属和合金，氧化铝和医用碳素等近于惰性的生物陶瓷以及近于惰性的生物医用复合材料等。

② 生物活性材料。可诱发或调节特殊生物活性的材料。

③可生物降解和吸收的生物材料。在生物环境中可全部或部分被降解和吸收的材料。如可生物降解高分子，磷酸三钙等可生物降解陶瓷。

上述材料与组织间的结合可分为三种类型。

① 形态结合（morphologic fixation）。材料或植入体与组织间的机械嵌连。要求材料与组织间紧密配合，多发生于生物惰性材料与组织间的结合。

② 生物结合（biological fixation）。组织长入材料或植入体孔隙而实现的材料与组织间的结合。为避免材料与组织间发生相对位移切断长入孔隙的组织，应避免材料-组织间过大的切应力。

③ 生物活性结合（bioactive fixation）。材料表面在生理环境中发生选择性生物化学反应和组织间形成的一种化学键性结合。主要发生于表面生物活性材料，如致密的生物活性陶瓷和组织之间，这种键合非常之强，常常高于自然组织和材料自身的结合强度，以致断裂常发生于组织或材料之内，而很少发生于界面。

在材料和植入体设计之中，几种结合类型常被结合使用。

13.3 生物医用材料的评价试验和应用发展

国际标准化组织对医疗器械和生物医用材料的生物学评价试验作出了具体的规定，如ISO 10993-1（1997），详见表1.3-72。

现代医学正在向再生和重建被损坏的人体组织和器官、

表1.3-72　医疗器械生物学评价试验指南（ISO 10993-1:1997）

器械分类			基本评价的生物学试验								补充评价的生物学评价试验			
接触部位		A：一时接触（≤24 h） B：短、中期接触（>24 h至30 d） C：长期接触（>30 d）	细胞毒性	致敏	刺激或皮内反应	全身急性毒性	亚慢性亚急性毒性	遗传毒性	植入试验	血液相容性	慢性毒性	致癌性	生殖与发育毒性	生物降解
表面接触	皮肤	A	X	X	X									
		B	X	X	X									
		C	X	X	X									
	黏膜	A	X	X	X									
		B	X	X	X									
		C	X	X	X		X	X						
	损伤表面	A	X	X	X									
		B	X	X	X									
		C	X	X	X		X	X						
由体外与体内接触	血路间接	A	X	X	X	X				X				
		B	X	X	X	X				X				
		C	X	X		X	X	X		X	X	X		
	组织/骨/牙	A	X	X	X									
		B	X	X				X	X					
		C	X	X				X	X			X		
	循环血液	A	X	X	X	X				X				
		B	X	X	X	X		X		X				
		C	X	X	X	X	X	X		X	X	X		

续表 1.3-72

器械分类			基本评价的生物学试验								补充评价的生物学评价试验			
接触部位		A：一时接触（≤24 h） B：短、中期接触（>24 h至30 d） C：长期接触（>30 d）	细胞毒性	致敏	刺激或皮内反应	全身急性毒性	亚慢性亚急性毒性	遗传毒性	植入试验	血液相容性	慢性毒性	致癌性	生殖与发育毒性	生物降解
体内植入	组织/骨	A	X	X	X									
		B	X	X				X	X					
		C	X	X				X	X		X	X		
	血液	A	X	X	X	X			X	X				
		B	X	X	X	X		X	X	X				
		C	X	X	X	X	X	X	X	X	X	X		

注：X 为需选择的项目，后面表同。

恢复和增进人体生理功能、个性化和微创伤治疗等方向发展。传统的无生命的医用金属。高分子、生物陶瓷等常规材料已不能满足医学发展的要求，生物医学材料科学与工程面临着新的挑战。由于材料科学与技术、细胞生物学和分子生物学的进展，深化了材料及植入体与机体相互作用的认识，加之医学进展和需求的驱动，当代生物医学材料的发展已进入一个崭新的阶段，并处于实现意义重大的突破的边沿——在不远的将来设计和制造整个人体器官。其核心是赋予材料生物结构和生物功能；其基础是通过材料科学与生物学的紧密融合，在细胞和分子水平上深化材料和宿主间相互作用的认识；关键是运用生物学原理设计和构建材料及植入体，使其能容纳活体细胞、细胞产物、细胞外基质，并运用干细胞、生长因子、组织工程等生物技术赋予其生物功能，以充分调动人体自我康复的能力。

生物医用材料的前沿研究集中于两个方面：①设计和制造可诱导组织或器官再生的材料或替代人体组织的活体“部件”，即含有活体细胞、细胞素或生长因子等的材料和植入体；②发展用于先天性基因缺陷、老年病、肿瘤等难治愈疾病的药物，特别是生物活性物质（疫苗、组织细胞、基因等）控释系统，其中控释载体和半渗透膜的设计和制备是关键。生物医用材料前沿研究和发展，为其应用和学科的发展开拓了新的空间，材料的优化和制备则是其进一步发展的前提和基础。目前研究的科学重点是：基于生物学原理、赋予材料和植入体生物结构和生物功能的设计；可靠地试验材料生物安全性和预测材料长期寿命的科学基础；先进的工艺和制造方法学。生物医用材料的前沿研究亦是传统材料技术更新和产品换代的方向，即应用生物学原理和生物技术改造和更新材料和植入体，近期则以临床需求的用于特定目的和部位的材料和制品为重点。

生物医用材料科学和工程前沿的发展，虽为永久修复人体病变或缺损组织或器官展示了美好的前景，但目前的研究大部分处于基础研究阶段。医用金属、高分子、生物陶瓷等常规材料，仍将至少是 21 世纪前半叶临床应用的生物医用材料的主体。但是，常规材料长期临床应用已暴露出不少的问题，如免疫性问题，凝血问题，磨损碎屑引起的人工关节失效以及试验和评价植入体长期寿命的模型等问题。究其原因，就生物医用材料自身而言，是其基础研究，特别是生物学基础薄弱，表现在目前大量使用的约 20 种生物医用材料中，没有一种是专为生物医用材料研究开发的。回顾和强化常规生物医用材料的基础，特别是生物学基础，根据生物学原理指导常规材料的改进和更新，是当代生物医用材料科学和工程发展的另一个重要方面。

世界医疗卫生费用 2000 年约为 2 万亿美元，其中医疗器械约占 10%左右。而在医疗器械类中，高技术生物材料及制品 2002 年的市场规模约为 500 亿美元。美国在其中约占 40%的份额，中国所占的份额仅为 3%左右，而中国人口却占世界人口的 20%左右。随着中国国民生产总值和人民生活水平的不断提高，中国巨大的生物医用材料及其制品潜在市场已引起世界的瞩目。事实上即使在现阶段，由于中国国产生物医用材料及其制品的技术含量低，90%左右的高技术生物医用材料及制品（如血管支架、心脏瓣膜、心脏起搏器、介入导管、生活相容性人工关节等）依靠进口，这从近 10 年来外商在华注册的生物医用材料产品及厂商数（见表 1.3-73）可以看出其增长速度十分明显，而且在中国的销售价格昂贵，沉重地增加了国家和人民的负担，面临丧失这一个国民经济新增长点的机会，所以迫切需要进一步加强生物医用材料及制品的重大基础研究、高技术研发以及产业化的培育工作，使我国在这一领域激烈的国际竞争中能占有一席之地。

表 1.3-73 外商在华注册生物医用材料产品及厂商数

年份	1995	1996	1997	2003
品种	60	500	1 500	2 192
厂商	30	250	580	995

14 生态环境材料

14.1 生态环境材料的基本概念

生态环境材料应是同时具有满意的使用性能和优良的环境协调性或者能够改善环境的材料。所谓环境协调性是指资源和能源消耗少，环境污染小和循环再利用率高。

长期以来，人们总是单纯地追求材料的高性能、高产量以及高附加值，破坏性地向地球索取，而往往忽略了材料在提取、制备、生产、使用直至废弃的全过程对地球的资源与环境造成的影响和后果。各种统计表明，从资源、能源消耗的比重和造成环境污染的根源分析，材料产业是造成能源短缺、资源过度消耗乃至枯竭和环境污染的主要责任者之一。正是在对传统的材料制备思想与方法深刻反思的基础上，20 世纪 90 年代初日本学者山本良一率先提出了生态环境材料（ecomaterials）的概念。认为生态环境材料应是将先进性、环境协调性和舒适性融为一体的新型材料。在此基础上，经过我国众多学者长时间地讨论，初步达成如下共识，即“生态环境材料应是同时具有满意的使用性能和优良的环境协调性或者能够改善环境的材料。所谓环境协调性是指资源和能源消耗少，环境污染小和循环再利用率高”。这里既包括按生

态环境材料的基本思想和设计原则开发的新材料，也包括对传统材料的生态化改造，即在材料环境协调性评价（life cycle assessment，LCA）的基础上，通过对材料制造工艺的不断调整和改造，逐渐实现传统材料的生态环境材料化。显然，生态环境材料是与原有的材料相比较而产生、相比较而发展的新型材料，其判断和认知往往是相对的、动态的和不断发展的。它是人类充分考虑材料在其整个生命周期中对生态环境影响的基础上，在生态设计（ecodesign）思想和原则的指导下，采用革新的低环境负荷工艺（或者说绿色制造工艺）开发出来的具有资源和环境意识的新材料。因此，生态环境材料符合人与自然和谐发展的基本要求，是人与自然协调发展的理性选择，也是材料产业可持续发展的必由之路。它不仅是源头治理或减轻环境污染的实体材料，而且应当是新时代材料研制与生产的发展方向。生态环境材料的研究与开发是多学科的复杂的系统工程，随着人们对生态环境和资源问题认识的不断深化，相关学科相互交叉、相互渗透，以及科学技术的不断发展，将会使材料及其制备过程出现革命性的变化，并以崭新的面貌出现在本世纪。

材料的生产、使用和废弃过程（即材料的寿命周期）涉及以下一系列过程：矿物的采掘、冶炼和合成、材料的制造、加工、使用、废弃等。在这些过程中由于消耗了资源、能源并向环境排放了废弃物，造成了资源的消耗和生态环境的损害。这种情况可以简单的概括为图1.3-32。

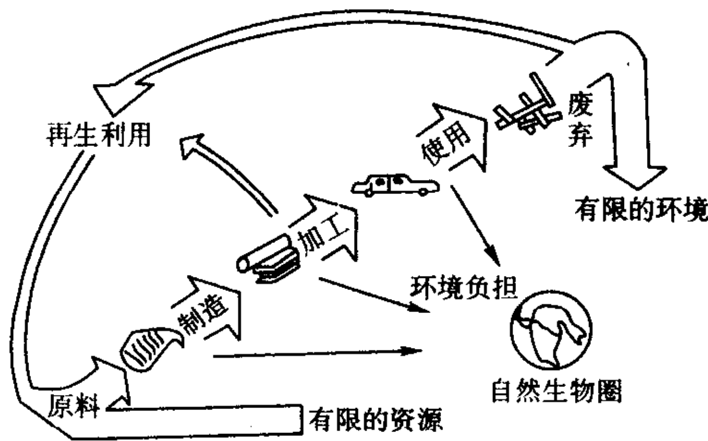


图1.3-32　材料的寿命周期与资源、环境的关系

对资源的消耗可以概括为以下几方面。

·能源的消耗。

·矿物资源的消耗。

·生物资源的消耗。

·水资源的消耗。

对生态环境的损害可以概括为如下几个方面。

·向大气环境排放的有害物质，如大气污染物的排出；具有温室效应气体的排出；破坏大气臭氧层物质的排出；酸性气体物质的排出等。

·向水域排放的有害物质，如危害人体和生物体健康物质的排出；影响生态环境物质的排出等。

·向土壤排放的有害物质，如污染土壤物质的排出；恶化土质物质的排出；危害生物链物质的排出等。

·其他损害环境物质的排放，如固体废弃物的排放；恶臭物质的排放；噪声、振动、电磁波的产生等。

由此可见人类正面临日益严峻的资源枯竭和环境污染问题，如图1.3-33所示。

某些重要金属的全球储藏量见表1.3-75，由表可见这些重要金属全球储量的消竭大部分均在数十年之内。中国主要矿产资源储量在世界上的位次见表1.3-74。由表可见，我国主要矿产资源总体上看还是比较丰富的，但若按人口平均的潜在价值比较，则中国并无优势可言，1990年时排序仅为第80位，如表1.3-76所示，而且随着人口基数的增加，人均资源呈下降趋势，所以中国资源人均值的国际地位还将有所下降。

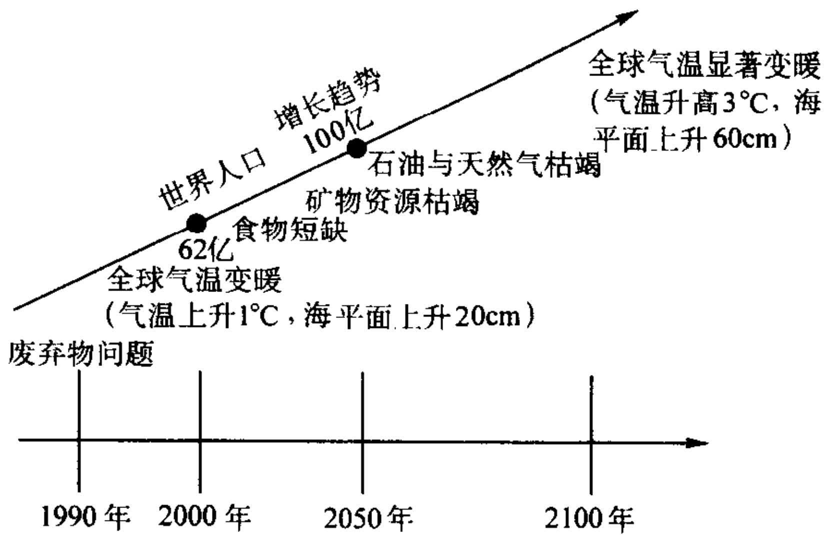


图1.3-33　人类面临的资源和环境问题

表1.3-74　中国主要矿产资源储量在世界上的位次

矿种	位次	储量较丰富的国家或地区
铁	5	俄罗斯、巴西、加拿大、澳大利亚、中国
锰	5	南非、俄罗斯、澳大利亚、加蓬、中国
铬	13	南非、津巴布韦、芬兰、俄罗斯等
铜	5	智利、美国、赞比亚、秘鲁、中国
钒	1	中国、南非、俄罗斯、澳大利亚、智利
钛	1	中国、印度、加拿大、挪威
铅	5	美国、澳大利亚、俄罗斯、加拿大、中国
锌	1	中国、加拿大、澳大利亚、美国、俄罗斯
镍	5	新喀里多尼亚、加拿大、印度尼西亚、俄罗斯、中国
钴	3	扎伊尔、赞比亚、中国、俄罗斯、古巴
钨	1	中国、加拿大、俄罗斯、美国、澳大利亚
锡	2	印度尼西亚、中国、泰国、澳大利亚
钼	2	美国、中国、智利、俄罗斯、加拿大
铝土矿	7	几内亚、澳大利亚、巴西、牙买加、印度、喀麦隆、中国
铋	1	中国
锑	1	中国、玻利维亚、南非、俄罗斯、墨西哥
汞	2	西班牙、中国、俄罗斯、南斯拉夫、英国
金	7	南非、俄罗斯、美国、澳大利亚、菲律宾、津巴布韦、中国
银	7	俄罗斯、加拿大、美国、墨西哥、澳大利亚、秘鲁、中国
铂族	6	南非、俄罗斯、美国、加拿大、津巴布韦、中国
锂	1	中国、智利、美国、加拿大、扎伊尔
铌	3	巴西、俄罗斯、中国、加拿大、尼日利亚
钽	2	扎伊尔、中国、澳大利亚、尼日利亚、俄罗斯
稀土	1	中国、美国、俄罗斯、印度、加拿大
菱镁矿	1	中国、俄罗斯、朝鲜、巴西、澳大利亚
萤石	1	中国、南非、墨西哥、英国、美国
磷	2	摩洛哥、中国、俄罗斯、南非、美国
硫铁矿	1	中国、加拿大、俄罗斯、美国、波兰
砷	1	中国、美国、加拿大、秘鲁
重晶石	1	中国、美国、印度、加拿大、墨西哥
滑石	2	美国、中国
石棉	2	美国、中国、南非
石墨	1	中国、美国
石膏	1	中国、美国、加拿大
叶腊石	3	日本、韩国、中国

表 1.3-75　某些重要金属的全球储藏量

金属	储藏量/10^6t	年消耗增长率/%	可开采时间/a
Fe	1	1.3	109
Al	1.17	5.1	36
Co	308	3.4	24
Zn	123	2.5	18
Mo	5.4	4.0	36
Ag	0.2	1.5	14
Cr	775	2.0	112
U	4.9	10.6	44
Ti	147	2.7	51

表 1.3-76　世界前10位国家的45种矿产资源储量的潜在价值比较（1990年）

国家及地区	矿产总价值		按人口平均	
	比例/%	位次	平均值/万美元·人$^{-1}$	位次
前苏联	17.73	1	5.06	8
美国	16.56	2	5.67	6
中国	14.64	3	1.19	80
沙特阿拉伯	6.59	4	62.58	1
英国	4.12	5	5.52	7
加拿大	3.96	6	12.58	3
伊朗	3.4	7	7.22	5
澳大利亚	3.35	8	17.57	2
南非	2.94	9	7.94	4
巴西	2.94	10	1.90	9

另一方面，从能源消耗和废弃物排放方面来看，我国材料行业的能耗远较先进工业国家为高，如表 1.3-77 所示；而废弃物排放量则是世界环境污染的大户，见表 1.3-78。

鉴于中国是基础材料（传统材料）的生产大国，2003 年我国的钢铁、有色金属、水泥、玻璃、合成纤维等的产量均居世界首位，所以关于环境材料的研究、发展与应用对我国具有特殊重要的战略意义和十分迫切的现实需求。

表 1.3-77　几种材料的能耗比较

材　料	能耗/（t标准煤/t产品）		比值
	中国	先进工业国家	
钢材	1.64	1.00	1.64
水泥	0.201	0.113	1.78
纸	1.20	0.59	2.04
玻璃	0.046（标准煤/箱）	0.02	2.30
电解铝	14 916 kW·h·t^{-1}	12 956	1.15

表 1.3-78　全国工业废弃物排放量

项　　目	1993 年	2000 年	增长率/%
工业固体废弃物数量/10^8t	6.17	8.2	32.9
工业废气排放量/$10^8 m^3$	109 604	—	—
燃料燃烧产生的废气/$10^8 m^3$	75 401	—	—
水泥厂排放的 CO_2/10^4t	24 500	—	—
工业 SO_2 排放量/10^4t	1 795	1 995.1	11.1
工业烟尘排放量/10^4t	1 416	1 165.4	-17.7
全国废水排放量/10^8t	355.6	415	16.7
工业废水排放量/10^8t	219.3	194	-11.5

14.2　生态环境材料研究发展的主要方向和内容

生态环境材料研究发展的主要方向和构成要素如下：

① 减少年人均材料流量，减少材料集约化程度；

② 减少寿命周期中的环境负荷，使用生态化的生产工艺；

③ 开发天然能源，使用藏量丰富的矿物和天然材料；

④ 避免使用有害物质，使用“清洁”材料；

⑤ 使用长寿命材料，强化再生利用，强化生物降解性；

⑥ 修复环境，强调生态效率（性能-环境负荷比）。

典型的环境协调性研究分类见表 1.3-79。

表 1.3-79　典型的材料环境协调性研究分类

设计原则	研究方向	改进材料体系实例	改进材料制备技术实例
减少资源和能源消耗	现场循环利用		废料循环利用
	再生循环利用	通用合金，通用树脂	金属中的杂质分离与无害化技术，塑料的再生循环技术
	减少枯竭性元素使用	低合金度合金	
	降低能耗		近净尺寸成型技术，软溶液制备技术，低能耗烧结技术
减少污染排放	降低污染程度	可降解塑料	熔融还原炼铁
	废弃物综合利用		废弃物资源化
	去除有毒有害元素	无铅钎焊合金，非铅基压电陶瓷	
提高材料使用价值	提高材料性能		超级金属
	延长材料使用寿命	长寿命金属材料，长寿命高分子材料	陶瓷涂层

在生态环境材料的研究和发展中，一个首要的问题是：什么样的材料才算得上是生态环境材料。这涉及如何评价材料的环境协调性，即其环境表现或环境性能，由此而产生了材料的环境协调性评价研究。目前通常采用生命周期评价（Life Cycle Assessment，LCA）的基本概念、原则和方法对材料或产品进行环境行为评估。1997 年国际标准化组织（ISO14040）对 LCA 的定义是这样一种评价方法，通过确定和量化与评估对象相关的能源消耗、物质消耗和废弃物排放，来评估某一产品、过程或事件的环境负荷；定量评价由于这些能源物质消耗和废弃物排放所造成的环境影响；辨别和评价改善环境的机会。评估过程应包括该产品、过程或事件的寿命全过程，包括原材料的提取与加工、制造、

运输和销售、使用、再使用、维持、循环回收，直到最终的废弃。根据ISO14040对LCA标准定义的技术框架，LCA应包括目的与范围的确定（goal and scope definition）、清单分析（inventory analysis）、影响评价（impact assessment）和生命周期解释（life cycle interpretation）4个组成部分，见图1.3-34。

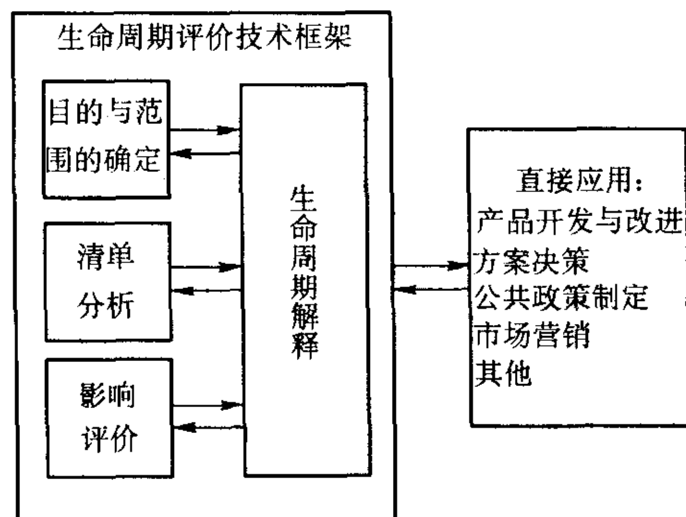


图1.3-34　LCA的评估过程及技术框架

14.3　生态环境材料的研发趋势及应用

1) LCA研究进展。从LCA提出到现在的10多年间，LCA的研究取得了巨大的进展。在LCA的理论研究中，由国际标准化组织制定的环境管理国际标准（ISO14000系列）中的LCA相关研究是最有影响的研究计划，体现了世界范围内LCA研究的共识。

其中有关LCA的基本原则较早就已经确立下来，近几年也没有大的改变（ISO14040），编目分析的标准也已基本成型（ISO14041），这两个标准都已经公布了国际标准草案DIS。近年对环境影响评估（impact assessment）的研究很多，提出了许多的方案与模型，也存在着许多争论和分歧，但基本上都包含了分类（classification）、表征（characterization）和评价（evaluation）这三个环节，而对评估解释（interpretation）还需进一步的研究。

目前LCA方法仍处于发展阶段，LCA的实施方案仍有待完善。在找到一个更实用、更有效、更普遍，并能为大家共同接受的LCA实施方案之前，LCA研究者将仍然以LCA标准和基本框架为指南，根据自己的理解和经验，结合具体产品系统的特点进行LCA的研究与应用。

除了对LCA理论的研究之外，以特定产品为对象的LCA应用研究也非常普遍，尤其是编目分析研究和相关的数据收集。但由于环境影响评估理论上的不足，总的来说，现阶段还没有普遍适用的完整的LCA实施标准，也没有一个能够被普遍接受的实施方案或指标体系。这方面还需要进一步地研究。

此外，还有大量的相关研究尽管没有使用LCA这样的名称，但也涉及产品的寿命周期概念和LCA评估方法，从广义上讲也应该被视为LCA研究的一个部分。

由于LCA评估中需要处理大量的数据，借助于计算机可以更好地完成LCA评估，近年来已经开发出数十个用于LCA、LCI的计算机软件，以及用于环境管理系统（EMS）的管理软件。其中最著名的例如SimaPro4.0、GaBi3.0和KCL-ECO3.0等。它们支持用户管理大量的数据，为产品系统建立模型，能够进行不同类型的计算，并帮助生成评估报告。但评估质量仍取决于用户。大多数LCA评估软件的功能都符合ISO14040对LCA方法的定义。

2) 生态设计（环境协调性设计，ecodesign）。ecodesign是LCA方法应用到工业产品设计过程中产生的新概念，其实质是在产品设计过程中加入对产品环境影响的考虑，而不是像传统的设计方法一样只考虑产品的质量、性能和成本等因素。从方法上看，ecodesign是将传统的产品设计方法与LCA方法相结合，从环境协调性的角度对产品设计提出指标和建议。

在设计阶段就考虑到产品的环境影响很重要，因为随着产品设计和生产过程的进行，防止污染的机会越来越小，而治理污染的代价越来越大。事实上，产品的环境污染大多数在产品设计时就已经决定了。所以ecodesign是从根本上解决环境问题的一个关键。

从汽车到家用电器，从建筑到生活日用品，ecodesign已经被广泛地应用到工业设计中，并得出了如下一些普遍的设计准则。

① 设计产品意味着设计一个对环境安全的产品寿命周期，所以设计者应该了解在产品的整个寿命周期中所消耗的资源和能源，以及对环境造成的负面影响。

② 对产品环境影响的理解应该建立在LCA分析的基础之上，才能得到安全可靠的结论。例如，通常人们认为使用天然材料比使用人工材料更有利于环境，但实际上这并不是绝对的。从LCA的观点来看，功能相同的产品系统才具有可比性。所以首先需要衡量完成同样功能的不同材料的使用量，还要考虑不同材料需要的预处理和废弃处理过程，这样才能得出完整和正确的结论。

③ 增加产品的使用寿命是减少环境污染的重要途径。除了将产品设计得更耐用之外，通过更换部件使产品升级的方法也能达到减少污染的目的。另外从调查中发现，有时消费者会因为厌倦而抛弃能够正常使用的产品，所以将产品设计得更富有吸引力也有利于延长产品的使用寿命。

④ 减少能源的使用对保护环境有直接的效果，因为使用能源不仅意味着资源的消耗，也意味着大量污染的产生。例如，产生10 kW·h的电能需燃烧2 kg的燃油，对于在产品寿命周期中会消耗大量电能的产品而言，通常大部分的环境负担都是由使用能源引起的。

⑤ 材料的选择对产品的环境负荷影响很大。在保证产品性能和质量的前提下，减少材料的用量不仅能降低产品成本，也有利于降低环境负荷。

⑥ 为了使产品的回收和循环利用成为可能，在产品设计时应做特殊的考虑。例如，产品结构应便于拆解，尽量使用单一及易回收的材料等。另外，在产品中主动使用回收材料也是一个非常重要的措施。

针对材料的生态设计的范例有。

金属材料中的通用合金、简单合金、无铅钎焊合金等；有机聚合物的可再生循环设计（聚合物合金）；有机复合材料的可再生循环设计（液晶聚合物）。

针对产品的生态设计的范例有产品的模块化设计（计算机、打印机、复印机等IT产品）；汽车的轻量化设计等。

进一步开发了计算机软件来进行计算机辅助产品可再生循环设计。

3) 生物降解塑料。目前，全世界完全生物降解塑料年产量为150万吨，是指非光降解型的，即能够在土壤或水中进行完全生物降解，对环境不产生任何污染的一类塑料，并非是淀粉填充性的半生物降解型的。国际上已有数家公司处于发展和商业化产品阶段，它们的产品包括：聚乙烯醇、丙烯酸共聚物、聚乳酸、微生物聚酯、热塑性淀粉、聚环氧乙烷共混物、聚己内酯。据报道，这些材料可以在传统设备上非常容易地进行加工，产品性能和现有塑料具有可比性，最终产品可在土壤或水中完全生物降解。

部分生物降解高分子的制备方法与特性见表1.3-80。降

解高分子所期望的用途见表 1.3-81。在 1989～1996 年期间国际上有关降解塑料所颁发的标准见表 1.3-82。

表 1.3-80 部分生物降解高分子的制备方法与特性

分类	生物降解高分子	制备方法	特性
微生物合成高分子	聚 3－羟基丁酸[P(3HB)]	以糖、植物油、乙醇等为原料，用微生物合成得到P(3HB)	熔点约 180℃，破坏强度 43 MPa，类似于 PP。但与断裂伸长率为 400%的 PP 相比，P(3HB)只有 6%，是硬而脆的高分子材料
		用遗传基因修饰的微生物合成超高分子量达 1 万以上的P(3HB)	用这种P(3HB)制成的薄膜，其破坏强度和断裂伸长率与 PA66 相当，将在高性能薄膜和纤维方面得到应用
化学合成高分子	聚乳酸(PLA)	以发酵的 L－乳酸为原料，聚合 PLA	熔点约 180℃，其性质与 PET 相似，即具有室温以上的玻璃化转变温度和透明性，能在纤维、薄膜和瓶类得到应用
	聚琥珀酸丁烯(PBS)	以 1，4－丁二醇与琥珀酸为原料，在催化剂和结合剂的作用下聚合 PBS	具有与 PE 相似的熔点和成型方法，可通过注射、挤出、中空吹塑、发泡等成型方法，制成包装材料。为了控制 PBS 的性能和生物降解性能，已在开发生产出己二酸共聚的生物降解塑料
	聚琥珀酸乙烯(PES)	以乙二醇与琥珀酸为原料，在催化剂和结合剂的作用下聚合 PES	
	聚 α－氨基丁二酸(PAA)	先将 α－氨基丁二酸缩聚制得聚琥珀酰亚胺，再通过加水分解、合成	是水溶性生物降解高分子。在水中是可溶性高分子电解质，具有类似与聚丙烯酸洗涤增效剂和废水处理分散剂的性能。由交联反应可形成吸水性溶胶，所以可作为尿布、卫生纸等的材料

表 1.3-81 降解高分子所期望的用途

领域	用途	
在自然环境中使用或流向自然环境可能性大的领域	农、林、水场材料	多用途薄膜、植物苗圃移植用灰料兜、钓鱼线、渔网、紫菜网等
	土木和建筑材料	隔热材料、荒地和沙漠绿化用保水膜、工事用保水片材、植物网等
	野外运动制品	高尔夫、钓鱼用鱼饵、跳绳等一次性制品
	水处理材料	沉淀剂、分散剂、洗净剂等
使用后回收困难的领域	食品或容器包装材料	食品包装膜、鲜新食品盒、一次性餐具等
	卫生用品	尿布、卫生纸等
	办公用品、衣料、杂货	笔盒、笔芯、理发梳子、垃圾袋、各类杯子、衣料等
要求特殊功能领域	逐渐释放性	药品、农药、肥料、种子等逐渐释放性包覆材料
	保水性、吸水性	荒地和沙漠绿化造林用材料
	身体内分解性	手术缝合线、骨折稳定材料、医用薄膜、医用不织布等
	低氧气透过性、非吸附性	食品用包装膜、饮料桶内部的涂料等
	低熔点	包装、出版物、制袋时所用的黏合剂

表 1.3-82 1989～1996 年间有关降解塑料的 ASTM 标准

标准编号	标准名称
ASTM G22—1987	测试合成高分子材料抵抗细菌的标准操作法
ASTM G21—1990	测试合成高分子材料抵抗真菌的标准方法
ASTM D3826—1991	采用拉伸试验测定可降解聚乙烯及聚丙烯降解终点的标准规则
ASTM D5071—1991	可光降解塑料曝晒用水氙灯弧型曝晒仪标准操作规则
ASTM D5152—1991	降解塑料残余固体物水萃出物的毒性试验标准
ASTM D5208—1991	可光降解塑料曝晒用（荧光）紫外线及冷凝仪标准操作规则
ASTM D5209—1991	城市污染淤泥中，测定可降解塑料需氧生物降解性的标准试验方法
ASTM D5210—1991	城市污染淤泥中，测定可降解塑料厌氧生物降解性的标准试验方法
ASTM D5247—1992	采用特定的测定可降解塑料需氧生物降解性的标准试验方法
ASTM D5272—1992	光降解塑料户外标准试验方法曝晒试验标准规则
ASTM D5338—1992	受控堆肥化条件下测定可降解塑料需氧生物降解的试验方法
ASTM D5437—1993	塑料在海洋漂浮曝晒状态下耐候试验标准规则
ASTM D5509—1996	塑料曝晒于模拟堆肥环境中的标准规则
ASTM D5512—1996	塑料曝晒于采用外加热器的模拟堆肥环境中的标准规则
ASTM D5951—1996	固体废弃物中的塑料可生物降解性试验方法
ASTM D6002—1996	环境降解塑料堆肥性评价的标准规则
ASTM D6003—1996	固体废弃物中的塑料经可生物降解试验方法进行毒性和堆肥质量试验后配置的剩余固体物的标准规则

4）环境替代材料。目前研究开发的环境替代材料有发泡塑料替代材料、氯氟烃化合物替代材料、石棉替代材料等。发泡塑料替代材料可用于制造各种一次性餐具，家电产品的包装材料和缓冲包装物，代替目前广泛使用的塑料及泡沫塑料包装材料及制品，较好地解决了石油资源短缺和“白

色污染”等问题。目前研究和开发的氯氟烃化合物替代材料有两大类，一类是过渡性替代材料，另一类是永久性替代材料。过渡性替代材料主要有氟代烃类化合物（HCFC）、丙烷、异丁烷等，永久性替代材料目前开发出来的有环戊烷，HFC_{134a}等。

目前开发的石棉替代材料已有好几种，如用含石墨和无机纤维的材料制作汽车刹车垫，用硅酸铝、硅酸锌陶瓷纤维制作隔热垫或其他保温绝热材料，用芳族聚酰胺纤维替代石棉纤维制作高温防护服等。

5）环境净化材料和环境修复材料。研究和开发的环境净化材料主要有大汽污染控制材料、水污染控制材料和其他污染控制材料等。环境修复材料可用于修复已破坏的环境。常见的环境修复材料有固沙植被材料、二氧化碳固化材料、臭氧层修复材料等。

编写：李成功（中国材料研究学会）

第 4 章　材料与机械制造

机械制造涉及的范围极其广泛，大到航空母舰、重型机械，小到微米机械（mems）。涉及的领域也很多，如轻纺制造业的食品加工、家具制造、资源加工业的石化产品、金属提炼与制造，都离不开机械。机械电子制造业，包括机床及锻压装备、交通运输设备、电子机械及电子通讯和仪器仪表等，更是机械制造业的主体。所以，机械制造业不但在国民经济中占有十分重要的位置，在国防建设中也无所不在。本章不包括微米机械和纳米机械，因为这些目前正在发展过程中，而且所涉及的设计原理和材料与我们通常所说的机械制造有很大区别。

机械制造有四个要素，那就是设计、材料、制造和检测。图 1.4-1 表示四者的关系。

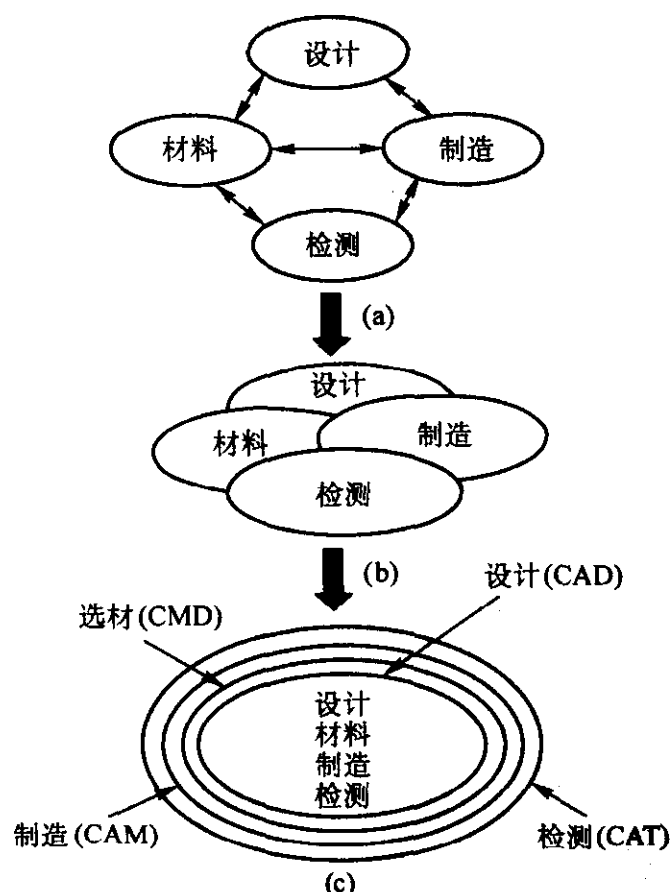


图 1.4-1　机械制造四要素及其关系的演化

CAD—computer aided design; CMD—computer materials data system; CAM—computer aided manufacturing; CAT—computer aided testing

在这四个要素中，设计是灵魂，因为设计决定机械产品的性能、品质与先进性；材料是基础，没有材料，机械产品就无从谈起，因此选材的正确与否，直接影响机械产品的性能、耐久性和成本；制造是关键，理想的设计、正确的选材，只有通过制造才能做出实用、有竞争力的机械；检测是保证，机械产品质量的控制决定于检测装置和手段的精确程度。

在正常情况下，任务确定后，先提出构思，而后设计，根据设计再选材，进行制造，最后经检测合格后出厂。20 世纪 80 年代，美国国防部提出并行设计工程（concurrent engineering），即通过合理管理，优化管理，有可能缩短产品的生产过程，从而提高劳动生产率、增强竞争能力，因而四个要素可搭接进行，如图 1.4-1b。最后，由于计算机的普遍采用，形成计算机集成制造系统（CIMS）（computer integrated manufacturing system），有可能实现如图 1.4-1c 的理想状态。

1　材料在机械制造中的重要性和地位

机械制造没有材料便造不出机械，由此可以看出材料在机械制造中的重要性，但不同类型的机械，材料所处的地位不同。现仅以人们最关注的汽车及高技术产品飞机与航天器为例。材料在汽车生产费用中占有很大比例，如图 1.4-2。但是不同类型汽车，材料所占比例不完全相同，对豪华汽车来说，电子器件所需费用占有更大比例，结构材料的费用相对下降。

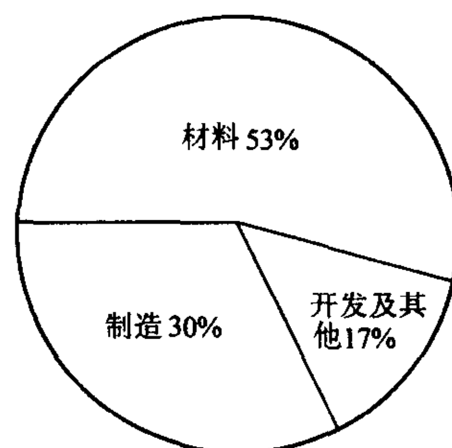


图 1.4-2　汽车工业材料生产与开发管理等费用分布图

汽车是当前最主要的交通工具之一，发展十分迅速。当前汽车制造业的最主要任务是降低能耗、减少污染和降低成本，此外还要保障安全和提高舒适度。为了达到降低能耗、减少污染，需要采取的措施：一是改造动力系统，二是减少车的自重，三是合理设计，改变汽车外形，以减少空气阻力，四是采用先进的电子控制系统和智能系统，不但可以降低能耗，而且可提高车的舒适度与安全性。图 1.4-3 是影响汽车节油的几个主要因素。

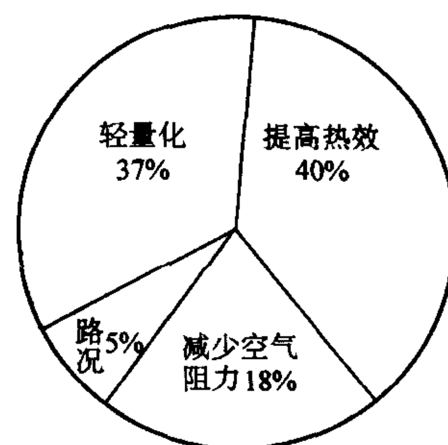


图 1.4-3　降低汽车能耗的几个途径

汽车轻量化主要决定于材料，而提高发动机的热效率有很多途径，都与材料密切相关。从热能量平衡来计算，在采取不同措施后，一个汽车发动机的热效率、尾气排出的热能及因冷却而损失的热能，分别示意如图 1.4-4。

应该指出，图 1.4-4 所示只是理论计算，而与实际情况存在不小差距，如目前汽车热效率不到 30%；但无论如何其相对值还是有参考价值的。当采用增压器后，热效率可能有 5%左右的提高，采用绝热发动机则有 10%以上的收益。但实现这些目标，都需要更高档的材料，如涡轮增压器需要耐热合金、高温合金或正在开发的钛铝化合物。绝热发动机对材料的要求更高，如采用热导率很低的 ZrO_2 工程陶瓷。

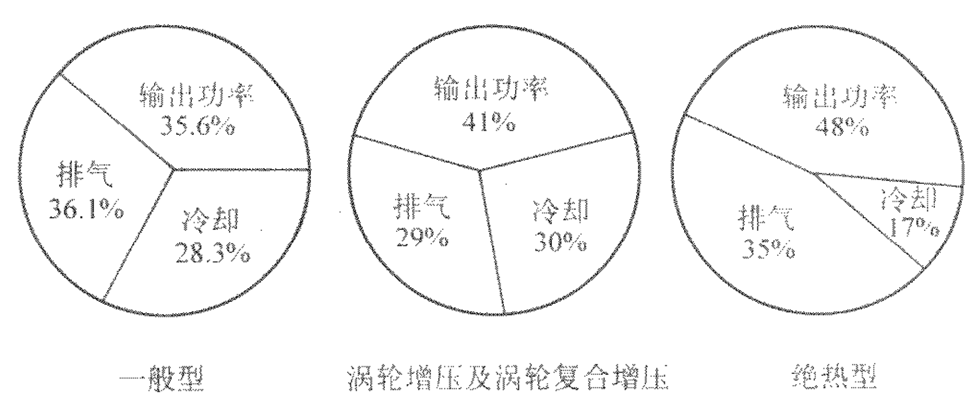


图 1.4-4 不同发动机热平衡（理论计算）

不管采用什么办法，只要用化石燃料作为动力，都避免不了尾气排放的污染问题，这是当前人们最为关注的。因此，为了降低尾气对环境的危害，必须采用催化剂，以改变尾气的成分。治理汽车尾气的催化剂以铂族金属为主，虽然可加入一些氧化铈以改善催化性能并降低成本，但世界铂金属主要消耗在汽车工业。

提高能源利用效率、减少污染的另一途径便是采用电动汽车或燃机与电动混合型动力。图 1.4-5 为原油直接用来发电，而后输入电网，充入电池，再用来开车；和原油精炼为汽油，作为发动机燃料，每升油所能行走的路程，二者相比，约为 2.5 倍。说明电动汽车在燃料利用率方面有明显的优势。但是发展电动汽车的关键是要有大功率、高效、长寿命的蓄电池，蓄电池材料便成为首要考虑的问题。

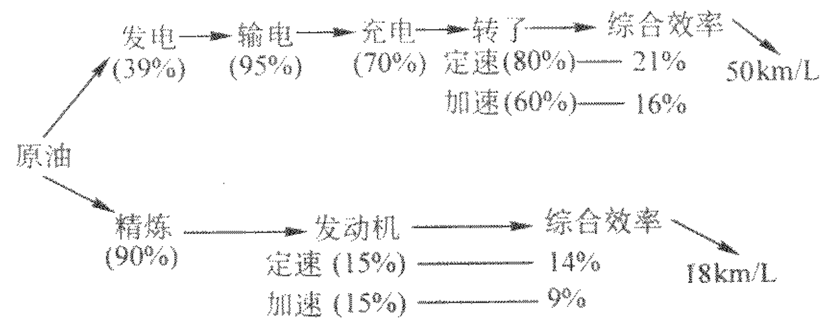


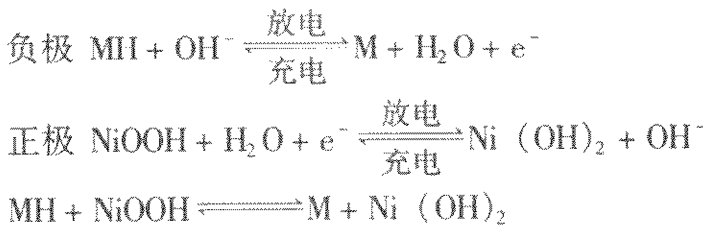
图 1.4-5 汽油发动机与电发动机效率对比

最常见的蓄电池是铅酸电池，但它的能量密度太低，最好的也不过 40 W·h/kg 左右，而内燃机的能量密度为 400 W·h/kg，为电池的 10 倍，因而必须开发高能量密度的蓄电池。1990 年由美国先进电池联合体（USABC）规定了可用于电动汽车的初步标准，见表 1.4-1，同时列出 1993 年由 OBC（Ovonic Battery Co.）所做出的 Ni－MH 电池达到的指标，可以看出，除了价格尚存在差距以外，全部达标，所以目前电动汽车或内燃机－电池混合汽车主要都采用 Ni－MH 电池。

表 1.4-1 USABC 对电动汽车电池的要求及 OBC 开发 Ni－MH 所达到的指标

性能	USABC 规定要求	OBC 达到指标
比能量/W·h·kg⁻¹	80（最好 100）	80
能量密度/W·h·L⁻¹	135	215
功率密度/W·L⁻¹	250	470
比功率/W·kg⁻¹	150（最好 >200）	175
循环寿命/次	600	1 000
使用寿命/a	5	10
使用温度范围/℃	－35～65	－30～65
充电所需时间/h	<6	1
自然放电指标	<15%（48 h）	<10%（48 h）
出厂价格/美元·kW⁻¹·h⁻¹	<150	200

Ni－MH 电池的反应原理为：



负极 $MH + OH^- \underset{充电}{\overset{放电}{\rightleftharpoons}} M + H_2O + e^-$

正极 $NiOOH + H_2O + e^- \underset{充电}{\overset{放电}{\rightleftharpoons}} Ni(OH)_2 + OH^-$

$MH + NiOOH \rightleftharpoons M + Ni(OH)_2$

（负极为储氢合金，如 $LaNi_5$，正极为氢氧化镍）

镍氢电池近几年又有很大发展，而更有竞争能力的还有几种，示如表 1.4-2 和图 1.4-6 所示。

表 1.4-2 电动汽车可能用的几种电池的主要参数

类型	比能量 /W·h·kg⁻¹	能量密度 /W·h·L⁻¹	比功率 /W·kg⁻¹	循环寿命 /次数	预计成本 /美元·kW⁻¹·h⁻¹
Ni－Cd	40～60	80～110	150～350	600～1 200	300
Ni－Zn	60～65	120～130	150～300	300	100～300
Ni－MH	60～70	130～170	150～300	600～1 200	200～350
Zn－空气	230	200	105	—	90～120
NaS	100	150	300	800	250～350
Li 离子	90～130	140～200	250～450	800～1 200	>200
USABC 要求	200	300	400	1 000	<100

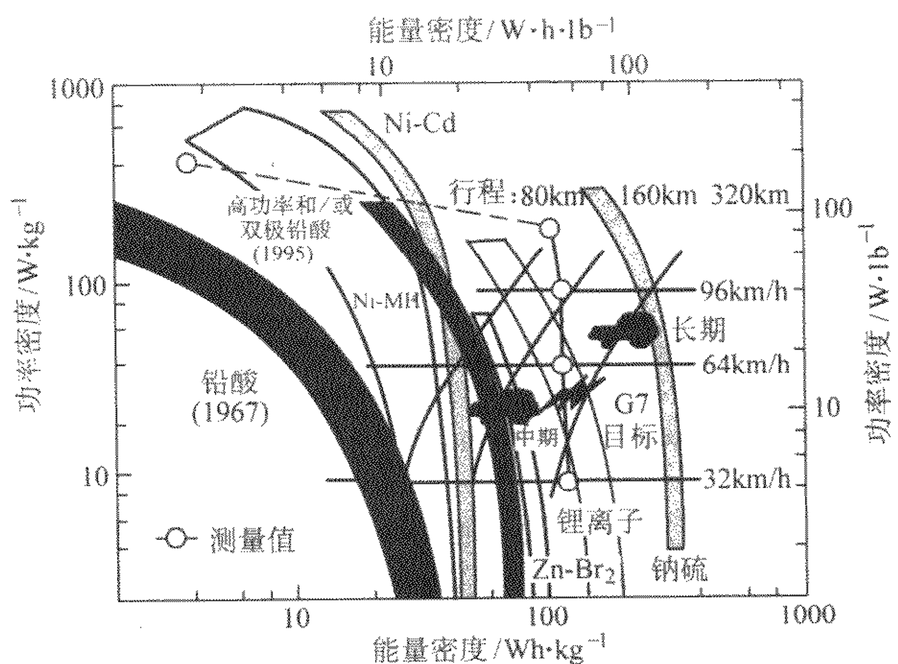


图 1.4-6 几种电池比能量与比功率

从比能量和比功率来看，锂离子电池和钠硫电池都有更大优势，但前者价格较高，而钠硫电池需要在 300℃ 工作，而且有爆炸危险。从目前情况看，锂离子电池由于材料的不断改进及生产规模的不断扩大，提高了竞争能力。

更为理想的是燃料电池，因为它直接由化学能转化为电能，不需蓄电池，而且效率可高达 60%，无污染或少污染。其主要类型如表 1.4-3。

目前有的已成功地用于宇航，但用于汽车的燃料电池仍在开发阶段，其中质子交换膜燃料电池及固体氧化物电池有望用于电动车。但前者需要铂催化剂，资源受到限制，又要质子交换膜，技术难度较大。固体氧化物燃料电池的电解质

为掺杂 Y_2O_3 的 ZrO_2（YSZ），工作温度太高，亟待改善，这些都是材料问题。

表 1.4-3 目前燃料电池的类型与现状

项目 \ 电解质	磷酸盐	熔融碳酸盐	ZrO_2	高分子膜	碱金属
工作温度/℃	150~210	550~650	1 000~1 100	80~110	70~200
燃料	天然气 CH_4	天然气 煤气	天然气 煤气	城市发生气，H_2	天然气 H_2
效率/%	35~42	50~60	≥60	—	>60
所处阶段	成熟	工业实验	实验室	工业实验	工业实验

综上所述，不管是以蓄电池为动力的电动汽车或以燃料电池为动力的电动汽车的瓶颈主要是受材料的限制，这些问题得不到解决，是难以与汽油为动力的汽车竞争的，除非污染指标限定十分严格，达到燃油汽车难以解决的程度。不过燃油与电动混合汽车已形成批量生产。在市区使用电池起动快（差2倍），无污染，在公路上使用汽油，电池得到充电。如此效率高、少排放，但需解决公路旁有充电设施等问题。

随着汽车自动化与智能化程度的提高，所需功能材料与电子材料大量增加，如用于汽车的微型马达高达40多个，钕铁硼高磁能积磁性材料便发挥很大作用；为了自动化控制精度的提高，需要灵敏度高，可靠性好的传感器等。这些都说明材料对汽车工业的重要性和所处的地位。

造船工业与材料的关系更是密切，如载重1万吨民用船，需要钢铁近4 000 t，有色金属22.5 t，油漆52.5 t，70 km电缆，玻璃53箱和水泥100多吨，还不包括6 400 hp的内燃机和近400 kW的发电机。

前面以较长的篇幅描述材料在汽车各个发展阶段的重要性和地位，但没有涉及选材问题，这将在以后以案例分析的形式对汽车车体材料进行说明。

下面对材料在宇航事业中的作用和地位进行简单分析。在此之前，先引用美国L.Schwartz发表在"前进中的材料研究"的一篇文章中的一段话，说明材料在空间技术的重要性。当苏联在1957年10月发射人造卫星以后，他在1958年召开的"材料研究与开发协调委会"（Coordinating Committee on Materials Research and Development）上，向国防部强调："看来今后国民经济和国防安全愈来愈依赖于新技术，而在很多情况下要求更可靠的新材料"。随即相继成立了若干交叉学科与材料研究的实验室，12年后交国家基金会（NSF）管理。这个事实说明材料在美国宇航事业中的作用。

宇航材料可分为航空材料与航天材料，因为它们的工作状态不同，所处环境不同，从而对材料的要求不一样。例如所有飞行器都要求高的比强度，但比强度的提高所收得效益却随飞行速度的提高而以数量级的增加，如图1.4-7。所以飞行速度愈高，对材料减重的要求愈严格。

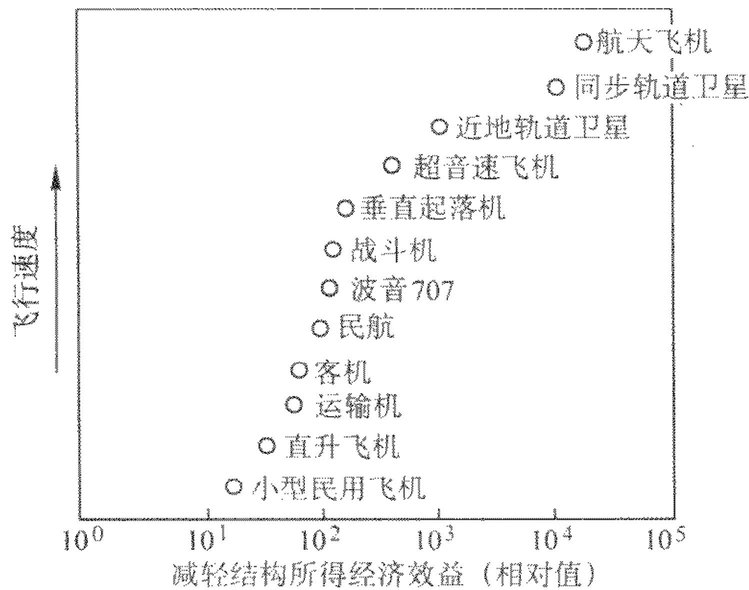


图 1.4-7 飞行器飞行速度与收益的关系

对飞机来说，减小质量要从以下四个方面着手：即自动化设计、采用先进结构，主动载荷控制，振动抑制与新材料的采用和现有材料的改进，如图1.4-8，可以看出，飞机减重要以材料为主，占2/3左右。

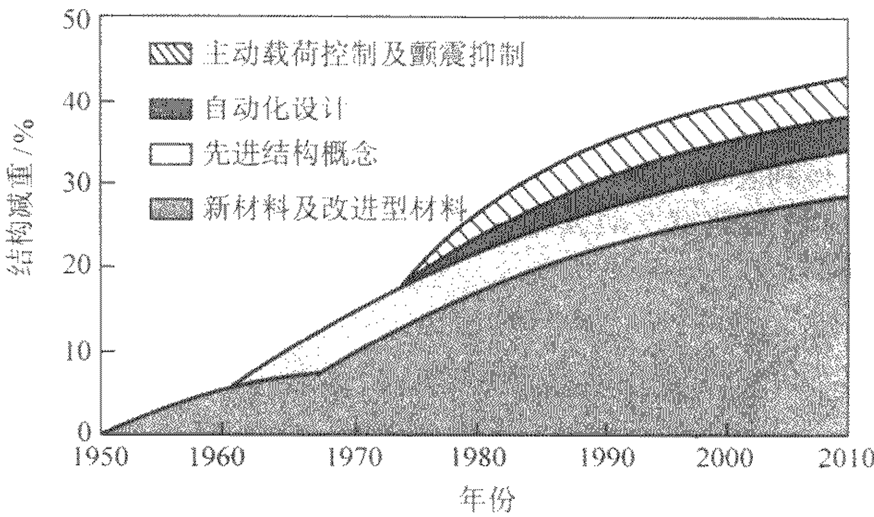


图 1.4-8 新材料与材料的改进对军用飞机减重所占比例

根据近年来的发展，飞机材料品种的变化以空中客车为例，其发展历程如图1.4-9。

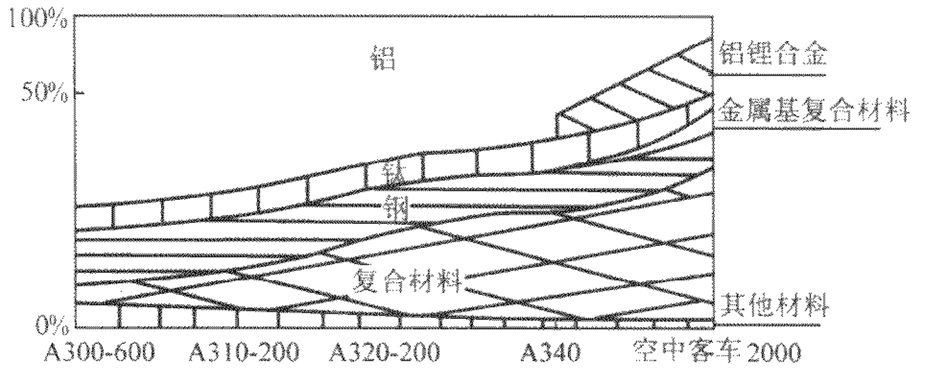


图 1.4-9 空中客车飞机材料应用发展

从图上明显地看出，复合材料的比例逐年增加，部分铝合金被Al-Li合金所代替。由于冶炼工艺及加工工艺的改进，解决了合金元素偏析及成形问题，开发出含锂2%~3%的Al-Li合金。合金中每增加1%锂，合金密度降低3%，弹性模量提高6%，性能明显提高，这对航空材料来说至关重要。通过材料的改变，从1980~2000年结构减重20%，也就是说有效载荷提高了20%。飞机的减重对发动机的要求降低了，载油量减少了，航程提高了，都会产生明显的经济效益，如波音747在一个寿命期内，每减重1 kg，可提高收益2 000美元，一架上百吨的大型飞机，由于减重而得到的收益就十分可观了。

为了飞机的减重，在材料方面有各种设想。由于有机材料密度小，如达到分子复合，比强度为钢的10倍，用于制造飞机，马赫数可达15，从纽约到东京只需2 h。当今纳米科技的呼声甚高，碳纳米管的比强度也10倍于钢，而且可耐更高温，有更大的现实性。总之，为降低能耗，减少污染，提高运力，改善飞机材料性能是最主要措施。

航空发动机是飞机的心脏，是确保飞机使用性能、可靠性和经济性的决定因素。航空发动机的工况十分复杂，高温是最重要的特点，因为提高工作温度可大幅度提高发动机的效率，如当涡轮前温度为1 000℃时的效率为35%，提高到1 500~1 600℃，其效率可达50%。还有资料表明，涡轮前温度每提高100℃，发动机推力提高20%，都说明高温材料在航空发动机的重要性。衡量航空发动机的综合指标为推重比，因此，材料的比强度和比刚度又是另外一类重要指标，因发动机每减少1 kg，飞机可减重4 kg，飞机升限可提高10 m。航空发动机是一高速旋转机械，因此，高低温度的疲劳

性能，包括应力疲劳与应变疲劳，是航空材料又一重要指标。航空发动机的使用寿命很长（几百小时到几千小时），且由于多次起飞降落，材料的裂纹扩展速率成为衡量材料的另一类参数。此外，根据不同工况，材料的耐蚀、抗氧化与耐磨损能力也是不容忽略的性能，如此种种，不一一列举。作为航空材料，更重要的一个特点是可靠性，因为载人航空器，人命关天，有关材料质量控制将在后面专门论述。

由于航空发动机的设计与制造日新月异，发动机材料也随之得到发展，图1.4-10a、b分别说明英国Rolls Royce及美国NASA对航空发动机材料从20世纪中叶到21世纪的进展与预测。

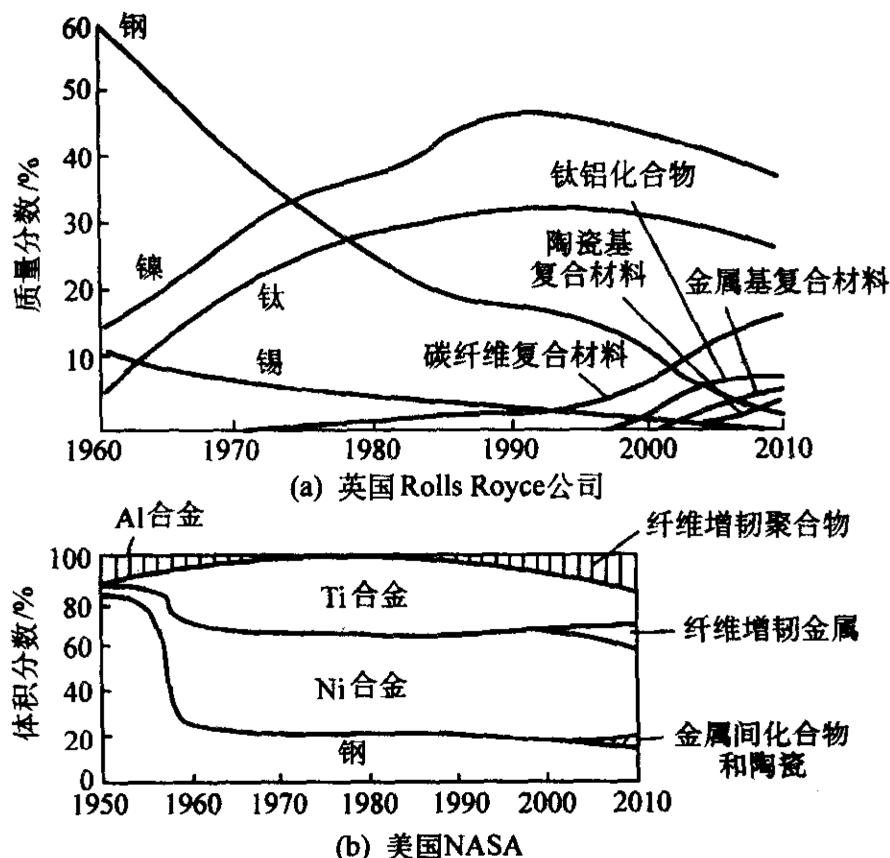


图1.4-10　航空发动机材料的发展与预测

航空飞行器实际由三部分组成，除前述飞机和发动机以外，还有机载装置，特别是雷达和电子设备，是保障飞机安全所必须，对战斗机来说尤为重要，因这是提高生存率和战胜对方的关键，所用电子元件和功能器件都与材料密切相关，而且品种更多，要求更高，限于篇幅，不加讨论了。

航天材料除对比强度和比刚度的要求更高以外，还有更为特殊的要求，如在大热流下工作的耐超高温或烧蚀材料；在超高真空工作的材料；耐粒子云及原子氧侵蚀的材料；能耐受高温与低温交变的材料。

图1.4-7已说明材料的减重对高速运行的人造卫星、航天飞机的重要性。图1.4-11为导弹壳体材料与射程的关系示意图，可以看出，从金属壳体换为C-C复合材料，其射程大幅度增加。

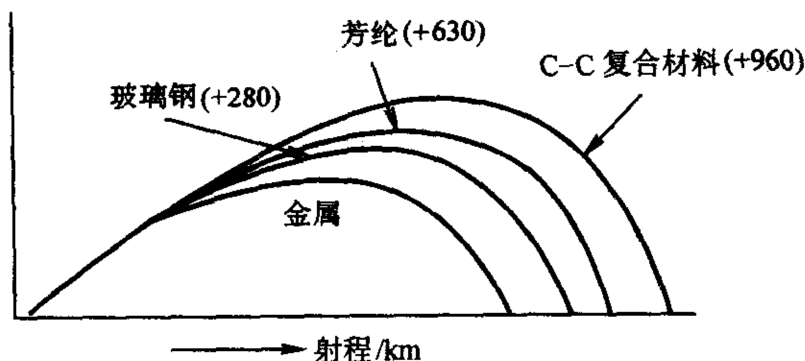


图1.4-11　导弹壳体材料与射程关系

我国长征火箭由铝合金（744 kg）换为复合材料（401 kg）减重343 kg，射程增加1 476 km。还有资料表明，洲际导弹每减重1 kg，可提高第一级有效射程1.8 km，二级为3.6 km，三级则为18 km，这些都说明材料的轻量化对航天的重要性。

减重1 kg，对上面三类产品（汽车、飞机、空间飞行器）来说所取得的效果不相同，但考虑到世界生产总量，汽车收益最大，如表1.4-4。

表1.4-4　汽车、飞机及空间飞行器的结构材料

产品	在全寿命内每减重1 lb[①]所得收益/美元	世界每年销售数目	总收益/百万美元
汽车（10万公里）	2	30 000 000	600
民机（10万小时）	200	2 500	150
空间飞行器（1次）	20 000	100	20

① 1lb = 0.45 kg。

2　机械设计过程中选材的几点考虑

设计是机械制造的灵魂与核心，它不但决定产品的性能，也影响产品的价格和市场竞争能力。机械产品的功能决定于设计、材料及加工工艺，其关系如图1.4-12所示。可以看出，产品要达到所需的功能，不但与选材有关，与制造工艺也有直接关系。有些产品选材得当，但工艺过不了关，产品也制造不出来，而且工艺的难易，直接影响产品的成本和市场竞争能力。

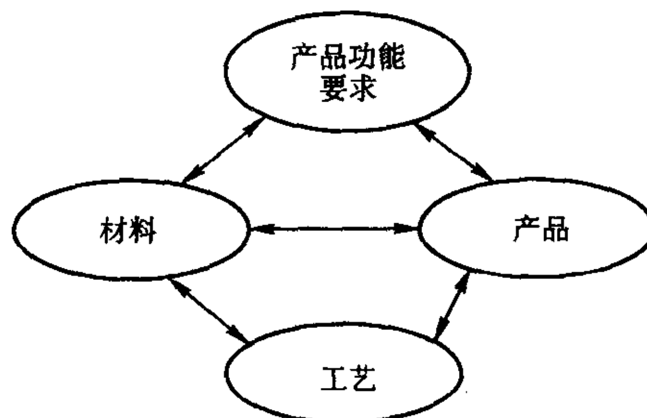


图1.4-12　机械设计中材料与产品性能要求及其与工艺间的关系

机械设计步骤及其与材料的关系如图1.4-13所示，这是根据剑桥大学M.F.Ashby教授提出的方案综合而成。他强调设计是一个反复的过程，最初决定于市场的需求，而后根据需求提出概念。如果在诸多概念选择中，认定一个是可行的话，设计便进行到具体化阶段，即选好工作原则，确定布局及规模，完成效能及效益的初步估算。如果结果是成功的，再进入细节设计阶段，即效能的最佳化，关键部件的认真分析，准备细节生产图、公差规范、精度、连接方法及表面粗糙度等，每个阶段都有选材问题，但是考虑的宽度与精度不同。

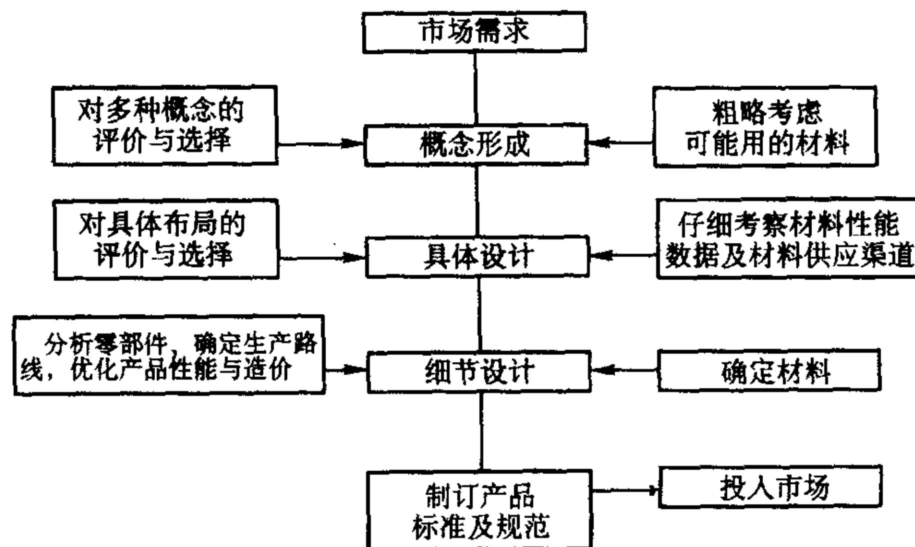


图1.4-13　机械设计步骤与选材的关系

在概念设计阶段，所有材料和生产工艺都是考虑的范围，要求很快得到有关数据，尽管不要求很精确。图 1.4-14～图 1.4-26 选自剑桥大学 M.F.Ashby 所著《Materials Selection in Mechanical Design》，为不同材料各种性质间的相互关系，可以作为机械设计初步选材的参考；但到精细设计阶段，需要更准确的数据和更多的信息作为最终设计的依据。在这些图中，把材料分为几大类，如泡沫塑料（polymer foam），木材（woods），工程塑料（engineering plastics），合成橡胶（elastomers），工程复合材料（engineering composites），多孔陶瓷（porous ceramics），工程陶瓷（engineering ceramics）及工程合金（engineering alloys）等几大类，其中还标出具体材料所处地位。

本文共引用了 13 幅图，其中包括：模量－密度，强度－密度，断裂韧度－密度，模量－强度，比模量－比强度，断裂韧度－模量，断裂韧度－强度，衰减系数－模量，热导率－热扩散率，线胀系数－热导率，线胀系数－模量，强度－线胀系数、强度－温度。

图中有几条虚线是设计控制线（guide line），对不同要求的构件有不同的控制线，在每个图下都有说明。至于公式的导出请参阅 Ashby 的原书有关章节。

图 1.4-14　材料弹性模量与密度

(1) $E/\rho=C$（适用于设计高刚性螺栓及在高离心力载荷下最小弯曲的部件）；(2) $E^{1/2}/\rho=C$（适用于最小质量的高强螺栓、轴和柱）；(3) $E^{1/3}/\rho=C$（适用于高强、质量小的板材）

下面是选材需要考虑的几个问题。

(1) 材料的性能

材料性能包括使用性能和工艺性能。使用性能因产品的性质而不同，如用于空间飞行器及高速旋转的部件，比强度和比刚度是首要考虑的问题。用于锅炉或电站的过热器管材，除了耐热与耐蚀以外，材料的热导性能对热效率的影响很大。航空发动机用的高温合金叶片，工作状态复杂而苛刻，除了要求在高温下的低蠕变速率，低裂纹扩展率和高抗疲劳、抗氧化和热腐蚀性能以外，热导率也很重要，因为它可以减少局部过热的危害，特别是为了提高耐热强度而采用气冷措施时，材料的导热性能便是一个重要参数了。在有些情况下，线胀系数是关键，如用于急冷急热的部件，高线胀系数容易产生热疲劳或脆性材料的炸裂；但是在某些情况下，利用线胀系数的变化实现管材的连接，如此种种，都说明材料的力学性质、物理化学性质对机械制造有着决定性的作用。

选材过程中，材料的工艺性能也不容忽视，有时甚至是决定性的。工艺性能不佳，成品率就低，因而增加产品的成本。我国有关机械制造的热加工工艺落后，如表 1.4-5 所示，铸锻件的成品率与国外存在明显差距，因而造成能耗高、污染严重、产品成本高。为此，必须重视精密成形，由此可以节材 50%～90%，减少加工工时 30%～70%，生产成本可降低 20%～50%。

材料的工艺性能涉及面极其广泛。对金属材料来说，从冶炼到铸造，到轧制成板材，锻造成锻件，都存在收得率或合格率问题；板材的深冲性能与焊接性能，以及冷加工性能都影响着产品的质量和成本。对塑料来说，塑料种类繁多，但从工艺性能来分，可分为热塑性塑料和热固性塑料，前者加热后软化，可塑性好，冷却后变硬，再加热又会变软，如此反复，而其性能不变；热固性塑料则在一定温度下加热或加入固化剂后，即可固化成形，固化后塑料性质稳定，不再溶于溶

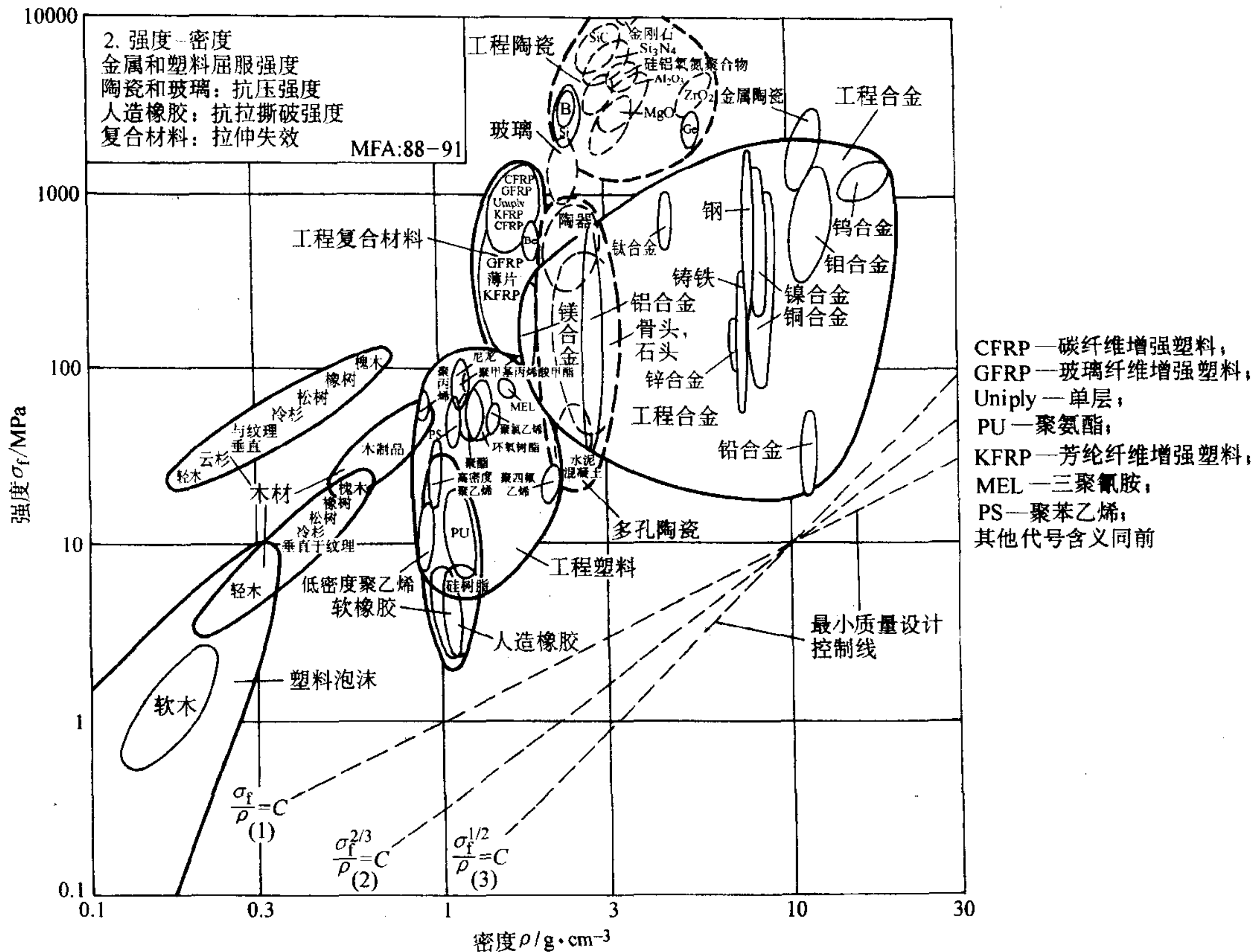


图 1.4-15 材料强度与密度

(1) $\sigma_f/\rho=C$（适用于最小质量的高强拉杆，最大转速的盘形设计）；

(2) $\sigma_f^{2/3}/\rho=C$（适用于最小质量的高强梁或轴的设计）；(3) $\sigma_f^{1/2}/\rho=C$（适用于最小质量高强板的设计）

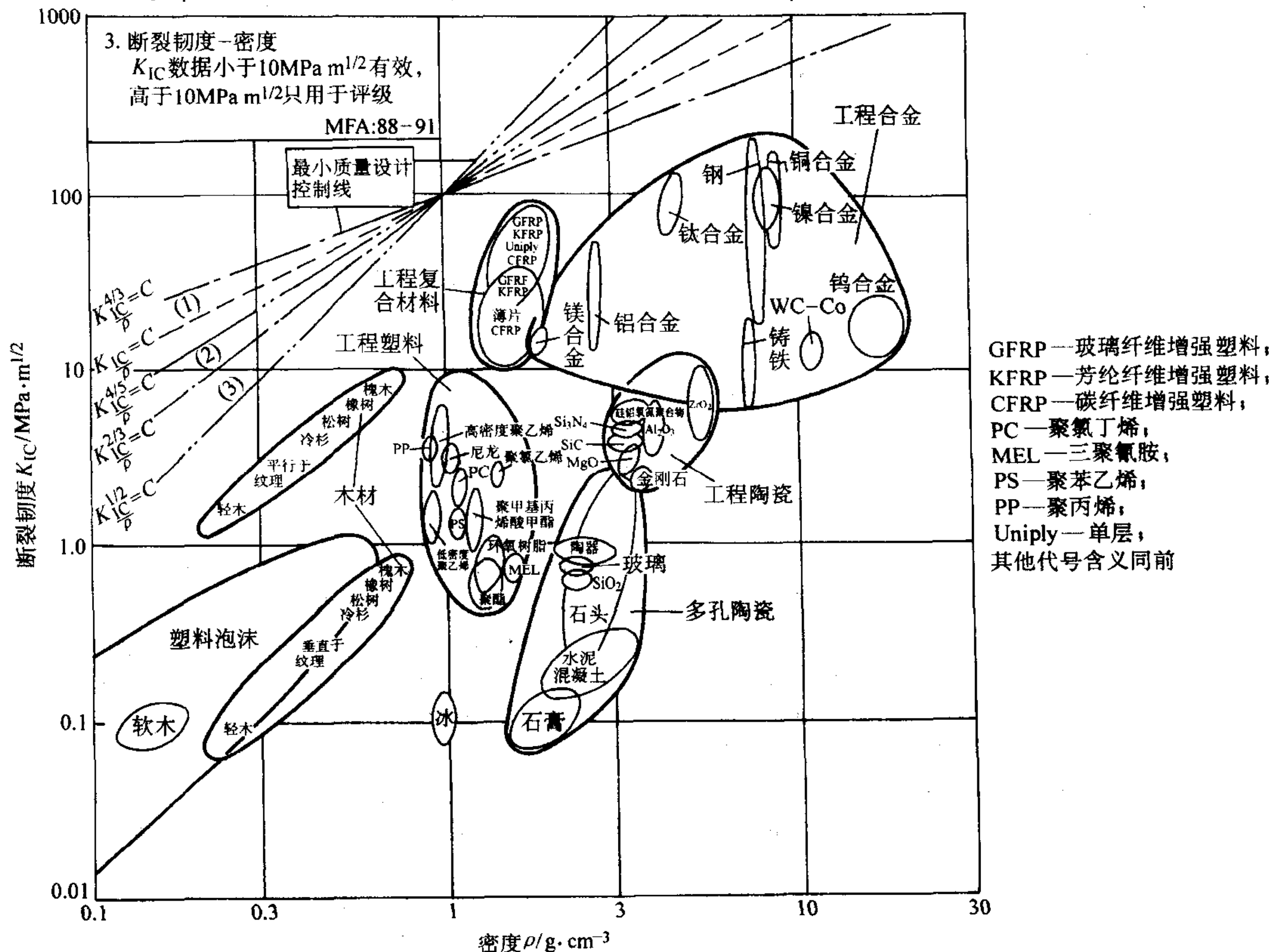
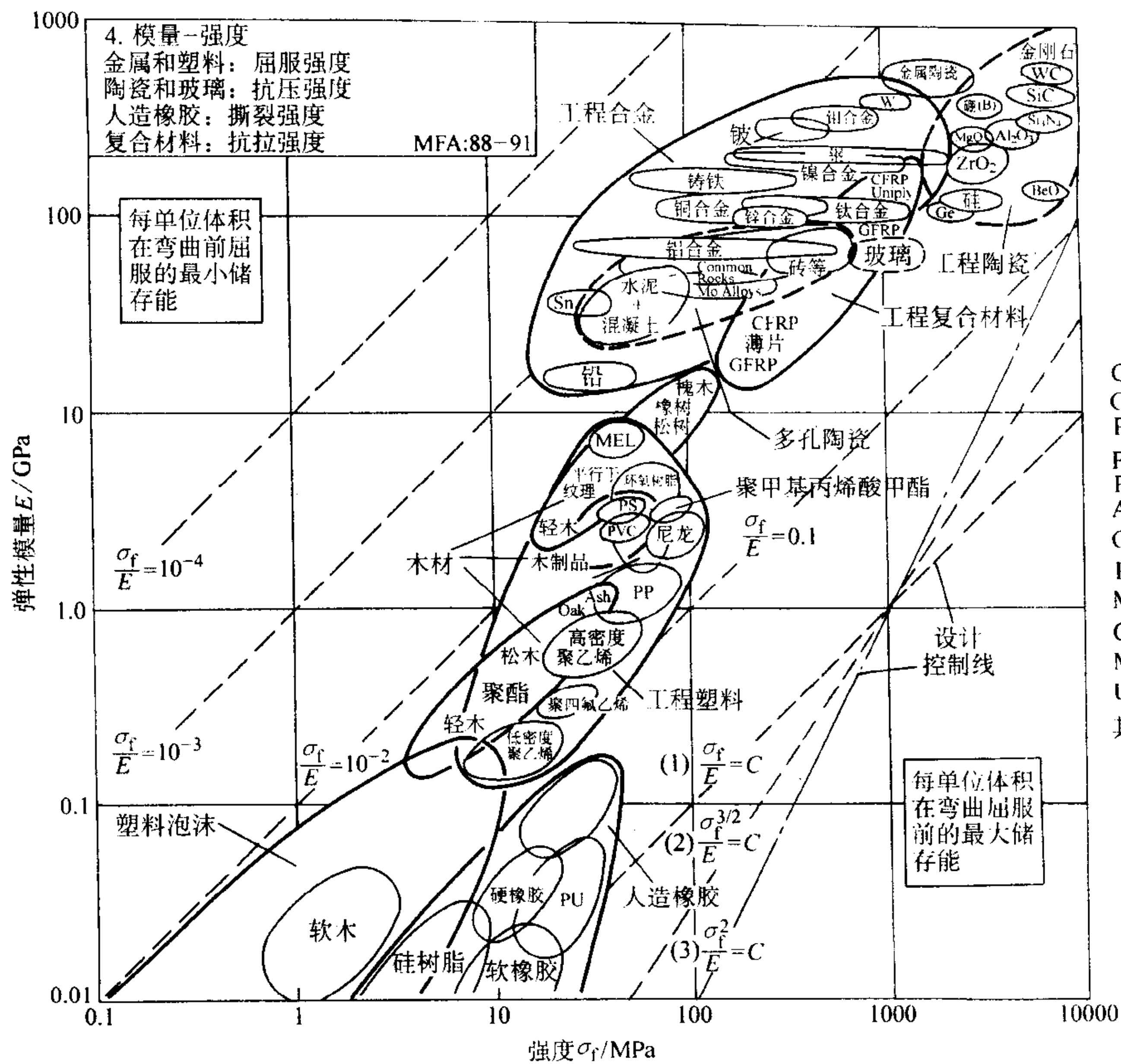


图 1.4-16 材料断裂韧性与密度

(1) $K_{IC}^{4/3}/\rho=C$（脆性拉杆、脆性盘最大转速的最小质量设计） $K_{IC}/\rho=C$；

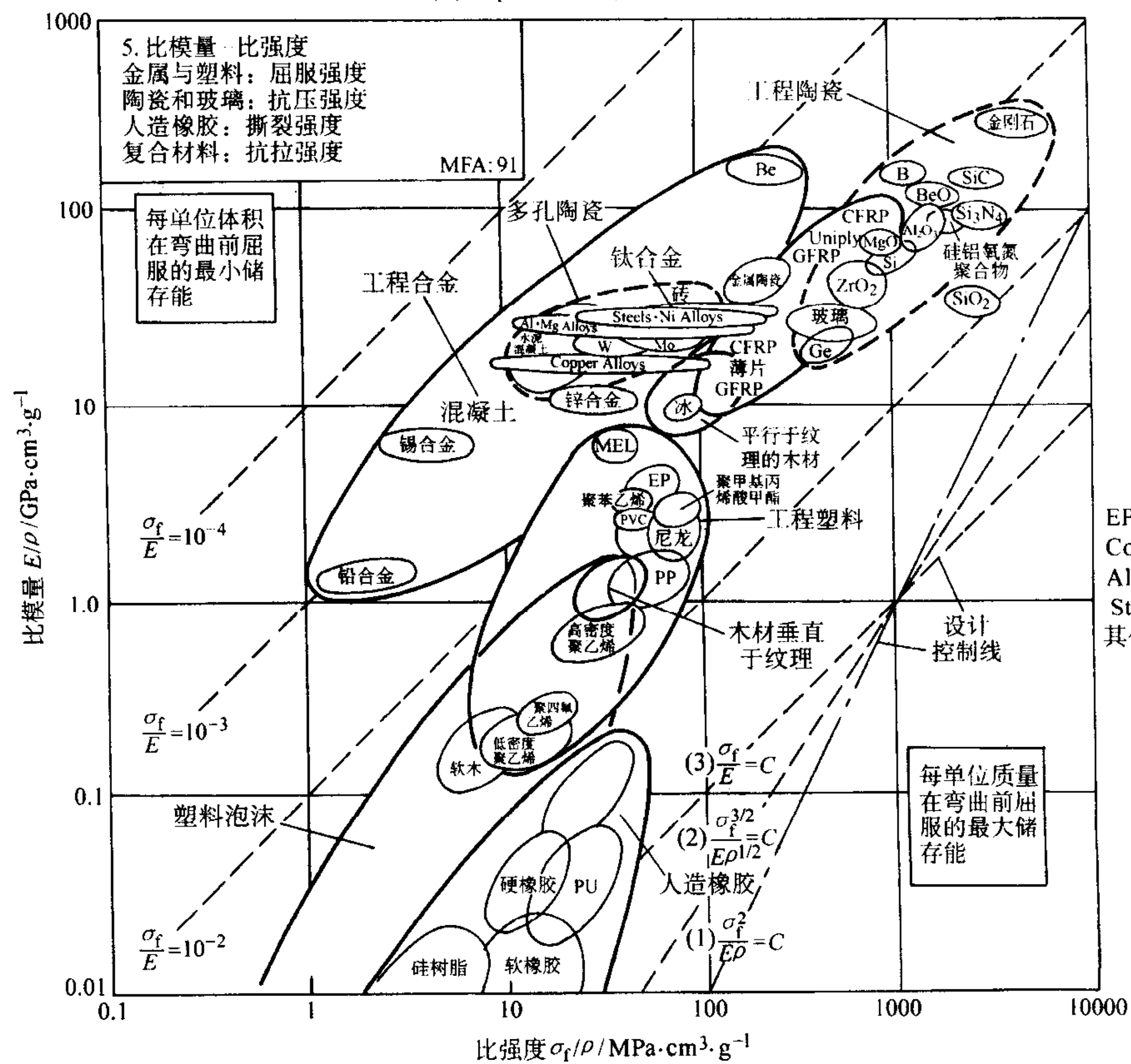
(2) $K_{IC}^{4/5}/\rho=C$（脆性梁及轴的最小质量设计） $K_{IC}^{2/3}/\rho=C$；(3) $K_{IC}^{2/3}/\rho=C$（脆性板的最小质量设计） $K_{IC}^{1/2}/\rho=C$



CFRP—碳纤维增强塑料；
GFRP—玻璃纤维增强塑料；
PS—聚苯乙烯
PVC—聚氯乙烯；
PP—聚丙烯
Ash—槐木；
Oak—橡木；
PU—聚氨酯；
MEL—三聚氰胺；
Common Rocks—普通石头；
Mo Alloys—钼合金；
Uniply—单层；
其他代号含义同前

图 1.4-17　材料模量与强度

（1）$\sigma_f/E=C$（弹性转轴）；（2）$\sigma_f^{3/2}/E=C$（弹性常数的选择，如刀刃、弹性膜等）；
（3）$\sigma_f^2/E=C$（弹簧、单位容积可存储的弹性能）



EP—环氧树脂；
Copper Alloys—铜合金；
Al–Mg Alloys—铝镁合金；
Steels·Ni Alloys—钢·镍合金；
其他代号含义同前

图 1.4-18　材料比模量与比强度

（1）$\sigma_f^2/E\rho=C$（拉杆、最小质量的弹簧、最大转速的圆盘）；
（2）$\sigma_f^{3/2}/E\rho^{1/2}=C$；（3）$\sigma_f/E=C$（弹性转轴设计）

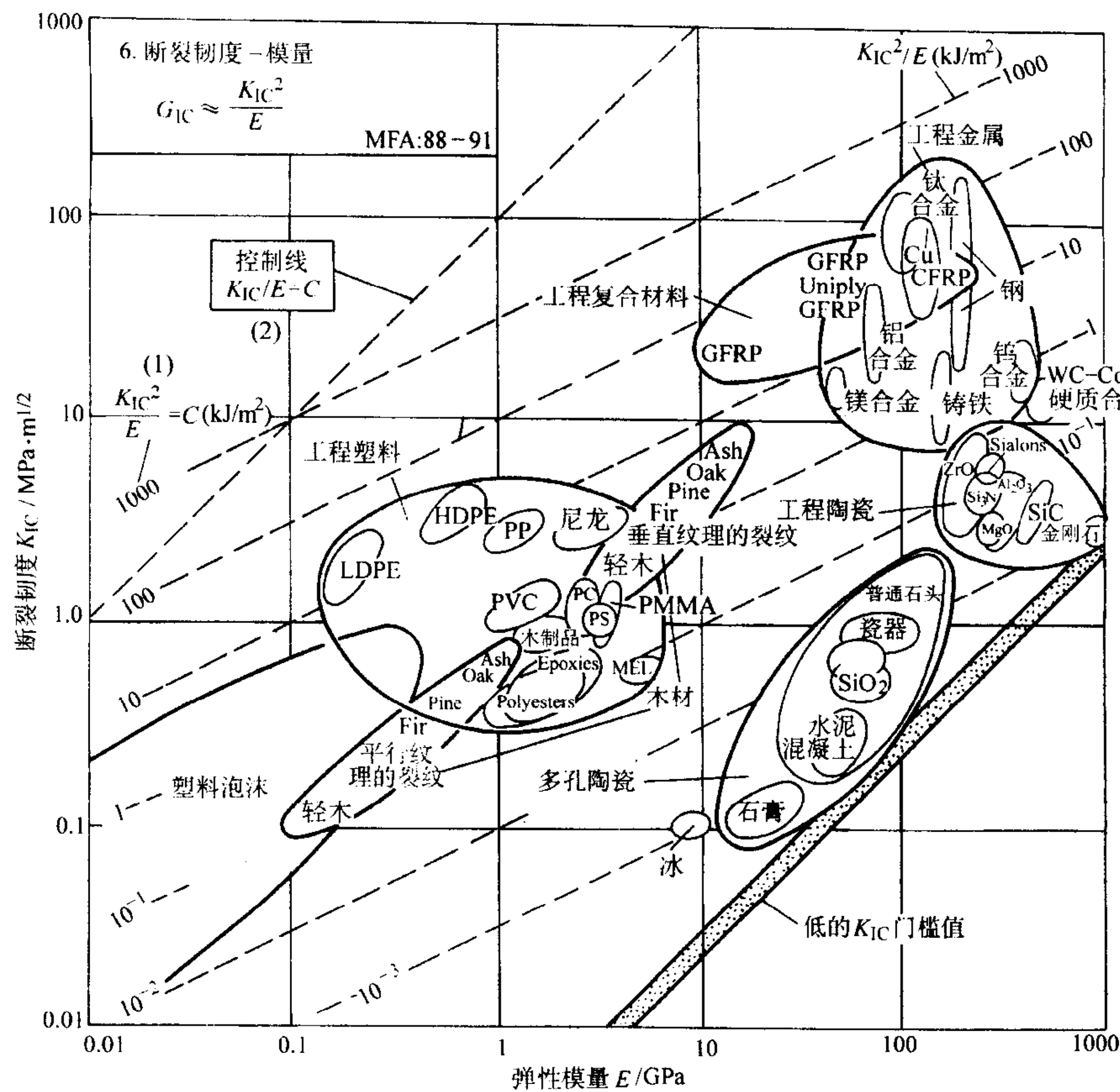


图 1.4-19 材料断裂韧度与模量

(1) $K_{IC}^2/E=C$（等韧性线）；(2) $K_{IC}/E=C$（有限位移脆断的标志线）

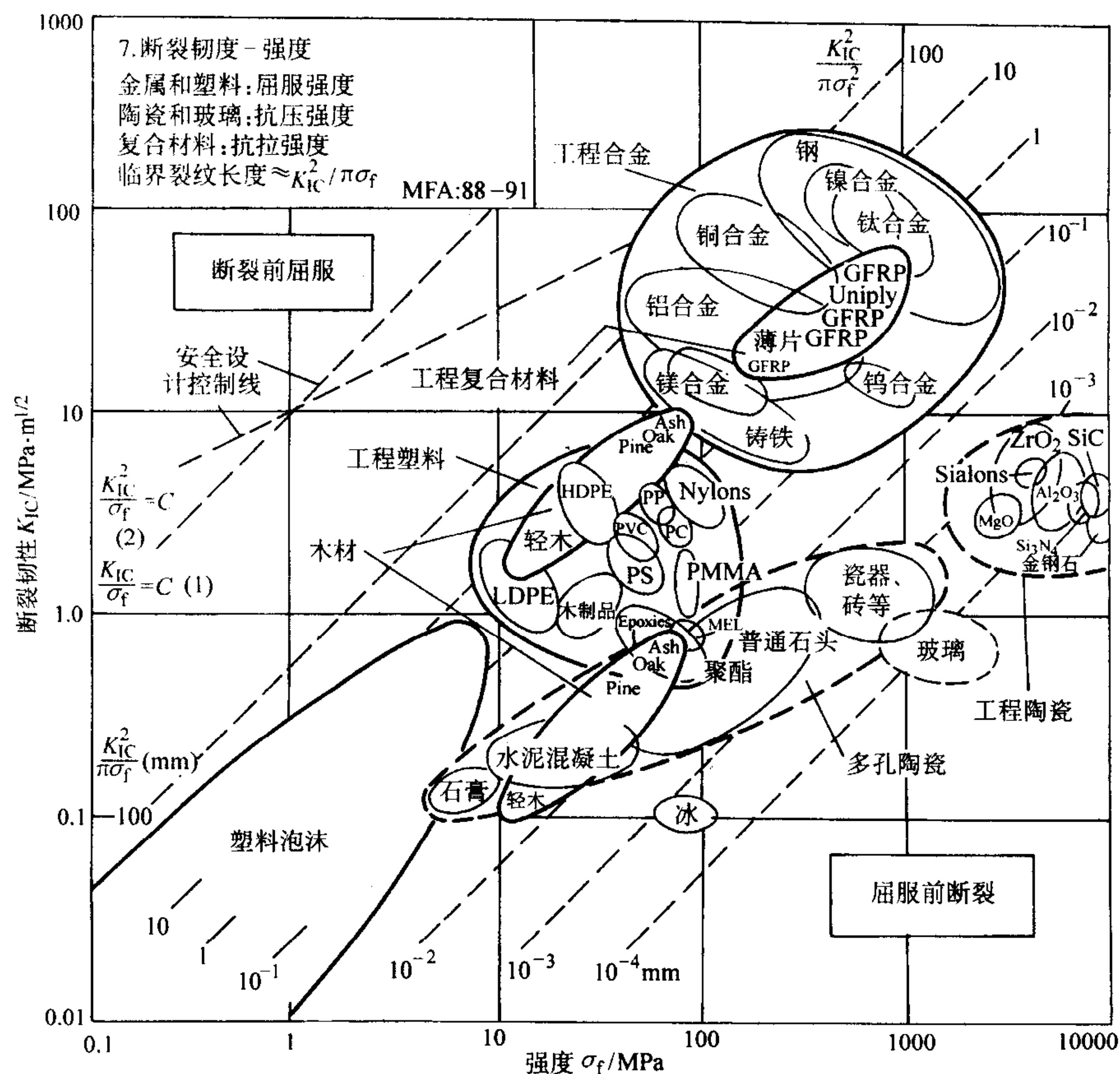


图 1.4-20 材料断裂韧度与强度

(1) $K_{IC}/\sigma_f=C$（断裂前的屈服点）；(2) $K_{IC}^2/\sigma_f=C$（断裂前的渗漏）

代号含义同前

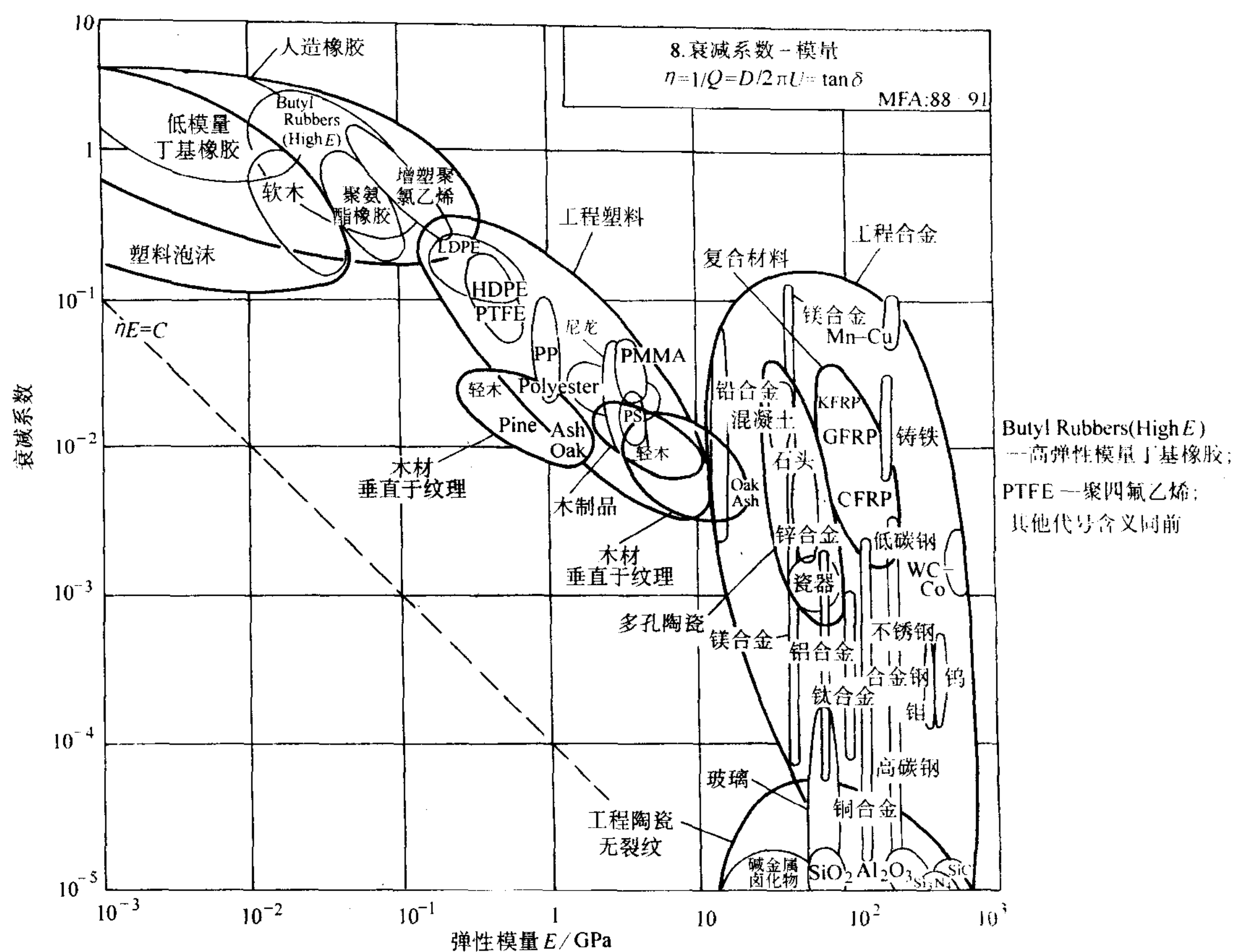


图 1.4-21 材料衰减系数与模量

$\eta E=C$（这是高分子聚合物估计衰减的法则）

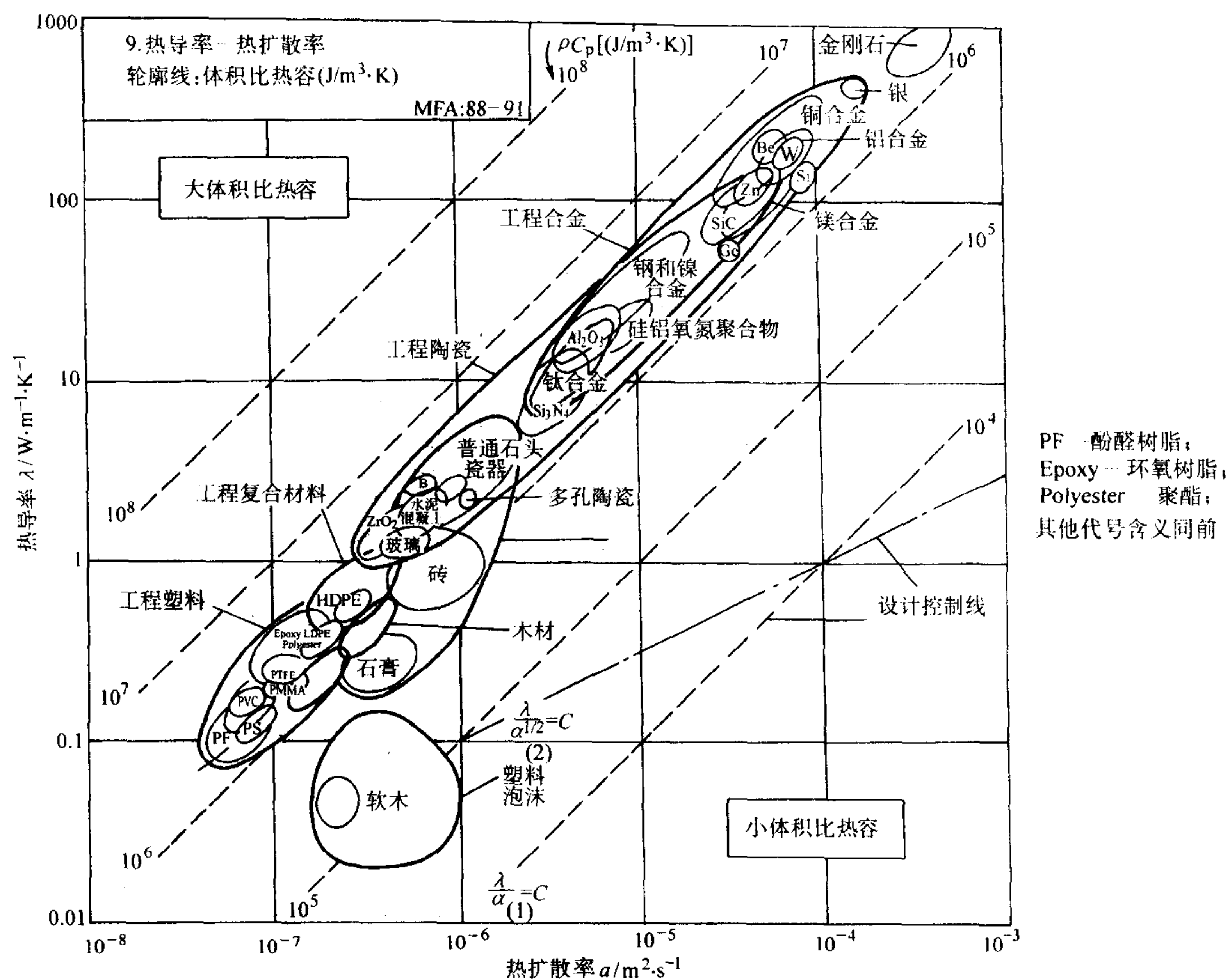


图 1.4-22 材料热导率与热扩散率

(1) $\lambda/\alpha=C$（恒定体积比热容）；(2) $\lambda/\alpha^{1/2}=C$（有效阻热程度）



图 1.4-23 材料线胀系数与热导率

$\lambda/\alpha = C$（热变形参数）

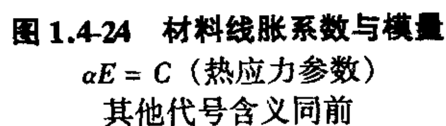


图 1.4-24 材料线胀系数与模量

$\alpha E = C$（热应力参数）

其他代号含义同前

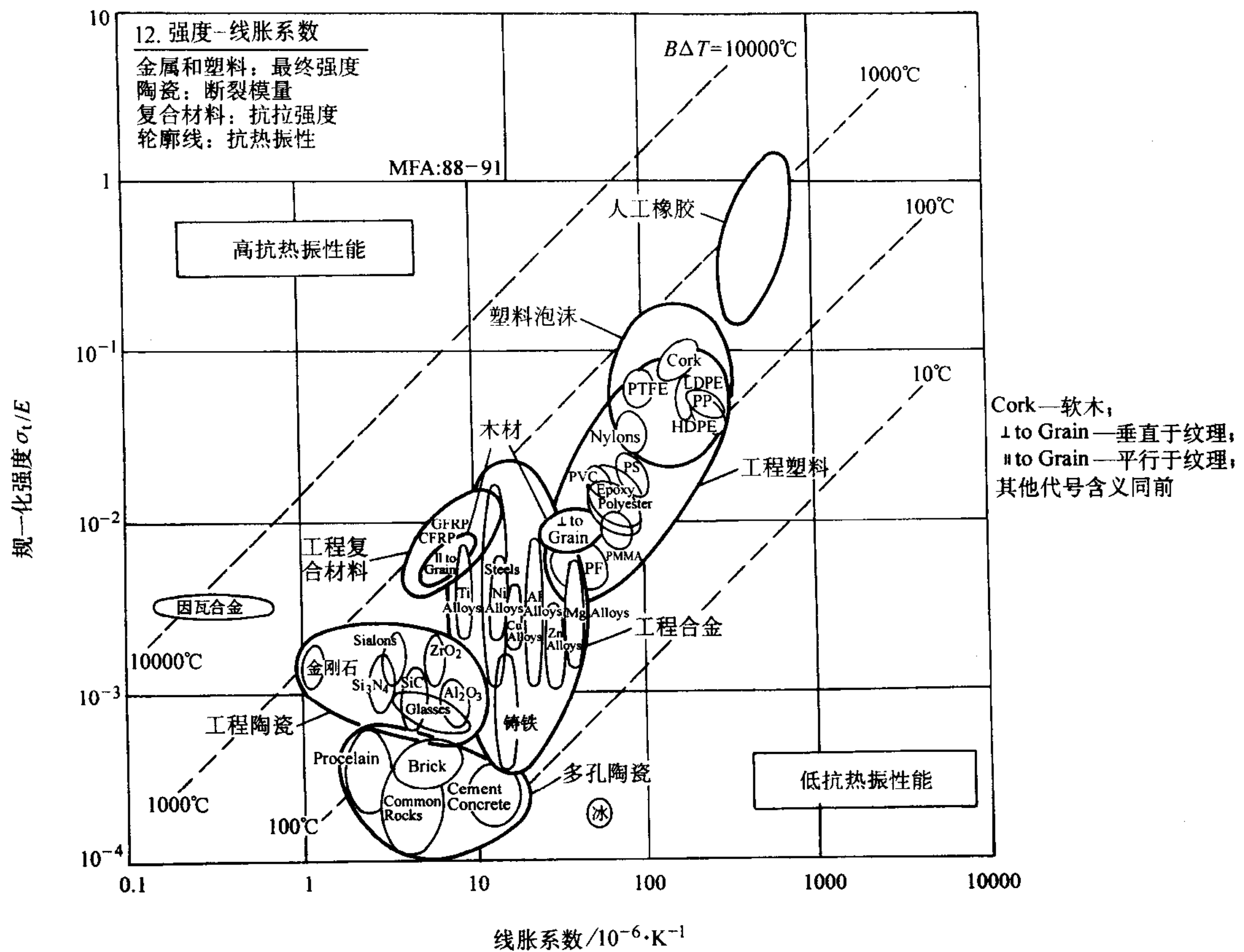


图 1.4-25　规一化强度与线胀系数

热震参数 $B\Delta T=\sigma_f/\alpha E$　B 为与约束及热导相关的因子

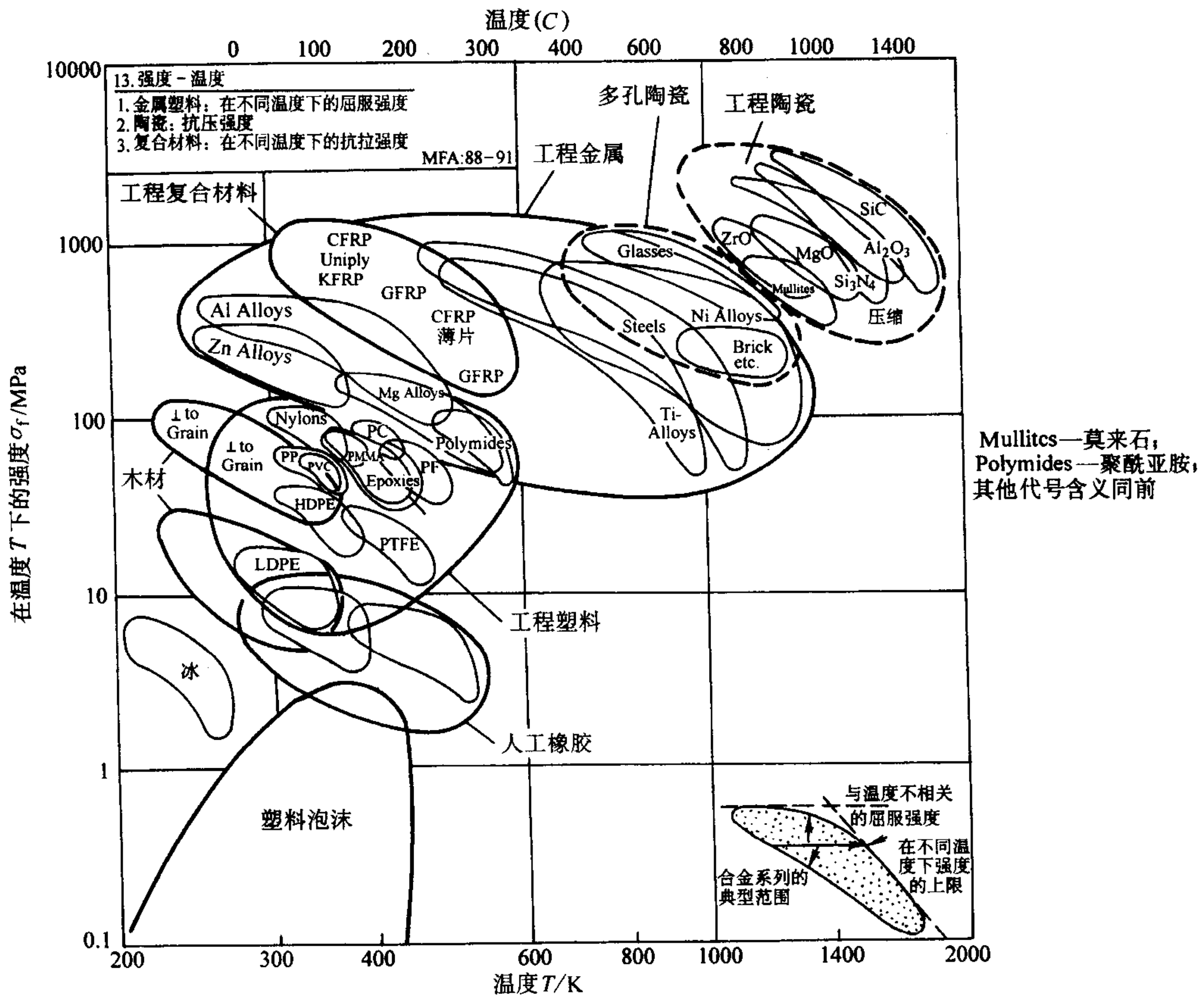


图 1.4-26　材料高温强度与温度

不同温度下的强度，金属及塑料为屈服强度，陶瓷为抗压强度，复合材料为抗拉强度

表 1.4-5 我国与工业发达国家材料加工主要指标对比

性能	中国	工业发达国家
精铸件收得率[①]/%	3~4	25~30
精锻件收得率/%	37	65~70
单位产值能耗	7	1（以日、美为基础）
原材料消耗	2.5	1（以日、美为基础）
工业增加值率	0.50	1（以日、美为基础）

① 精铸航空涡轮叶片收得率我国可达 50%，美国为 70%~90%。

剂中，加热也不会变软，直至分解。塑料的成形方法大致可分为注塑成形，主要用于热塑性塑料和压制成形，主要用于热固性塑料；也有些塑料采用烧结成形，如聚四氟乙烯。塑料的工艺性能比金属材料优越，因而得到广泛应用。工程陶瓷因其熔点高、质脆、工艺比较复杂且成本高，一般采用粉末压制成形而后烧结；为了使产品的致密度提高，在烧结过程中采用热等静压，占成本比例更高的是机械加工，如图 1.4-27 中所示。

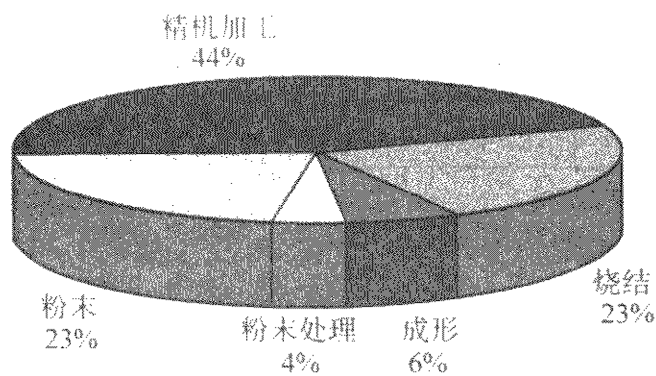


图 1.4-27 陶瓷汽车阀门制造工艺成本分解

复合材料的工艺更为复杂，以树脂基复合材料为例，首先根据复合材料使用部位的工况进行力学性能设计，而后决定增强体的布置形式，再进行熔融塑料的浸渍工艺。金属基复合材料及陶瓷基复合材料的制备工艺多种多样，对材料性能及成本都有决定性的影响。

近年来，金属泡沫及陶瓷泡沫材料逐渐进入市场。这类材料密度小，阻尼性能好，有很高的比强度和比刚度，有广泛应用前景，如图 1.4-28，但泡沫金属与陶瓷的制造工艺复杂，质量控制难度大，是一类正在开发中的材料。

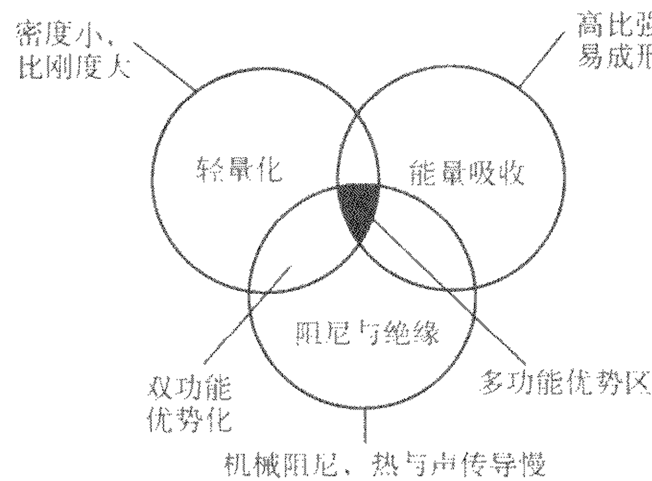


图 1.4-28 泡沫金属的特性及应用

(2) 材料的价格（成本）

根据产品的种类和性质的不同，材料价格差异很大。图 1.4-29 为不同产品所需材料的相对价格。如建筑材料为 1，则汽车、军机材料分别为 10 和 1 500，生物医用材料高达 50 000。因此对后者质量要求十分严格，而对建材来说，产量和廉价工艺更为重要。

影响产品价格的因素很多，材料是其中之一。一般来说，材料在机械产品中的成本在 10%~20%，如表 1.4-6。

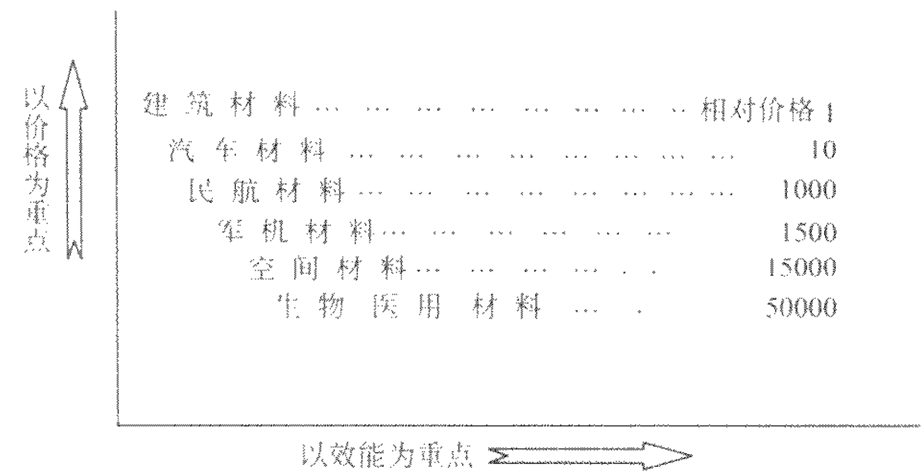


图 1.4-29 不同产品所需材料的相对价格

表 1.4-6 机械构件典型成本分布

项目	在整个成本中的比例/%
原材料（不包括复合材料）	10~20
设计与工程	10~20
制造（锻、机加、连接等）	20~40
非破坏检验与质量控制	10~20
管理	10

但不同产品的材料在成本中所占比例不同，如汽车中的结构材料占成本的一半以上；商船的比例更高。在这些产品中，设计与检测和质量控制所占比例就相应下降。

影响材料价格的因素很多，首先是资源的制约。有些地壳中储量少而又有重要用途的元素，如贵金属、稀有金属等，价格昂贵，一般都是限用或少用。其次是技术进步，工艺的不断改进，会使材料的价格上升缓慢。图 1.4-30 为几种重要材料在 1974~1994 年 20 年间的价格的变化。塑料和水泥的价格不断上升，而钢材的价格波动甚小，主要因为从炼铁到轧钢的工艺不断改进，从而劳动生产率不断提高，能耗不断下降，产品价格不随其他物价的上涨而明显提高。最后是与材料的产量密切相关，图 1.4-31 为美国各种材料年用量与单位重量价格的关系。在双对数坐标图上，基本是一个直线关系，但从斜度看，材料价格下降 2 倍，材料的应用量增加 4 倍，由此说明材料价格的下降会大大促进材料的应用。反之，扩大材料的应用，就促进材料价格的下降。

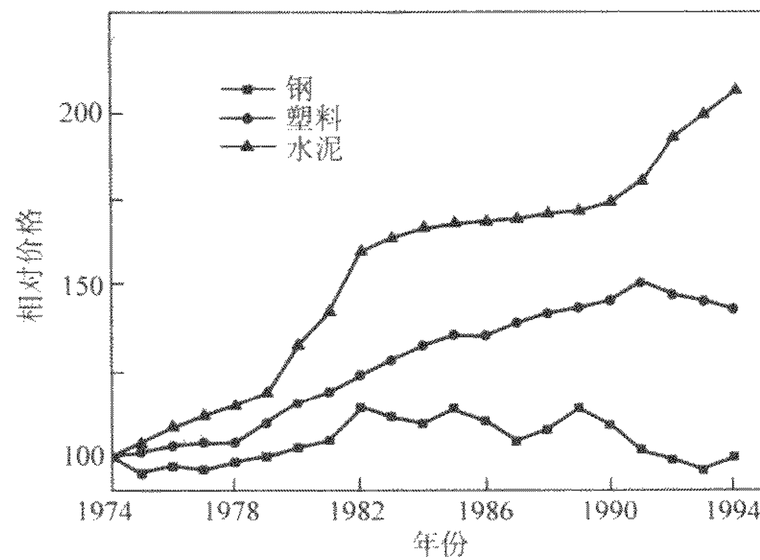


图 1.4-30 不同产品所需材料的相对价格

另外，钛合金的价格居高不下，主要归咎于产量太低，不能形成连续大批量生产线，生产效率不能提高，能耗降不下来，如图 1.4-32。

另外，价格相差不大，但习惯势力或其他因素影响了材料的推广应用，如用天然气做为汽车燃料有利于环保，应大力提倡。其中的关键需要有压力高达 20 MPa 的高压气瓶，

在英国采用的 Cr－Mo 合金钢，这种高压气瓶太重，从而降低汽车的有效载荷，如采用复合材料做衬里，可降低自重而又价格升高不多，如表 1.4-7，应该说高压容器使用高档复合材料也是有竞争能力的，但限于某些原因，由钢气瓶改为复合材料气瓶的阻力还是很大的。

图 1.4-31 材料用量与价格间的关系

① 1 lb ≈ 0.45 kg

图 1.4-32 一些金属与价格的关系

表 1.4-7 天然气高压容器几种材料对比

容器材料	相对质量/%	安全因子	1 加仑①柴油投入/美元
全钢	100	2.25	80
E－玻璃纤维增强铝衬里	50	3.5	135
E－玻璃纤维增强热固塑料	48	3.5	100
碳纤维增强热固塑料	24	3.5	122

① 1 加仑 = 3.785dm^3。

(3) 材料的质量与质量控制

材料质量对机械产品有决定性作用。不同产品对材料质量要求不同，主要体现在所规定的材料标准之中，包括国家标准、行业标准和企业标准。

标准是根据标准差来确定的。当材料或零件在大批量生产时，逐个检查费时费力，有时甚至是不可能的。一般采取抽样检查，根据检查结果来定是否可以投产或出厂。其计算方法如下：

取 n 个样品求其均值，即 $\overline{X} = (X_1 + X_2 + \cdots + X_n)/n$，而后再求标准差（$\sigma$），

$$\sigma_i = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

根据统计理论，频率分布的标准差为：

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

据此可以绘出零件的正态分布及在不同标准差（σ）范围内的误差分数，如图 1.4-33。

图 1.4-33 产品质量的正态分布

在正常情况下，工业控制限为 3σ，即产品中有 0.27% 不合格品，这是可以接受的。但对高技术产品来说，合格率为 99.73% 实在太低。如美国阿波罗飞船有 360 万个零件，如果按照 3σ 标准，就允许有近万个零件不合格；印刷电路有 1 万个焊点，允许 27 个为虚焊，这显然是不可行的。据此，1986 年摩托罗拉公司的 M.J.Harey 提出了 6σ 质量管理的概念，即产品质量合格品出现概率为 99.999 999 8%，即 100 万件中有 0.002 件不合格品，产品不合格率在 10^{-9} 量级。

6σ 质量管理不仅要求产品质量远离缺陷，而且成为一种全新的企业管理理念和管理模式，成为一种全新的企业思维和文化，包括决策、生产流程、质量分析和控制、产品改进与售后服务等，不但要求准确，而且要高速度。美国好多大公司都推行了 6σ 管理战略，以提高竞争能力。如通用电器公司（General Electric Co.）是世界上最大的航空发动机及工业燃气轮机制造商之一。这两种动力装置中最关键材料是涡轮叶片及涡轮盘用的高温合金。在 6σ 管理制度下，GE 对于用在高速转动部件用的材料采取了十分严格措施，首先在原材料的选择上予以高度重视，因为有些有害微量元素在冶炼过程中无法去除，所以在冶炼工艺上采取三联，即先用真空感应炉熔炼以调整成分和去气，铸成电极后再经电渣重熔以去除非金属夹杂物，最后再经过真空自耗重熔，以调整合金锭的组织结构，并减少偏析。这种复杂的冶炼程序主要为了保证合金质量，使最后制造出来的零件符合 6σ 管理标准。

6σ 管理的代价是巨大的，特别对材料的要求完全无缺陷在理论上是不可能的，因而材料存在缺陷也是正常的，关键是在装置工作寿命期内，缺陷是否发展到使装置达到失效的程度。特别自 20 世纪 80 年代引入损伤容限设计以后，材料在工况条件下的裂纹扩展速率是可测定的，因此，利用断裂力学原理可以保证安全使用。这就要求对材料的无损检测要准确、精密、及时。

目前常用的探伤方法很多，如荧光法、磁粉法及涡流法用于探测表面缺陷的存在，X 射线（γ 射线）及超声探伤法可用于判定材料内部缺陷的存在。材料在局部塑性形变发生弹性波，这种弹性波可用传感器接受下来，表现为声发射，如果几个传感器合理分布，可以定出缺陷发生的位置。前一类无损探伤可以定量地显示已存在缺陷的分布、大小和取向；而声发射说明已有缺陷发展的过程。但是这些都不能预测在什么部位会发生裂纹。中科院金属研究所经过了 40 年的研究与开发，采用红外技术可以确定发生裂纹的可能部位。材料在形变过程中

会产生热量而使温度升高，用红外仪监控温度的变化便可察知薄弱环节之所在。具体作法是压力容器在承受疲劳载荷下，薄弱部位因弹性形变而产生“热斑”（heating scar）。试验证明，产生热斑部位首先发生裂纹。用这种技术，成功地用于“神舟”号飞船高压氧容器质量控制，也为美国监控过输气管道。复合材料增强剂的断裂监控是一难以解决的问题，采用红外技术，有可能得到解决；用于叶片质量控制也有发展前景。

(4) 材料的供应问题（availability）。

材料是机械制造的基础，材料的供应便是机械制造的关键。首先遇到的是资源问题，资源往往是限制材料发展因素之一。

中国民用飞机开发公司董万钧对飞机材料由于缺陷在不同阶段所造成的损失有如下论述：如在草图阶段发生问题为1时，则生产阶段为10，终检阶段为100，出厂使用阶段为1 000，产品改型则为10 000，如因缺陷造成飞机失事，则损失难以估计，可能高达100 000乃至数百万倍。这充分说明航空材料质量控制的重要性。但是对材料质量的要求因产品而异，必须考虑经济性、现实性和必要性。对结构材料来说，没有缺陷的材料是不存在的，关键在使用条件下，缺陷的发展速度是否在寿命期内足以发展到失效的程度，只要在使用寿命期内裂纹的扩展在一定安全系数范围内，缺陷还是允许存在的。因此对材料质量的要求和加强检测之间有一个折衷的选择，如图1.4-34所示，当材料质量要求十分严格时，生产成本大为增加，对检测的投入可以少一些；而当对材料质量不太严格时，生产成本可以减少，但要加强检测，以保证符合标准所规定的要求，则检测设备与人力的投入大为增加。二者之间存在一个折衷选择（trade-off），使成本降到最低。当然对实行6σ管理时，重点在保证生产质量，而不能在很多合格与不合格品中挑选合格品，因为在混有不合格品中挑选合格品难免有漏检的机会。但是前者成本过高，所以一般在二者之间采用折衷选择。

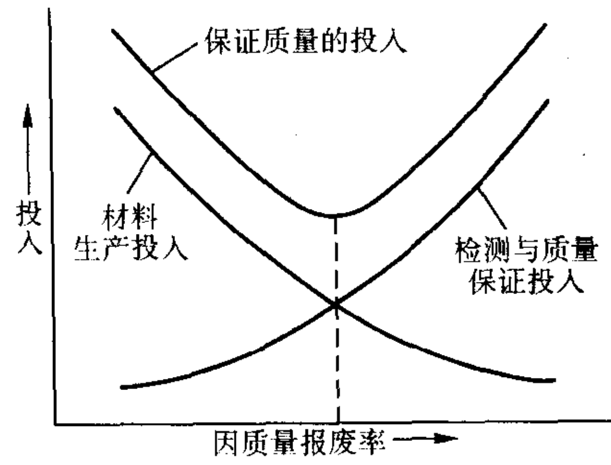


图1.4-34 材料生产投入与检测投入的折衷选择

材料中含有大量稀有元素，不但制约了材料的供应，而且价格也会上扬，因此只有在十分必要的情况下，才会采用。如定向凝固的高温合金中，一般加入铪以解决柱状晶间开裂问题；为了提高高温合金高温强度，最近开发出铼含量很高的合金。但是研究工作者都在努力改变这种状态而设法减少或免除这些稀有金属元素的用量。

采用市场上已大量供应、成熟的材料，还是采用专门材料，是目前机械制造业的一个重要考虑的问题。前者价格比较低廉、质量稳定、供应渠道有保障。但是新材料有更好的性能或功能，通过新材料的选用，可使机械产品的功能明显提高，因而对高技术机械产品来说，采用新材料有很强的诱惑力，但也要注意到新材料的研制周期长、投入大、而且有很大风险。一般来说，一种新材料的开发，从问题的提出到概念设计，到实验室验证，扩大试验到产业化需要很长时间，少则几年，多则十几年；有的因工艺问题难以克服而夭折，即使新材料达到批量生产的程度，材料用户也存在很多顾虑，除了价格昂贵以外，新材料的用途往往有限，生产量小，从而产品质量稳定不下来。另外，新材料生产厂家往往出于一个企业，有可能出现垄断现象而受制于人，因而新材料生产定点不应局限于一个企业。使用新材料一个更突出的问题是材料的供应连续性，谨防一种材料仅用于一种机械产品、一旦机械产品换代，材料随之消亡，这样损失太大，因而在现有材料的基础上进行改进比较稳妥可靠，因为工艺比较成熟，质量容易稳定，许多材料也是这样发展起来的。

(5) 材料与可持续发展

这是近几十年才提出来的一个问题，在机械制造中必须加以考虑。这是由于资源日益枯竭、环境日益恶化的结果。表1.4-8为地球表壳金属矿可开采的年数，可以看出，有些金属可供开采的年数，有的只有几十年，如铜、铅、锌等，最长的时间也不过几百年。

表1.4-8 地球表壳金属矿产资源可供开采的年数

矿石	储量		可开采年数
Fe	矿石/10^9 t	153.4	171
Mn	金属/10^6 t	907.2	97
Cr	富矿（$Cr_2O_3$35%～45%）/10^6 t	1 057	106
Ni	金属/10^6 t	52.6	68
W	金属/10^6 t	2.8	60
Mo	金属/10^6 t	5.44	55
V	金属/10^6 t	4.35	135
Cu	金属/10^6 t	340	40
Pb	金属/10^6 t	95	26
Zn	金属/10^6 t	170	26
Al	铝矾土/10^9 t	21	238
Ti	金属/10^6 t	17.2	53
Sn	金属/10^6 t	3.06	16
Au	金属/10^3 t	39.8	25
Pt	金属/10^3 t	31.1	135
Li	金属/10^6 t	1.9	487
Nb	金属/10^6 t	4.08	310
稀土	氧化物/10^6 t	45	781

对有机材料来说，目前主要以石油或天然气为原料，其可供年限也只有几十年，最长不过百年，但是有机物的组成主要为C、N、H、O，在宇宙取之不尽，用之不竭，它们可取之于生物体，资源可再生，但要加强技术开发。有些陶瓷材料，如Si_3N_4、SiC等的资源问题不太严重，但多数无机非金属材料的原料也不是取之不尽，用之不竭的，如我国水泥原料可持续开采的年限也不过四五十年。

为了有效利用资源，必须采取以下措施：首先是提高矿石的开采率和选矿技术。目前一般有色金属的开采回收率很低，从30%到70%，约半数残留废矿之中，再加选矿回收率不高而造成很大浪费。其次是提高材料的利用率，表1.4-5表示我国材料加工利用率与国外差距很大，有很大发展空间。第三是提高材料的性能，如新一代钢通过晶粒细化，改进显微结构，可使低碳钢及低合金钢的强度提高一倍。第四是实行环境协调性评估，即全寿命循环评估（life-cycle-assessment）（LCA），这在“机械制造与生态设计”案例中详述。

近年来，又提出再制造工程的概念。再制造工程是利用

原有零件并采用再制造技术，使零部件恢复尺寸、形状和性能，形成再制造的产品。再制造与维修不同，维修是在产品的使用阶段为了保持其良好技术状况及正常运行而采取的技术措施，具有随机性、原位性和应急性。再制造则是将大量相似的报废产品返回工厂拆卸后，按零部件的类型进行收集和检测，以之作为再制造毛坯，利用高新技术进行批量化修复，使性能升级，从而可靠性和寿命通过再制造而恢复或提高，达到节约原材料和加工工时的目的。这已成为当前的一种新兴产业，图 1.4-35 为再制造工程在产品全寿命周期中的位置。

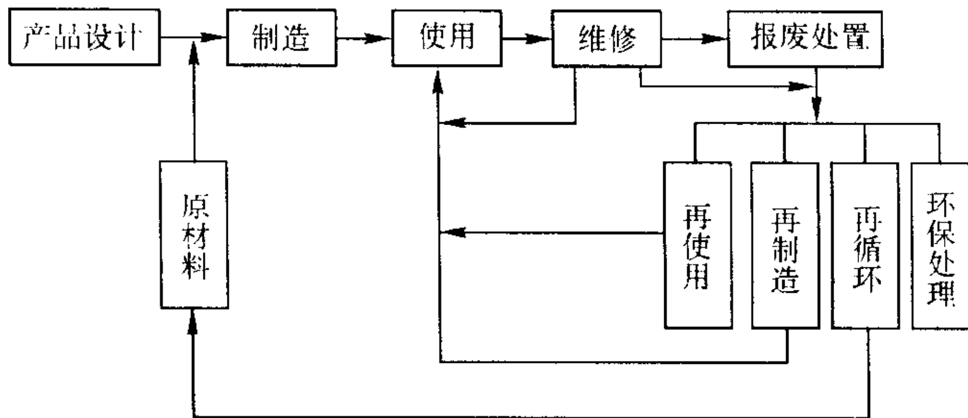


图 1.4-35　再制造工程在产品全寿命周期中的位置

不管维修或再制造工程，材料和工艺都是关键，特别是热喷涂、激光处理、离子镀、堆焊等技术，形成了当前十分活跃的表面工程学科体系。表面强化、表面改性是节约资源、降低成本的最主要手段。因为很多零件的损毁都始于表面，如疲劳裂纹的萌生、摩擦磨损、腐蚀与氧化。因此，即使是新生产的零部件、刀具或模具，也都进行表面处理，以延长服役寿命，增加精确度。对在高温工作的部件来说，采用热障涂层可提高使用温度 150℃左右。

3　在机械设计选材过程中容易出现的几个误区

1）要重视材料的综合性能，切忌单一追求高强度指标　在机械设计选材时往往喜欢追求高强度指标，认为强度高则安全系数高，孰不知一般来说，材料的强度愈高，裂纹扩展率也愈高，一旦出现裂纹，或构件表面已存在缺陷，在受力条件下，裂纹很快扩展而导致断裂。相反，强度低的材料对裂纹敏感性差反而安全可靠。而且高强度钢的氢脆敏感度大，因而钢中氢容许量因强度的提高而锐减，高含氢量容易发生滞后断裂。

2）要重视在接近使用条件下材料及构件的力学性能　实验室的材料或构件的性能往往与实际情况相去太远，如果不重视使用环境的条件便很难得出正确的选材。图 1.4-36 是不同强度、不同环境和不同状态的钢在疲劳应力下表现的示意图。

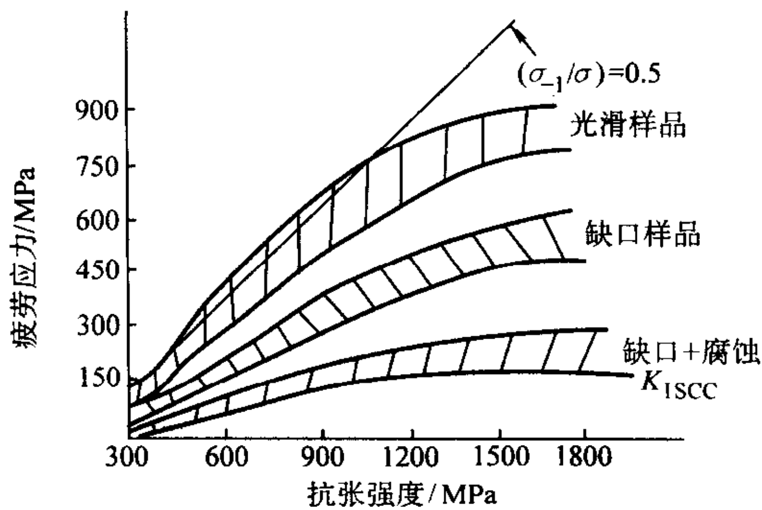


图 1.4-36　钢在不同受力状态和介质条件下的强度示意图

可以看出，钢的强度级别不同，缺口的影响也不同，腐蚀介质的作用也不一样，但是总的趋势是强度愈高，缺口敏感性愈严重，腐蚀介质的作用愈明显。如在缺口与腐蚀介质的共同作用下，效果就更突出。

此外，材料在不同疲劳载荷谱的作用下，都会有不同的表现，因为材料会发生锻炼效应而寿命有所提升，也可能因受到过度损伤而加速破坏，因此，材料或零部件只有在接近使用条件下的试验才可以得到正确结果。又材料在疲劳与蠕变的交互作用下，会加速其破坏过程，因此，单项测试数据不能作为设计的惟一依据。

再一个容易忽略的问题是不能用小样品测试的数据来做为设计大体积、宽厚结构的依据，因为二者的力学性质差异很大，特别是塑性和裂纹扩展率。

上述问题都是在材料工程化过程中急待解决的问题。

3）机械产品质量的减小主要决定于材料密度的减少，而强度、模量的影响较小　在设计过程中往往有一种错误的估计，即强度提高 1 倍，机械产品的质量可减小一半。根据力学计算，材料强度或模量的提高，只能部分降低产品的质量，如图 1.4-37。所以为了降低产品的质量，重要的是选择密度小的材料而不能主要依赖强度的提高。

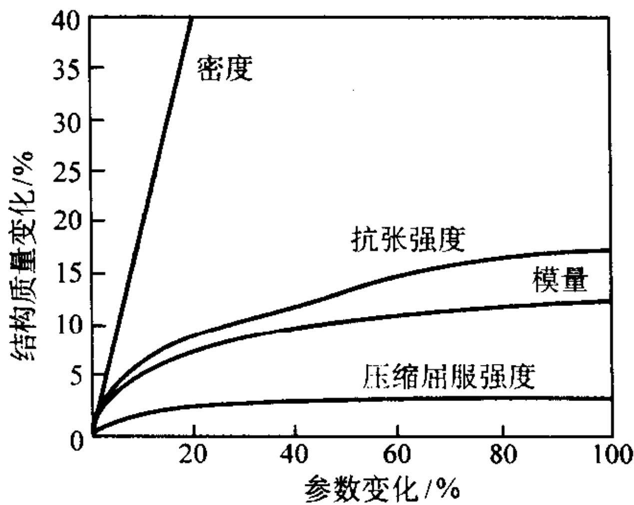


图 1.4-37　材料性能提高对结构质量的影响

4）要高度重视部件的应力集中因子　机械制造离不开孔、沟、槽及粗细过渡等应力集中部位，部件的断裂往往就从这些地方开始。表 1.4-9 为飞机机翼不同零件发生断裂的次数，可以看出，凡是带螺纹或孔的零件都容易发生断裂，这就是应力集中所致。

表 1.4-9　机翼不同零件发生断裂的次数

零件名称	发生断裂次数
螺栓、螺杆、螺钉	108
紧固件孔	72
轴肩、尖缺口	57
焊缝	53
腐蚀	43
螺纹	32
微振磨损	13

螺纹、螺栓、孔和轴肩都是应力集中之所在，它们占了所有破坏部件的70%以上，由此说明应力集中系数的选择十分重要。解决的办法一是曲率半径尽量加大，另一种办法是使圆弧部位产生压应力，如喷丸、渗碳、渗氮等，最重要的是选用塑性好，无缺口敏感的材料。应该指出，用光滑样品测出的塑性相同的材料，其缺口敏感性可能不相同，要通过实验来确定，这就要求设计人员与材料工作者的密切合作。

5）机械设计人员与材料工作者要密切合作，才能做到合理选材 设计人员往往根据材料的已有数据进行设计，而材料数据往往与其生产流程和微观结构密切相关，材料工作者只有严格把关，才能保证数据的可靠性和重复性，使设计人员所采用的数据切实可靠。另外，材料的某些弱点可以通过设计人员的优化设计来克服，如结构陶瓷的脆性问题至今仍难以大幅度地改善，陶瓷学家寄希望于设计人员合作，改变目前的设计方法和理念，采取措施使陶瓷能广泛应用于机械产品，利用陶瓷耐高温、耐磨损、膨胀系数小等特点，使机械制造更上一层楼。

4 案例分析

（1）关于汽车的选材

汽车在很多国家都列为支柱产业，因为数量大、带动性强，与人们的生活息息相关。表1.4-4说明三种运载工具每减重1 lb所取得的收益。可见在这三种工具中，每减重1 lb，汽车所得收益远大于商用飞机或飞船，因为汽车量大面广，所以汽车的减重必须受到重视。因为能源消耗减少，污染也就减轻了。具体讲，车身材料强度提高80%，车身质量可降低25%，每百公里油耗降低3 L。此外，对汽车设计来说，还必须考虑安全可靠，价格低廉，提高自动化与智能化程度，长寿命、可回收和可修理以及舒适度等。

汽车材料涉及面很广，据统计，美国汽车所用材料占其国内材料用量的比例很大如表1.4-10，可见汽车对材料的依靠性及汽车产业对材料产业的促进作用。

表1.4-10 汽车用材料占美国材料百分比

材料	占美国材料消耗量/%	
	1984年	1988年
钢	17.0	12.2
铝	16.0	8.3
铜	9.7	10.3
铅	59.0	67.3
铂	46.9	39.1
锌	26.0	23.0
合成橡胶	55.1	50.1
天然橡胶	76.6	76.6

从材料角度出发，汽车可分为三个部分：一是车体，二是动力装置，包括冷却和电力系统，三是控制系统。随着汽车档次的提高，控制系统所占比重越来越大，即功能材料的用量越来越多，品种也越来越多。表1.4-11为美国车体材料的变化。

从表1.4-11看出：车体材料主要是钢，而且高强钢用量增加速度大于铝，复合材料及工程塑料用量大幅度增加；陶瓷材料及先进复合材料刚刚开始，但近10年来碳纤维强化的先进复合材料的增加速度加快。

表1.4-11 美国车体材料的变化情况

项目	年份				年增长率/%	
	1970	1985	1990	1995	1985～1995	1970～1985
汽车销量/10^3 辆	6 547	8 002	8 200	8 400	1.3	0.5
平均质量/kg	1 585	1 375	1 312	1 270	-0.9	-0.8
材料用量/10^6·lb①						
高强度钢	758	785	1 085	1 398	11.3	5.9
不锈钢	37	106	114	120	7.3	1.3
铝	216	468	486	564	5.3	0.7
复合材料	-	33	136	232	48.0	21.8
先进复合材料	—	—	—	2	—	—
工程塑料	12	99	179	243	15.0	8.4
一般塑料	294	616	766	931	5.1	4.2
陶瓷	—	—	1	20	—	47.1

① 1 lb = 0.45 kg。

对客车来说，主要采用高强度钢以保证安全和减重。另外，车体前后部件要能吸收冲击能量，车门结构也要加强，以防侧面的碰撞。各部分连接一般采用点焊。整体做好后还要上漆，一般要在130～180℃烘烤15～30 min，钢的性能不能有明显改变。

车体钢的发展已有几代。最初采用加铝（0.08%）的软钢，以防止明显屈服平台的出现而避免制耳，其强度在300 MPa左右，以后又出现低合金结构钢，强度在400 MPa上下；到20世纪70年代，加磷（0.08%）钢板问世，其强度仍为400 MPa，但深冲等综合性能较好，我国在80年代初，鞍钢试生产成功，并在一汽推广。以后又加入少量合金元素生产马氏体与铁素体共存的双相钢。这几种钢板的深冲性能和强度示如图1.4-38。进入90年代最流行的是无间隙原子钢（IF钢），所谓无间隙原子是指碳和氮的含量低于0.005%，如此则钢具有非常好的成形性，因其屈服强度低，深冲性能好。如加入少量铌或钛后，可细化晶粒，提高强度，并进一步改善深冲性能。

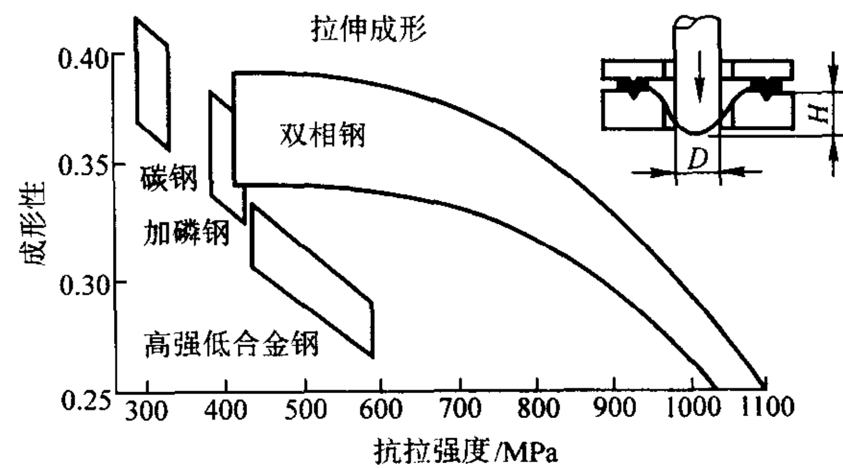


图1.4-38 几种汽车用钢的可成形性

进入21世纪，我国研制的“超级钢”，使碳钢的强度从200 MPa，低合金钢从400 MPa提高了1倍，作为汽车车体材料正在进行试验。

为了减轻车体重量，塑料及铝、镁等轻合金的用量逐年增加，碳纤维增强塑料也开始应用。用于汽车的工程塑料品种很多，除密度小、抗腐蚀以外，还不易出现压痕，但塑料部件尺寸要求严格，否则装配困难，特别是塑料回收困难，不符合环保要求，但最近把废弃塑料放入高炉炼铁，也算找到了出路。铝的密度只有钢的35%，看来很有竞争能力，但铝的价格在诸材料中算是最高的，如镀锌钢板每公斤约

0.70 美元，工程塑料约 1.75 美元，而铝接近 4.0 美元。所以有人估计铝车体的出现要等汽油的价格提高 1 倍，因采用铝所减重的收益才能抵得上昂贵材料铝的选用。镁的密度更小，且有良好的消振性能，镁在汽车中的应用方兴未艾，但目前用量还不多，一辆车不过几公斤，且都是压铸件。一方面是由于价格高，另一方面镁大量用于汽车还要解决腐蚀与防护、提高强度及成形性能等问题。无论如何，镁是地球表层惟一取之不尽用之不竭的金属结构材料，今后必将有更大发展。

车体采用复合材料也可能是一个选择方案。复合材料有两类：一是在聚酯中加入 25% 的短纤维或有取向的长纤维，一般用于汽车面板，另一类是碳纤维增强塑料。加入增强纤维以后，一方面提高强度，更重要的是提高模量，或增加吸收能量的能力。表 1.4-12 和表 1.4-13 分别列出不同复合材料的模量变化及吸收能量的能力。

表 1.4-12　不同复合材料的弹性模量

材　　料	模量/GPa
取向碳纤维增强	137.8
取向芳纶增强	75.8
取向玻璃纤维增强	31.0
50% 短玻璃纤维增强	15.8

表 1.4-13　不同材料吸收能量的能力

材　　料	单位质量相对吸收能量
高性能复合材料	100
一般复合材料	60 ~ 75
软钢	40

可以看出，以碳纤维增强的先进复合材料性能都远远高于玻璃纤维增强塑料。因此，如果用钢板做的汽车重 1 705 kg，用碳纤维增强复合材料做的汽车则只有 1 136 kg，减少了 569 kg。复合材料用于汽车的部位示如图 1.4-39。

显然用先进复合材料制造的车体减重明显，性能大为改善，但是由于材料价格昂贵，制造难度大，又不能回收，不可能用以制造有竞争能力的汽车，除非有特殊用途的汽车，如赛车，或用于汽车的某些关键部位；特别是玻璃增强的复合材料，价格低廉（每千克 1.80 美元），已大量用于汽车制造业。

汽车的动力系统对汽车的功能和节油都具有决定的作用。动力系统的气缸一般采用可锻铸铁，内衬耐磨金属，如 SiC 颗粒增强铝基复合材料。为了提高动力系统的热效率，美、日等国都曾研制出陶瓷绝热发动机。我国在 20 世纪 80 年代也曾列为重点攻关项目，研制成功后曾装车来往于北京

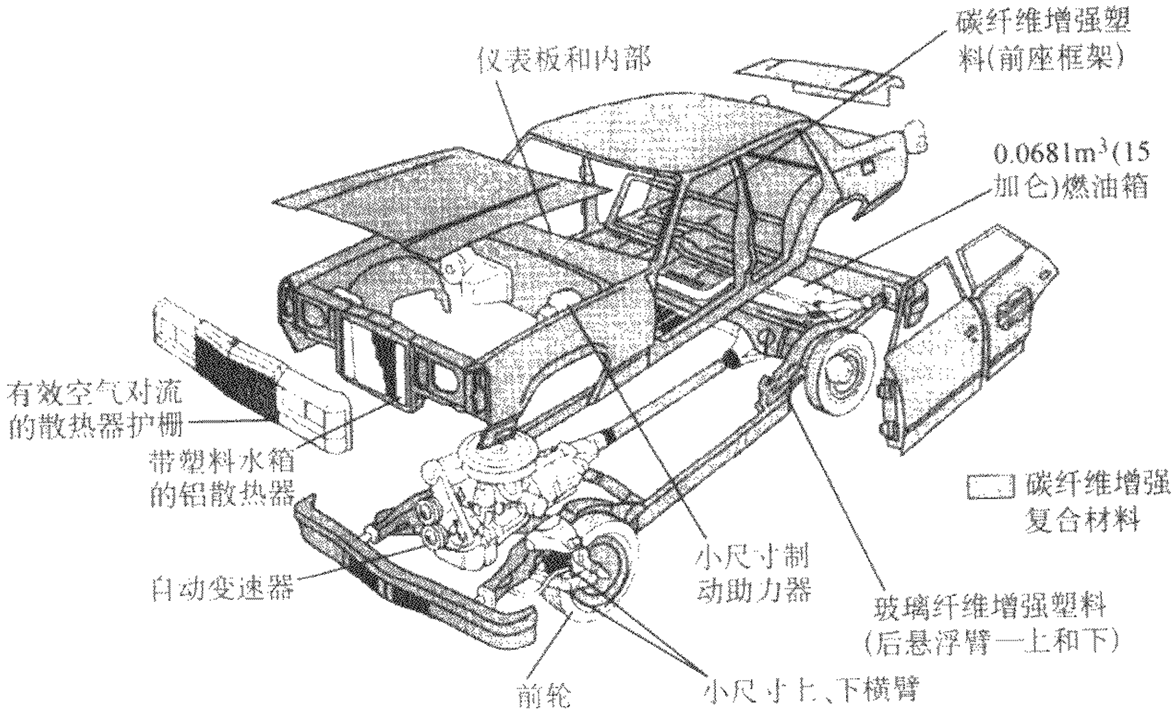


图 1.4-39　全复合材料汽车的分解示意图

与上海之间，但是价昂，回收困难，不可能有竞争能力而中断。由于陶瓷发动机不需水冷，用于沙漠地区有一定的优势。

陶瓷作为发动机零件却有明显优势，如图 1.4-40 为气缸中活塞所用阀门采用 Si_3N_4 材料后的两组特性曲线。可以看出，采用陶瓷阀门后，噪声和传动转矩都较金属阀门低。因此，可节油 5% 左右。

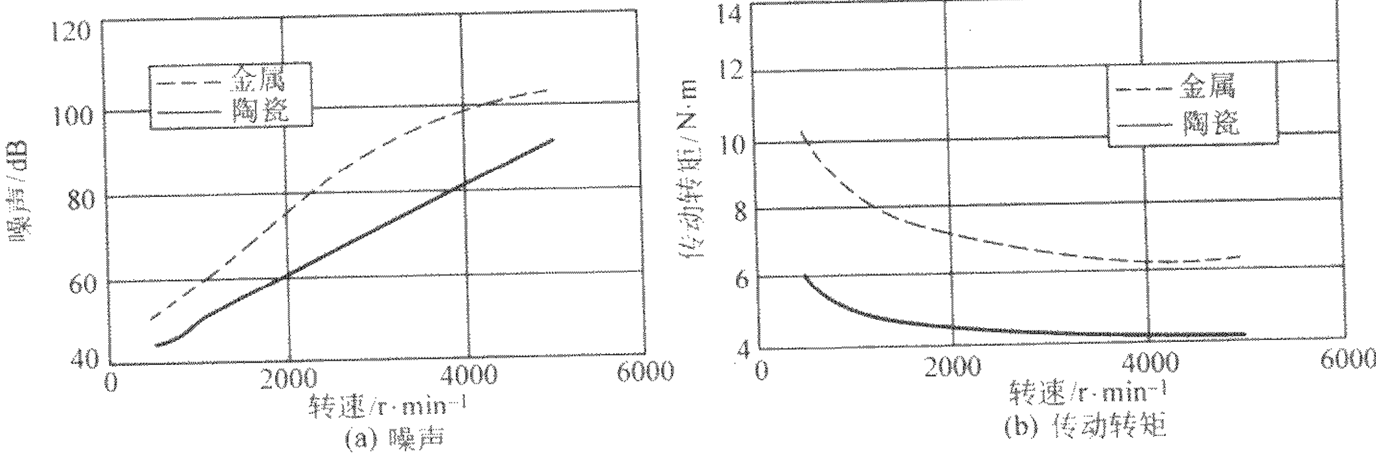


图 1.4-40　采用陶瓷阀门与金属阀门的特性对比

总之，汽车的选材比一般机械制造的选材要给以更高的重视，因为汽车生产量大，涉及人身安全，尽管微小的变化，可能造成严重后果；特别要重视生态设计，部件寿命要适当，易回收，易修理，为汽车再制造工程创造条件。

(2) 飞机机体结构设计及铝合金选材

先进铝合金是航空航天工业中的主要结构材料，它不仅具有高比强度、高比模量、良好的断裂韧性和疲劳强度以及较低的裂纹扩展速率，同时还具有优良的成形工艺性和良好的耐蚀性。在最新的民用飞机中，铝合金的用量仍要占到70%～80%。在新一代军用飞机上，由于树脂基复合材料和钛合金用量的增加，常规铝合金的用量相对有所减少，但高纯、高韧、高强的高性能铝合金用量却增加了。对于航天工业来说，铝合金是运载火箭的主体结构材料，主要用于推进剂储箱、尾段和箱间段及级间段等部位，在今后的相当长时间内，铝合金仍将是运载火箭、宇宙飞船、空间站等航天器的主体结构材料，也是导弹等武器系统的重要结构材料。

现以飞机机体结构设计和铝合金选材为案例来进行分析，看材料在航空制造业中的重要地位以及材料与飞机结构设计与工艺及应用之间的密切关系。从20世纪初开始有动力飞行的飞机以来到20世纪50年代，飞机基本上是按静强度的准则来设计，主要追求的目标是减重，最主要的材料性能要求只是强度。铝合金在航空上的应用最早可以追溯到20世纪初莱特兄弟用Al－Cu－Mn铸造铝合金制造发动机的曲柄箱体。早在1903年德国人Wilm就Al－Cu－Mg系合金进行了研究，并于1906年发现了铝合金的沉淀硬化现象，由此开发出第一个可热处理强化的铝合金，从而揭开了铝合金飞速发展的序幕。1920年德国率先开发出2017－T4合金，并应用于JunkerF13机身结构上。20世纪30年代美国ALCOA公司开发出2014－T6、2024－T4以及2024－T3铝合金，并将这些合金应用于DC－13飞机中。由于当时飞机设计主要追求高强度，从而开始了Al－Zn－Mg系合金的研究，但后来由于该系合金存在严重的应力腐蚀现象而未能得到广泛应用。20世纪40年代初人们发现Cu可以大大提高Al－Zn－Mg系合金的强度并可改善该系合金的耐蚀性，而Mn等元素则可显著改善Al－Zn－Mg系合金的抗应力腐蚀和抗剥落腐蚀性能。这样在二次大战期间成功地研发出Al－Zn－Mg－Cu（7000系合金）的高强铝合金7075－T651，并应用于B29大型轰炸机上，随后又研发成功更高强度的7178－T651合金并应用于B－707飞机上。1954年因3架应用7075合金制造机体的喷气式客机坠毁，从而使航空界认识到单一追求材料强度的危险性。20世纪60年代以来，飞机越来越大型化，在飞机上开始使用厚大截面结构，因此出现了机体结构的严重应力腐蚀问题。2024－T3和7075－T6这两种大量应用的合金均存在应力腐蚀开裂倾向。为了解决7000系铝合金的应力腐蚀问题，研发成功了7075－T73工艺，该工艺虽然解决了7075合金的抗应力腐蚀问题，但却使合金的静强度损失了15%左右。随后，又研发成功牺牲较小强度而应力腐蚀性能可以满足要求的7075－T76工艺，从而使7075合金更广泛地应用于DC－10和L－1011等飞机上。在20世纪70年代以前，航空铝合金的主力军是2024和7075合金，这些合金迄今仍然是许多现役飞机的主要机体用材。近30余年来，为了满足飞机失效安全设计的要求，又研发成功了7475合金，又称损伤容限型合金，广泛应用于F16战斗机和一些民用飞机上。1971年ALCOA公司在美国海军的支持下又研发成功7050－T4合金，该合金在F18舰载飞机上获得了应用。同一时间内欧洲研发成功7010合金，被空中客车等大型客机和军机大量采用。20世纪70年代后期，为满足波音B－757和B－767飞机的设计要求，ALCOA公司又研发成功7150－T6合金，为满足提高抗腐蚀性能的要求，研发成功T61热处理工艺，已应用于B－757、B－767、A－310和MD－11等民用客机上。进入20世纪80年代，7150－T77热处理工艺研发成功，它第一次在铝合金中实现了既满足腐蚀性能要求而同时又不牺牲合金强度。随后通过调整合金成分和降低Fe、Si等杂质元素含量和改进生产工艺等手段，ALCOA终于研发成功7055－T77超高强铝合金板材和2524－T3超高疲劳强度板材。7055－T77合金的强度比7075－T651高25%，比7150－T77高8%～12%，而断裂韧性和抗应力腐蚀性能则与7150－T77相当。2524－T3铝板的断裂韧性和抗疲劳性能均高于2024－T3铝板30%以上，被认为是当今最为理想的飞机蒙皮材料。这两种新合金在波音公司B－777飞机上首先应用。上述飞机设计、制造工艺和选材应用的发展历程可以综合归纳为图1.4-41、图1.4-42和表1.4-14、表1.4-15。波音公司B－777飞机上先进铝合金及其他材料的选用情况见图1.4-43。

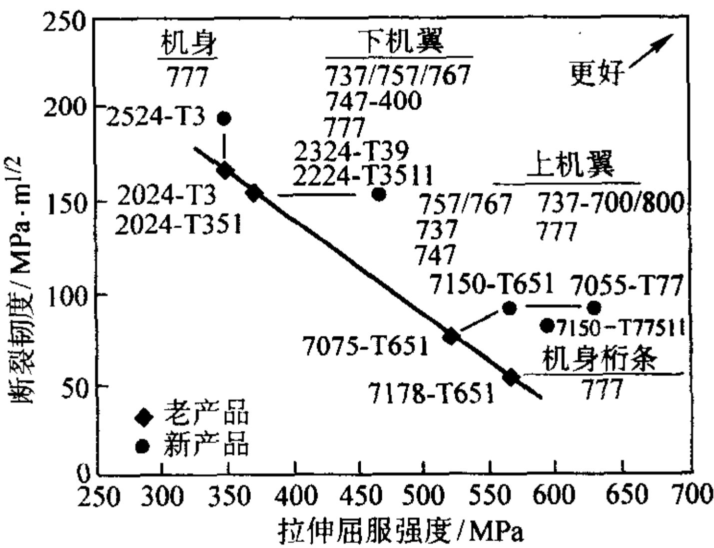


图1.4-41　先进铝合金在飞机上的应用部位及其性能改进

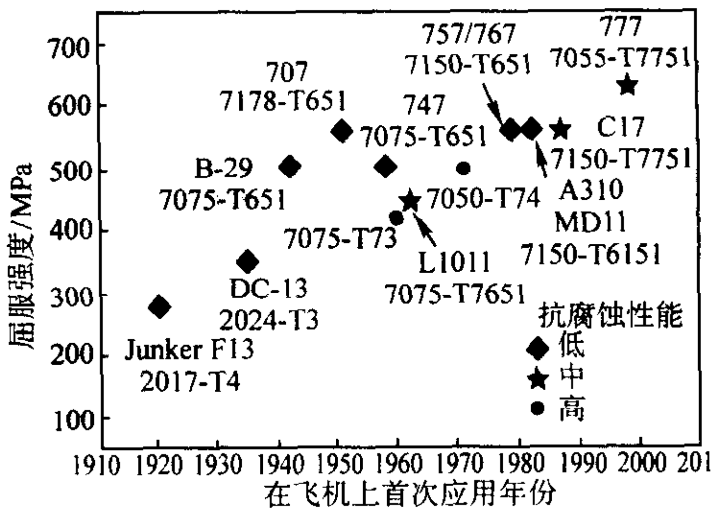


图1.4-42　航空铝合金的发展简史

表1.4-14　飞机机体结构设计及对铝合金性能要求

年份	飞机设计使用需求	对铝合金性能要求
1903年至20世纪50年代	静强度设计，减重	强度
60年代	抗腐蚀性能需求，疲劳	抗应力腐蚀，抗剥离腐蚀，疲劳
70年代	综合性能要求，损伤容限	强度，抗腐蚀性能，损伤容限，高温性能
80年代	强烈的减重需求	比刚度、比强度
90年代	降低制造和维修成本的需求	强度、韧度、疲劳、可成形性能、可维修性能

表 1.4-15　飞机结构选材

飞机结构部位	关键性能要求	典型铝合金选用
上机翼	抗压强度，刚度，损伤容限	7075-T6，7150-T6，7055-T77
下机翼	损伤容限，疲劳	2024-T3，7475-T73
机翼梁、肋	疲劳，抗剪强度	2024-T3，7010-T76，7150-T77
机身蒙皮	疲劳，损伤容限，腐蚀	2024-T3，7075-T6，7475-T6，2524-T3
机身桁条	疲劳，抗压强度	7075-T6，7075-T73，7475-T76，7150-T77
机身隔框	刚度，疲劳，抗压强度	2024-T3，7075-T6，7050-T6

(3) 航空涡轮盘强度设计及高温热强合金选材

航空发动机的涡轮盘是关键的热端转动部件，其工作条件十分苛刻，在高温、高转速、高气流速度及压力下工作。它是航空发动机结构完整性大纲中至关重要的部件，其结构强度设计、选材、制造工艺和性能检测相互关系十分紧密，并对航空发动机在服役中的可靠性和寿命具有重要的影响。

军用航空燃气涡轮发动机通常以推重比（推力/质量）来评定发动机的水平，而涡轮前燃气温度对发动机的推重比有最直接、最显著的影响。随着发动机水平的不断提高，其选用的涡轮前温度也是不断提高，如图 1.4-44 所示。

因此，涡轮前温度是将发动机"划代"时最重要的判别依据，如表 1.4-16 所示。从表可见相应的选材级别也不断提高，从一般变形高温合金到粉末高温合金、纤维增强超合金、金属间化合物乃至陶瓷和碳/碳复合材料。

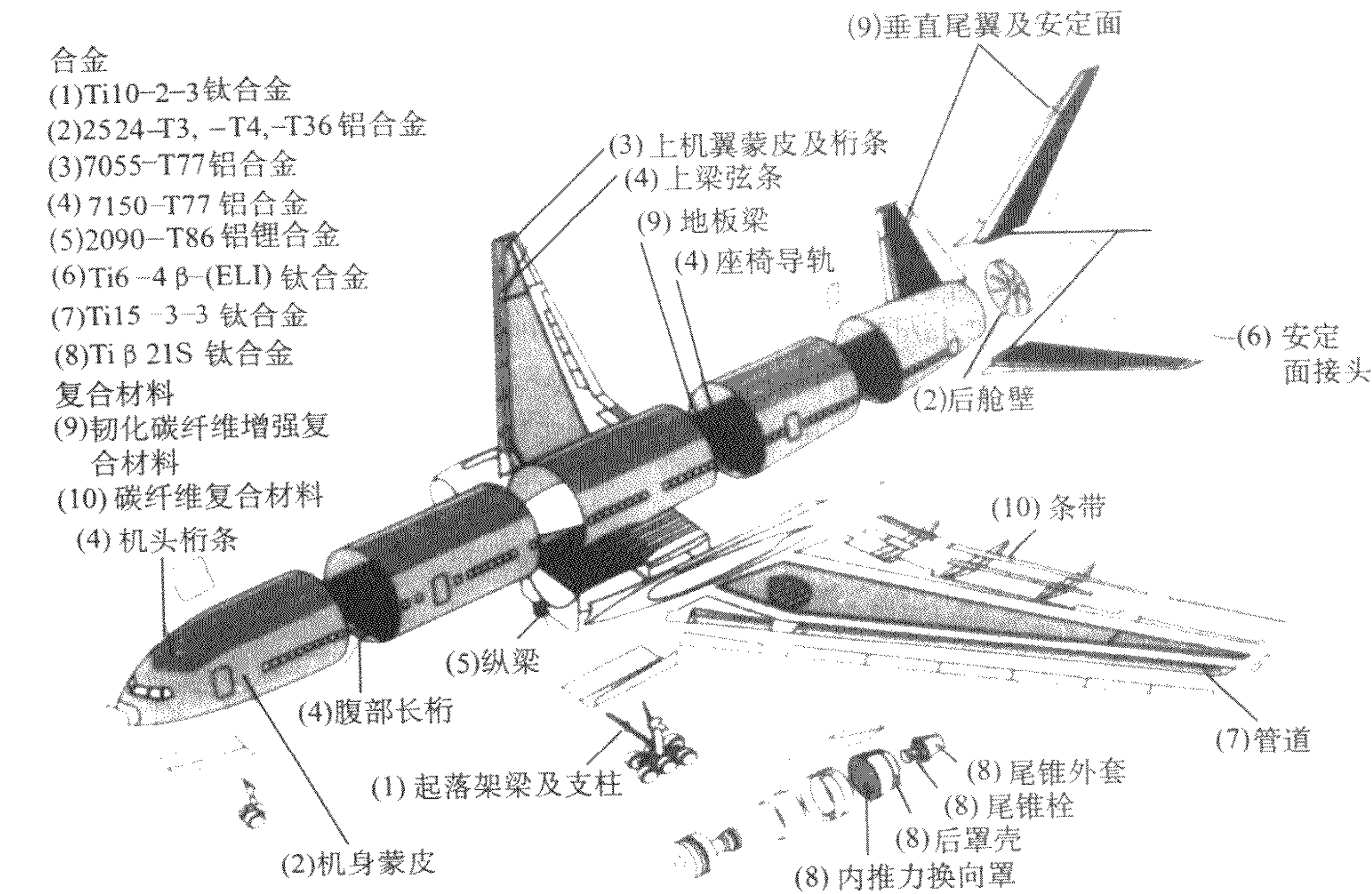


图 1.4-43　B-777 飞机上先进铝合金及其他材料选用情况示意图

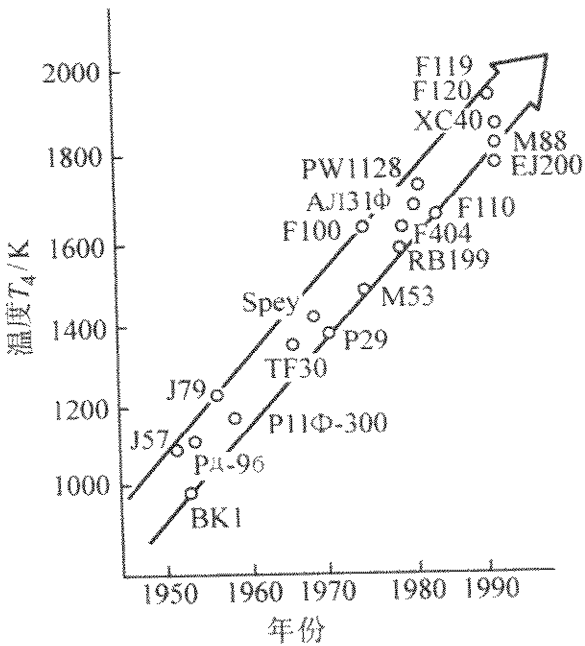


图 1.4-44　涡轮前温度逐年增加趋势

由于涡轮前温度已超出工程用高温合金材料的承受能力，对涡轮部件一般采取冷却措施。涡轮盘在采取隔热措施后，其工作温度一般相当于或略低于压气机出口温度。在对涡轮部件的冷却过程中，伴随着能量损失，使发动机推力减少，耗油率增大，所以涡轮前温度越高，冷却空气用量就越大，能量损失也就增大。因此，从设计角度来看，总是希望能有承受更高温度的材料可供选用，以降低冷却空气用量，提高发动机的循环效率。涡轮盘在高温工作环境下承受高的轮缘载荷，这是由涡轮叶片的离心负荷引起的，且本身在高

表 1.4-16　各代发动机的涡轮工作环境条件①

机型/代	Ⅱ	Ⅲ	Ⅳ	Ⅴ
推重比	5 ~ 6	7 ~ 8	9 ~ 10	15 ~ 20
涡轮前温度/K	1 250 ~ 1 400	1 600 ~ 1 750	1 850 ~ 1 950	2 250 ~ 2 350
压气机出口温度/K	650 ~ 700	800 ~ 900	850 ~ 950	950 ~ 1 200
涡轮盘工作温度/℃	600	650	700	780 ~ 950
材料	变形高温合金	变形高温合金 粉末冶金高温合金	粉末冶金高温合金 纤维增强超合金 变形高温合金	双性能粉末冶金高温合金 纤维增强超合金 TiAl 合金 陶瓷和 C/C （耐 1 093 ~ 1371℃）

① 表中为国外典型军用发动机数据，民用发动机的压气机出口温度更高。

速下旋转。提高涡轮转速可以减少涡轮级数，但在设计时，涡轮转速的提高主要受涡轮叶片和涡轮盘材料强度水平的限制。涡轮盘的轮缘载荷大小可以用表征涡轮叶片根部应力的 AN^2 来衡量，其中 A 为涡轮转子通道面积，cm^2，N 为涡轮转速，r/min。推重比 8 以下发动机通常 AN^2 不超过 25；推重比 10 发动机，AN^2 已高达 30 以上，这意味着涡轮盘轮缘载荷将增大 20%以上。鉴于发动机涡轮盘的研制成本很高，故要求其具有长寿命，军用发动机涡轮盘寿命要求达到 2 000～4 000 h 以上，而国外先进民用发动机涡轮盘寿命则要求达到 10 000 h 以上。

综上所述，对涡轮盘材料的基本要求可以归纳为：高许用温度，包括在高温使用下的长期稳定性；高的比强度和比持久强度；低裂纹扩展速率；高蠕变极限；优良的疲劳性能和低缺口敏感性；较低的材料成本和制造及维修成本。随着发动机总体设计技术的进步，涡轮盘强度设计技术也有了很大进步，从 20 世纪 60 年代早期的静强度估算方法，当时对材料性能的要求主要是抗拉屈服强度和模量等物理常数；后来进入安全寿命设计，不仅要考虑强度力学载荷，要求提供 $S-N$ 疲劳曲线，还要考虑热分析技术的要求；20 世纪 70 年代以后进入采用应变疲劳理论进行有限寿命设计，此时要求材料能提供 $\varepsilon-N$ 应变疲劳曲线和热－机械疲劳性能；进入 20 世纪 80 年代以后，开始应用断裂力学理论和损伤容限设计，此时材料的性能数据中要包括 K_{IC} 和 da/dN 等数据；20 世纪 90 年代以来，进入到可靠度设计和概率寿命设计，此时对材料性能数据要有统计分析，包括提供 -3σ 数据等，同时对无损检测要求提供缺陷特征及其分布等数据。表 1.4-17 及表 1.4-18 所示为航空涡轮盘强度设计的不同年代和方法及其对材料性能数据的要求和变形高温合金、粉末高温合金和钛铝金属间化合物三类典型涡轮盘材料的比持久强度数据的对照表。

表 1.4-17　航空涡轮盘强度设计方法及所需材料数据

年代	强度设计方法	所需材料数据
20 世纪 60 年代初	静强度设计	E、μ、δ、α、σ_b、$\sigma_{0.2}$、$\sigma_{-1}{}^Q$、ψ、G、a_K
60 年代末	加上安全寿命、热分析、振动分析	加 $S-N$ 曲线、古德曼图、λ、c、E_d、$\sigma_\varepsilon{}^Q$、σ_D
70 年代	加上应变疲劳理论、有限寿命设计	加 $\epsilon-N$ 曲线、热机械疲劳
80 年代	加上断裂力学理论、损伤容限设计	加断裂性能、K_{IC}、J_E、da/dN、ΔK_{th}、da/dt、K_{ISCC}、缺口敏感性、$q_t=\sigma_b/\sigma_{bH}$
90 年代至今	加上可靠度设计、概率寿命设计	加 -3σ 数据、缺陷特征及分布

表 1.4-18　典型航空涡轮盘材料性能对比

性能＼合金	GH4133 (ЭИ437Б)	YZGH4169 (In718)	FGH95 (Rene' 95)	FGH96 (Rene' 88-DT)	TiAl
σ_ε/ρ	87.45	104.6	127	117.6	127
σ_b/ρ	124	121.4	150.4	168.1	180
$\sigma_{50\,h}^{650℃}/\rho$	75.3	87.4	124.4	123.6	$\sigma_{100\,h}^{750℃}/\rho$ 54.5

美国“综合高性能涡轮发动机技术（IHPTET）”（1988 年启动，计划 15 年完成）计划中，发动机结构评估对材料研制途径和强度设计和试验间的综合关系见图 1.4-45，从中可见材料、制造、检测和设计之间的有机关系。

(4) 机械制造与生态设计

目前资源逐渐枯竭，温室效应愈来愈明显，污染愈来愈严重，生态环境问题愈来愈突出，因此，早在 20 世纪 70 年代，罗马俱乐部就指出：人口爆炸式的增加，资源枯竭与环境破坏，必将制约经济的增长。此后，就提出了全生命周期评价方法（life cycle assessment，LCA）以监控资源与能源的有效利用程度及污染情况。最近又提出生态设计（ecodesign）以保证资源的有效利用和最大限度地减少污染，最终形成一个可持续发展的社会。

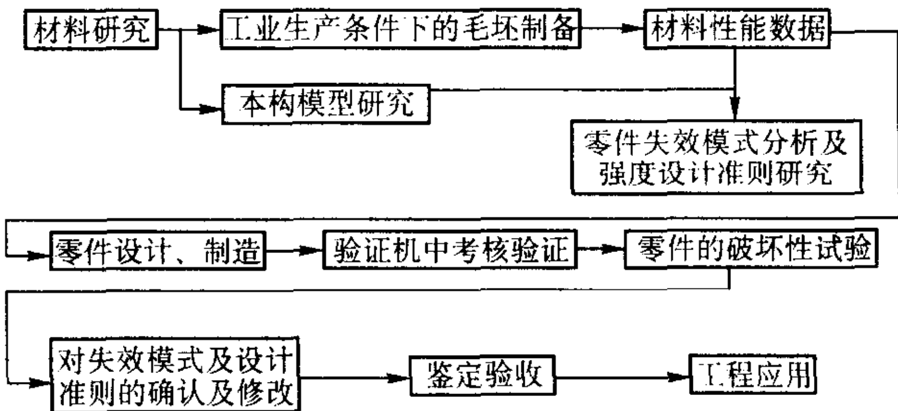


图 1.4-45　材料研制和鉴定验收过程示意图

生态设计基于全生命周期评价的结果，而后达到减量（reduce）、重复使用（reuse）、再生利用（recycle），即 3R 的目标。

生态设计有三个要素，即成本（cost）（C）、环境影响（impact）（I）、效能（performance）（P）。

$$\text{产品的综合价值指标}=\frac{\text{效能}\ (P)}{\text{成本}\ (C)\times\text{环境影响}\ (I)}$$

就是生态设计的产品必须是使 P 趋于最大，I 和 C 趋于最小。

下面是关于住房空调机的全寿命周期的评价方法。

空调机全寿命周期评价可分为 5 个阶段，即材料生产、加工、运输到户、使用及报废处理，如图 1.4-46。

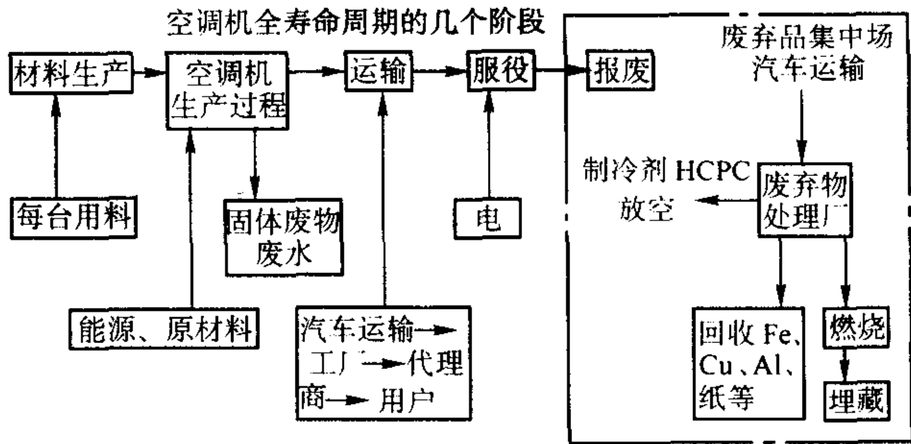


图 1.4-46　产品全寿命周期示意图

在全寿命周期内气体的排放量见表 1.4-19。

表 1.4-19　在全寿命周期内各种气体排放量/%

气体	材料生产	空调机生产	运输	服役	报废
CO_2	1.4	0.2	<0.01	99.2	0.75
NO_x	22.6	1.5	0.34	87.0	11.5
SO_2	12.3	0.14	<0.01	89.1	1.6

从上表说明，生产节能型空调是减少有害气体排放量的最主要环节，其次就是空调机的选材问题了。

生态设计的目标在于提高资源的利用率和减少对环境的冲击。提高资源效率的途径很多，实行非物质经济和加强管理是重要途径之一。目前将资源使用效率作为一个指标，如使其资源的绝对消费量降至 1/4，则称之为因子 4（factor4）。为达到这一目的，可以通过使产品性能的提高，寿命的延长，以及资源回收，再利用，再加资源用量的减少以达到因子 4 的目标。人们把因子 4 作为发达国家最近达到的目标；与此同时，还有人提出因子 10 或因子 20 作为长远目标，如到 2050 年要实现因子 4 则环境负荷可减少 75%，而实现因

子10环境负荷则可减少90%。

实现生态设计，第一要设计者、生产者乃至全社会观念改变，把节约资源，减少污染放在重要位置，甚至首要位置。第二在设计完成后，在材料的选择上对材料中有害元素的含量要严格把关，把它们的含量限制在标准规定的范围以内，如铅、镉、汞、铬（六价），否则便不可能符合绿色生产标准。第三则考虑生产过程中的环境效率（节约材料、节约能源），从而生产出环境友好产品（nature-friendly products）。日本很多公司都定下环境友好产品的发展目标，如日立公司2000年为29种，2001年为243种，2002年为568种，分别占其产品的3%、22%和46%。第四，尽量利用信息技术进行管理，从而达到节约资源和能源的目的，如利用信息技术（IT）进行采购比常规办法可节约80%能源消耗。

总之，生态设计是一种新兴的设计思想，随着信息化程度的提高，资源与能源的枯竭与社会发展水平的提高，生态设计在机械制造业中必将有更大发展。

编写：师昌绪（国家自然科学基金委员会）
李成功（中国材料研究学会）

参考文献

1 C. Razim . Potential and Limitation of Materials Development and Application in the Automotive Industry, Frontiers in Materials Science and Engineering. Edited by S. Somigya etc, Elsevier 1993, 691 ~ 699

2 C.A.Vincent, Batteries, An Overview. The Encyclopedia of Advanced Materials. edited by D. Bloor, R. J. Brook, M. C Flemings, S. Mahajan, Pergamon (1994), 214 ~ 229

3 杨遇春. 电动汽车和相关电源材料的现状与前景. 中国工程科学, 5 (12), 2003, 1 ~ 11

4 周本廉, 陈立泉. 电池材料. 材料科学与工程国际前沿, 李恒德主编. 济南: 山东科技出版社, 2002

5 李成功, 傅恒志, 于翘. 航空航天材料. 北京: 国防工业出版社, 2002

6 Material Research to Meet 21 Century Defense Needs, National Research Council (National Academies Press) (Washington D, C), 2003, 27 ~ 54

7 M. F. Ashby. Materials Selection in Mechanical Design. The design process, 1999, 8 ~ 19

8 W. Boecker, H. Kruener. Silicon Carbide and Silicon Nitride Ceramics for High Performance Structurals Application, Advanced Materials (1994), Frontiers in Materials Science and Engineering (1994), Elsevier 865 ~ 872

9 J. Banhart. Aluminum Foams: On the Road to real Applications, MRS Bulletin (April 2003), 290 ~ 295

10 Material Research to Meet 21 Century Defense Needs, National Research Council (National Academies Press) (Washington D, C), 2003, 247 ~ 250

11 A.D.Batte . The Implementation of Composite Materials technology in an Engineering Environment. Advanced materials in Masketplace, edited by J. E. Caslle and M. J. Kelly, Institute of Materials, 1995, 105 ~ 117

12 M. E. Zohdl. Statistical Quality Control. Mechanical Engineerings Handbook. edited by M. kutz, 1998, 1175 ~ 1185

13 T. N. Goh. The Role of Statistical Design of Experiments in Six Sigma: Perspectives of a practitioner, Quality Engineering, 2002, 14 (4): 659 ~ 671

14 山本良一. 生态设计. 王天民译. 北京: 化学工业出版社, 2003

15 徐滨士. 再制造工程咨询报告, 徐滨士院士科研文选. 北京: 机械工业出版社, 2001

16 E. Aghion, B. Brobfin. Magnesium alloy development towards the 21 Century, Materials Science Forum, 2000, 19: 250 ~ 251

17 S. J. Findlay, N. D. Harrison. Why aircraft fail. Materials Today, 2002, 11: 18 ~ 25

18 F. R. Field and J. P. Clark. Automotive Body Materials, the Encyclopedia of Advanced Materials, edited by B. Bloor, R. Brook, M. C. Flemings, S. Mahajan, Pergamon 1998, 183 ~ 191

19 Automotive Application of Composite, the Encyclopedia of Advanced Materials, edited by B. Bloor, R. Brook, M. C. Flemings, S. Mahajan, Dergamon 1994, 173 ~ 183

20 李成功, 巫世杰, 戴圣龙, 杨守杰. 先进铝合金在航空航天工业中的应用与发展. 中国有色金属学报, 2002, 3 (12): 14 ~ 21

21 江和甫. 对涡轮盘材料的需求和展望. 航空材料及其构件成形技术研讨会文集, 北京: 中国锻压协会, 2003, 4 ~ 9

22 K. Yanagilani et. LCA Study Air Conditioners. Proc. of the 3 rd International Conference on Ecomaterials (1997, Tsukuba, Japan): 351 ~ 354

23 M. Hirano. Environmental Activities of Hitachi Group, A Report on China-Japan Symposium on Eco-Materials, Recycling-Oriented Industry and Environmental Management, (2004, 4 held in Suzhou, China)

24 冯端, 师昌绪, 刘治国. 材料科学导论—融贯的论述. 北京: 化学工业出版社, 2002

25 黄昆, 韩汝琦. 固体物理学. 北京: 高等教育出版社, 1988

26 师昌绪等主编. 材料科学技术百科全书. 北京: 中国大百科全书出版社, 1995

27 师昌绪, 李恒德, 周廉主编. 材料科学与工程手册. 北京: 化学工业出版社, 2004

28 薛增泉. 纳米科技探索. 北京: 清华大学出版社, 2002

29 张立德. 纳米材料. 北京: 化学工业出版社, 2001

30 李恒德, 师昌绪. 中国材料发展现状及迈入新世纪对策. 济南: 山东科技出版社, 2002

31 周廉等. 有色金属材料咨询报告. 西安陕西科学技术出版社, 2000

32 干福熹. 信息材料. 天津: 天津大学出版社, 2000

33 吴人洁. 复合材料. 天津: 天津大学出版社, 2000

34 雷永泉等. 新能源材料. 天津: 天津大学出版社, 2000

35 王天民. 生态环境材料. 天津: 天津大学出版社, 2000

36 俞耀庭, 张兴栋. 生物医用材料. 天津: 天津大学出版社, 2000

37 徐冠华等. 中国高新技术产业发展报告. 北京: 科学出版社, 1999

38 冯端等. 金属物理学: 第 1 卷, 第 2 卷. 北京: 科学出版社, 1990

第2篇

材料成形基础理论

主　编　董湘怀
编　写　董湘怀　吴树森　魏伯康
　　　　周华民　李　赞
审　稿　中国材料工程大典编委会

中国机械工程学会
中国材料研究学会
中国材料工程大典编委会

第 1 章　材料成形的冶金学原理

1　材料成形中的流体流动

1.1　基本概念

1.1.1　作用于流体的力和应力

(1) 质量力和表面力

作用于某一流体块或微元体积的力可分为两大类：表面力、质量力或者体积力。

1) 质量力　直接作用在流体块中各质点上的非接触力，如重力、惯性力等。质量力与受力流体的质量成正比，也叫体积力。单位质量流体上承受的质量力称单位质量力 F_m。当取直角坐标系时，F_m 在坐标轴向的分量记为：X、Y 和 Z。在重力作用下，当 z 轴垂直向上时，有

$$Z = -g \tag{2.1-1}$$

2) 表面力　作用于流体块外界面的力，如压力、固体壁面对流体的作用力和切应力等。表面力与流体的表面积成正比。

(2) 应力

单位面积上的表面力，称为应力。流体应力的特性有：流体的应力不仅随时间和点的位置而变化，并且与受力面的方向有关。受力面的应力可分解为切向应力和法向应力。

(3) 流体的压力、静压、动压和总压

1) 流体的压力　流体任一点 M 的三个法向应力的算术平均值，称为 M 点的流体压力 p，即

$$p = -\frac{1}{3}\ (p_{xx} + p_{yy} + p_{zz}) \tag{2.1-2}$$

式中，p_{xx}、p_{yy}和 p_{zz}的值分别随所取的坐标而变，但它们的算术平均值不随所取的坐标变化。

2) 静压　流场中任一点处的流体压力 p，又称为该点流体的静压。

3) 动压　流场中任一点处的单位体积流体所具有的动能$\frac{1}{2}\rho v^2$，称为该点流体的动压。

4) 总压　流速为零时的压力，称为流体的总压。对不可压缩流体来讲，流场中任一点处的静压与动压之和（$p + \frac{1}{2}\rho v^2$），即为该点流体的总压。

(4) 迹线、流线、流谱

1) 迹线　流体质点的运动轨迹，称为迹线。

2) 流线　在给定时刻 t，每一点都和流体质点的速度相切的线，称为流线。

在流场中，除奇点（包括驻点）外，流线不能相交。在定常流动中，流线和迹线重合。

3) 流谱　在流场中，流线族的几何图形，称为流谱。

流谱可描述流场的概貌（图 2.1-1）。流线各点处的切线方向就是该点流速的方向；流线的疏密表示流速的大小。对亚声速流动，同一流场中流线密的地方表示流速大；反之，流速小。

(5) 流量和平均流速

一个流管是由许多流束组成的，这些流束的流动参量并不一定相同，所以流管的流量应为

$$Q = \int_A v\mathrm{d}A$$

工程上引用平均速度$\bar{v}$的概念，根据流量相等的原则，

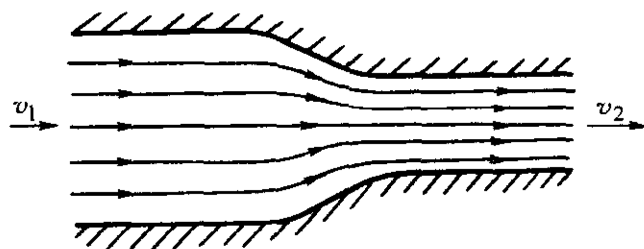


图 2.1-1　流谱

单位时间内匀速流过有效断面的流体体积应与实际流体通过同一断面的体积相等，即

$$\bar{v}\int_A \mathrm{d}A = \int_A v\mathrm{d}A = Q$$

则

$$\bar{v} = \frac{\int_A v\mathrm{d}A}{\int_A \mathrm{d}A} = \frac{Q}{A} \tag{2.1-3}$$

1.1.2　黏性流体和理想流体

(1) 流体的黏性和牛顿内摩擦定律

1) 流体的黏性　在作相对运动的两流体层的接触面上，存在阻碍两相邻流体层作相对运动的一对等值而反向的作用力，流体的这种性质叫做流体的黏性，由黏性产生的作用力叫做黏性阻力或内摩擦力。

2) 牛顿内摩擦定律　沿直线运动着的黏性流体，流体层间的内摩擦力 τ 的大小，与沿流体层法线方向（设为 y 方向）的速度梯度成正比，与流体的动力黏度 η 有关，而与正压力的大小无关。即

$$\tau = \eta\frac{\mathrm{d}v}{\mathrm{d}y} \tag{2.1-4}$$

动力黏度 η 的单位为 Pa·s。η 值越大，流体的黏性也越大。

在工程计算中也常采用流体的动力黏度与其密度的比，这个比值称为运动黏度，用 ν 表示，即

$$\nu = \frac{\eta}{\rho} \tag{2.1-5}$$

运动黏度 ν 的单位为 m^2/s。

(2) 非牛顿流体

不遵循牛顿黏性定律的流体，称之为非牛顿流体。常见的非牛顿流体有以下三类。

1) 宾海姆塑流型流体　宾海姆流体的切应力与速度梯度之间的关系为

$$\tau = \tau_0 + \eta\frac{\mathrm{d}v}{\mathrm{d}y} \tag{2.1-6}$$

在流变学等场合，常将稳定态下的速度梯度 $\mathrm{d}v/\mathrm{d}y$ 称为剪切速率，以 γ 表示。

如图 2.1-2 所示，要使这类流体流动，需要有一定的切应力 τ_0（塑变应力）。换言之，当切应力小于 τ_0 时，该流体处于固结状态；只有当切应力大于 τ_0 时才开始流动。例如，细粉煤泥浆、乳液、砂浆、矿浆等均属于这类流体。

2) 伪塑流型流体和胀流型流体　这类流体的特征方程为

$$\tau = \eta\ (\frac{\mathrm{d}v}{\mathrm{d}y})^n \tag{2.1-7}$$

式中，η 与 n 均为常数。当 $n < 1$ 时，为伪塑流型流体；当 $n > 1$ 时为胀流型流体。属于这类流体的有半固态金属液、石灰和水泥岩悬浮液等。

3) 屈服－伪塑流型流体　这类流体的特征方程为

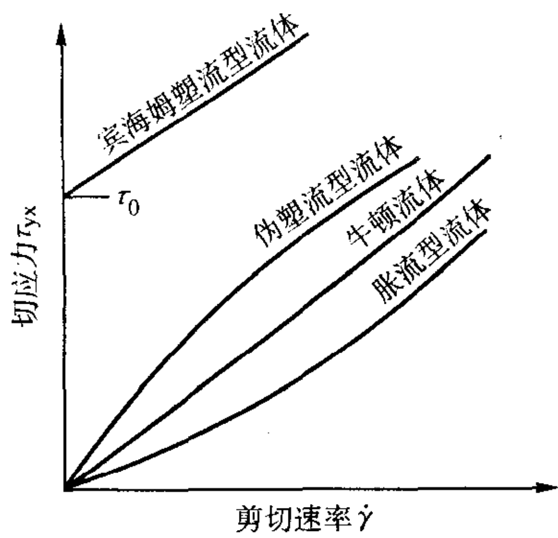


图 2.1-2 稳定流流体的切应力 - 剪切速率关系曲线

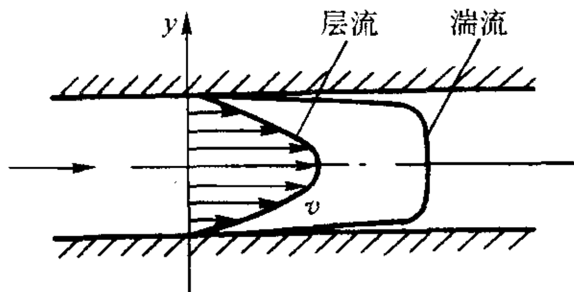


图 2.1-3 管内层流和湍流的速度分布

$$\tau=\tau_0+\eta\left(\frac{\mathrm{d}v}{\mathrm{d}y}\right)^n \tag{2.1-8}$$

这类流体与宾海姆塑流型流体相类似，但切应力与速度梯度之间的关系是非线性的。

此外，在研究半固态金属或铸造涂料时，会遇到在剪切速率固定不变的情况下，流体的切应力（τ）随切变运动时间的增加而减小的非牛顿流体，称为触变性流体。

1.1.3 可压缩流体和不可压缩流体

从是否考虑流体密度的变化出发，可将流体分为可压缩流体和不可压缩流体两类。

密度等于常数的流体为不可压缩流体；反之，为可压缩流体。

1）液体的压缩性远小于气体的压缩性。由于液体的压缩性很小，通常按不可压缩流体处理；若压力变化很大（如水锤问题），则必须按可压缩流体处理。

2）气体的压缩性虽然较大，但对压力及流速的影响均很小，从而所引起的密度变化也很小的问题，可按不可压缩流体来处理。

1.1.4 稳定流动和非稳定流动

如果流场的运动参数不仅随位置改变，并且随时间不同而变化，这种流动就称为非稳定流动（也称非定常流动）；如果运动参数只随位置改变而与时间无关，这种流动就称为稳定流动（也称定常流动）。

例如，对于非稳定流，流场中压力分布可表示为

$$p=p(x, y, z, t)$$

对于稳定流，上述参数表示为

$$p=p(x, y, z)$$

1.1.5 层流、湍流和雷诺数

（1）层流和湍流

从流体质点的运动是否紊乱出发，流动分层流和湍流两类。

1）层流　流体质点间相互不混杂、层次分明、平滑地流动，称为层流。

2）湍流　流体质点间相互混杂而无层次地流动，称为湍流（也称紊流）。

湍流时，由于流体质点的不规则运动，使空间任一点上的流速，无论其大小或方向都随时间而变化（其他流动参数也随时间而变化）。因此，湍流实质上总是非稳定流动。

3）速度分布　湍流时，过流断面上的速度分布比层流时均匀，但物面附近沿物面法线方向的速度变化$\partial v/\partial y$较层流时大（图 2.1-3）。

（2）雷诺数和临界雷诺数

1）雷诺数　雷诺数 Re 是一个表征流体惯性力与黏性力之比的量纲为 1 的数，即

$$Re=\frac{vl}{\nu} \tag{2.1-9}$$

式中，v 为流体的特征速度，在管流中取过流断面上的平均流速；l 为流体的特征长度，在管流中常取水力直径 d_s，对圆管内的流动即为管子内径 d；ν 为运动黏度。

若雷诺数很小，则相对于惯性力来说，黏性力是主要的，黏性对流动的影响，在整个流场中都是重要的。若雷诺数很大，则惯性力是主要的，黏性对流动的影响只有在边界层内或速度梯度较大的流动区域内才是重要的。

2）临界雷诺数

① 临界雷诺数　层流和湍流相互转化时的雷诺数，称为临界雷诺数。由层流转变为湍流时，为上临界雷诺数；反之，为下临界雷诺数。上临界雷诺数的值不稳定，变化范围大；下临界雷诺数的值基本不变。

② 层流和湍流的判别　实用上，以下临界雷诺数 Re_c 来判别流动状态是层流还是湍流。当 $Re>Re_c$ 时，流动为湍流；当 $Re<Re_c$ 时，流动为层流。

几种流动情况的临界雷诺数值见表 2.1-1。

表 2.1-1 几种流动情况的临界雷诺数 Re_c 的值

流动情况	临界雷诺数值
在内径为 d 的圆断面管内的流动	$Re_c=\frac{Vd}{v}\approx 2\,300$
在水力直径为 d_s 的任意断面管内的流动	$Re_c=\frac{Vd_s}{v}\approx 2\,000$
在缝宽为 δ 的环形或平板缝隙内的流动	$Re_c=\frac{V\delta}{v}\approx 1\,000$

工程实际中，多数为湍流。层流主要存在于缝隙流动和黏性大、流速小的流动中。

3）水力直径

① 湿周　过流断面 A 上被流体浸润的固体壁面周线，称为湿周 x。

② 水力半径　过流断面面积与湿周之比，称为水力半径 R_s。即

$$R_s=\frac{A}{x} \tag{2.1-10}$$

③ 水力直径　4 倍过流断面面积与湿周之比，称为水力直径（当量直径）d_s。即

$$d_s=4\frac{A}{x}=4R_s \tag{2.1-11}$$

1.2 基本方程

1.2.1 连续性方程

流体的连续性方程为

$$\frac{\partial\rho}{\partial t}+\frac{\partial(\rho v_x)}{\partial x}+\frac{\partial(\rho v_y)}{\partial y}+\frac{\partial(\rho v_z)}{\partial z}=0 \tag{2.1-12}$$

其物理意义是：流体在单位时间内流经单位体积空间输出与输入的质量差与其内部质量变化的代数和为零。这个方程实际上是质量守恒定律在流体力学中的具体体现。

对于可压缩性流体稳定流动，$\frac{\partial\rho}{\partial t}=0$，于是，式（2.1-

2）变为

$$\frac{\partial(\rho v_x)}{\partial x}+\frac{\partial(\rho v_y)}{\partial y}+\frac{\partial(\rho v_z)}{\partial z}=0 \qquad (2.1\text{-}13)$$

对于不可压缩流体，$\rho=$常数，则式（2.1-13）成为

$$\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}+\frac{\partial v_z}{\partial z}=0 \qquad (2.1\text{-}14)$$

对于一维流动的不可压缩流体，即 $\rho=$常数，当流体流经断面积分别为 A_1 和 A_2 的两个断面时，连续性方程可以写为

$$v_1A_1=v_2A_2$$

$$\frac{v_1}{v_2}=\frac{A_2}{A_1} \qquad (2.1\text{-}15)$$

式（2.1-15）为一维总流不可压缩流体稳定流动的连续性方程。它表明在稳定流动条件下一维总流沿流程体积流量保持不变，为一常值；各有效断面平均流速沿流程的变化规律是：平均流速与有效断面面积成反比，即断面大流速小，断面小流速大。

1.2.2　运动方程

（1）欧拉方程——理想流体的动量平衡方程

欧拉方程是从流体运动的动量守恒定律（牛顿第二定律 $F=ma$）导出的

$$\left.\begin{aligned}X-\frac{1}{\rho}\frac{\partial p}{\partial x}=\frac{\mathrm{d}v_x}{\mathrm{d}t}\\ Y-\frac{1}{\rho}\frac{\partial p}{\partial y}=\frac{\mathrm{d}v_y}{\mathrm{d}t}\\ Z-\frac{1}{\rho}\frac{\partial p}{\partial z}=\frac{\mathrm{d}v_z}{\mathrm{d}t}\end{aligned}\right\} \qquad (2.1\text{-}16)$$

若用向量表示，则为

$$W-\frac{1}{\rho}\nabla p=\frac{Dv}{Dt} \qquad (2.1\text{-}17)$$

式中，W 为质量力，$W=Xi+Yj+Zk$；∇p 为压力梯度，$\nabla p=\partial p/\partial x+\partial p/\partial y+\partial p/\partial z$，有时写成 grad p。请注意，压力本身是个标量，而压力梯度却是向量。向量和标量之积仍是向量；$\frac{Dv}{Dt}$为实质导数，即加速度。

欧拉方程（Euler equation）建立了作用在理想流体上的力与流体运动加速度之间的关系。它是研究理想流体各种运动规律的基础，对可压缩及不可压缩理想流体的稳定流或非稳定流都是适用的。

（2）奈维－斯托克斯方程——实际流体动量传输方程

$$\left.\begin{aligned}\frac{\mathrm{d}v_x}{\mathrm{d}t}=X-\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\left(\frac{\partial^2 v_x}{\partial x^2}+\frac{\partial^2 v_x}{\partial y^2}+\frac{\partial^2 v_x}{\partial z^2}\right)\\ \frac{\mathrm{d}v_y}{\mathrm{d}t}=Y-\frac{1}{\rho}\frac{\partial p}{\partial y}+\nu\left(\frac{\partial^2 v_y}{\partial x^2}+\frac{\partial^2 v_y}{\partial y^2}+\frac{\partial^2 v_y}{\partial z^2}\right)\\ \frac{\mathrm{d}v_z}{\mathrm{d}t}=Z-\frac{1}{\rho}\frac{\partial p}{\partial z}+\nu\left(\frac{\partial^2 v_z}{\partial x^2}+\frac{\partial^2 v_z}{\partial y^2}+\frac{\partial^2 v_z}{\partial z^2}\right)\end{aligned}\right\} \qquad (2.1\text{-}18)$$

应用拉普拉斯（Laplacian）运算子$\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$，并用向量符号$\frac{Dv}{Dt}$表示 v 对 t 的三个实质导数，则上式可改写为

$$\left.\begin{aligned}\frac{\mathrm{d}v_x}{\mathrm{d}t}=X-\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\nabla^2 v_x\\ \frac{\mathrm{d}v_y}{\mathrm{d}t}=Y-\frac{1}{\rho}\frac{\partial p}{\partial y}+\nu\nabla^2 v_y\\ \frac{\mathrm{d}v_z}{\mathrm{d}t}=Z-\frac{1}{\rho}\frac{\partial p}{\partial z}+\nu\nabla^2 v_z\end{aligned}\right\} \qquad (2.1\text{-}19)$$

或

$$\frac{Dv}{Dt}=W-\frac{1}{\rho}+\nu\nabla^2 v$$

这就是实际流体的动量守恒方程，也即不可压缩黏性流体的动量传输方程，又称奈维－斯托克斯方程式（Navier－Stokes equation）（也被称作 N－S 方程）。

1.2.3　伯努利方程

（1）理想流体的伯努利方程

对于只有重力场作用下的稳定流动、理想的不可压缩流体沿流线任意两点 1 和 2 之间对运动方程式积分，得

$$z_1+\frac{p_1}{\gamma}+\frac{v_1^2}{2g}=z_2+\frac{p_2}{\gamma}+\frac{v_2^2}{2g} \qquad (2.1\text{-}20)$$

式中，γ 为重度，N/m³，$\gamma=\rho g$。

上式即为理想流体的伯努利方程式。理想流体的伯努利方程是动量守恒方程（欧拉方程）在一定条件下的积分形式，它表述了运动流体所具有的能量以及各种能量之间的转换规律，是流体动力学的重要方程之一。

（2）实际流体的伯努利方程

单位质量的黏性流体沿流线从点 1 到点 2 的路程上由于黏性所产生的摩擦阻力（或摩擦阻力损失）为 h'_w，实际流体的伯努利方程为

$$z_1+\frac{p_1}{\gamma}+\frac{v_1^2}{2g}=z_2+\frac{p_2}{\gamma}+\frac{v_2^2}{2g}+h'_w \qquad (2.1\text{-}21)$$

或写成：

$$z_1g+\frac{p_1}{\rho}+\frac{v_1^2}{2}=z_2g+\frac{p_2}{\rho}+\frac{v_2^2}{2}+\frac{h'_w}{g} \qquad (2.1\text{-}22)$$

（3）实际流体总流的伯努利方程

实际流体（黏性流体）经流管或流道流动时，设 h_w 为通过流道截面 1 与 2 之间的距离时，单位质量流体的平均能量损失，则实际流体总流的伯努利方程为：

$$z_1+\frac{p_1}{\gamma}+\alpha_1\frac{v_1^2}{2g}=z_2+\frac{p_2}{\gamma}+\alpha_2\frac{v_2^2}{2g}+h_w \qquad (2.1\text{-}23)$$

或写成：

$$z_1g+\frac{p_1}{\rho}+\alpha_1\frac{v_1^2}{2}=z_2g+\frac{p_2}{\rho}+\alpha_2\frac{v_2^2}{2}+\frac{h_w}{g} \qquad (2.1\text{-}24)$$

式中，α_1，α_2 为动能修正系数。在一般工程中，大多数情况下，流速都比较均匀，α 在 1.05～1.10 之间。所以，在工程计算中，可取 $\alpha=1$。

（4）伯努利方程的几何意义和物理意义

1）几何意义　z 是指流体质点流经给定点时所具有的位置高度，称为位置水头，简称位头，z 的量纲是长度的量纲。$\frac{p}{\gamma}$是指流体质点流经给定点的压强高度（受到压力 p 而能上升的高度），称为压头，$\frac{p}{\gamma}$的量纲也是长度的量纲。$\frac{v^2}{2g}$的量纲为$\left[\frac{v^2}{2g}\right]=\frac{L^2T^{-2}}{LT^{-2}}=L$，也是长度的量纲，它表示流体质点流经给定点时，以速度 v 向上喷射时所能达到的高度，称为速度水头。

伯努利方程中位置水头、压力水头、速度水头三者之和称为总水头，用 H 表示：

$$H=z+\frac{p}{\gamma}+\frac{v^2}{2g}$$

由于伯努利方程中每一项都代表一个高度，所以可用几何图形来表示各物理量之间的关系。如图 2.1-4 所示，连接$\frac{p}{\gamma}$各顶点而成的线叫做测压管水头线（或静力水头线），它是一条随过水断面改变而起伏的曲线；连接$\frac{v^2}{2g}$各顶点而成的线叫做总水头线。由图 2.1-4 看出，理想流体运动中，因为不形成水头损失，故有 $H_1=H_2=H=$常数，即流线上各点的总水头是相等的，其总水头顶点的联线是一条水平线。而 $H_p\neq$常数（因为速度头$\frac{v^2}{2g}$是随过水断面的改变而变化的）。不过包括位头（连接各点 z 而成的）在内的三个水头

可以相互转化，而总水头却仍不变。

按式（2.1-23）可绘出实际流体总流的几何图形，如图2.1-5所示。可以看出，在黏性流体运动中，因为形成水头损失，故 $H_1 \neq H_2$，即沿着流向，总水头必然是降低的，所以其总水头线是一条沿流向向下倾斜的曲线。与理想流体运动的情形一样，此时其测压管水头线（静压水头线）还是一条随着过水断面改变而起伏的曲线。

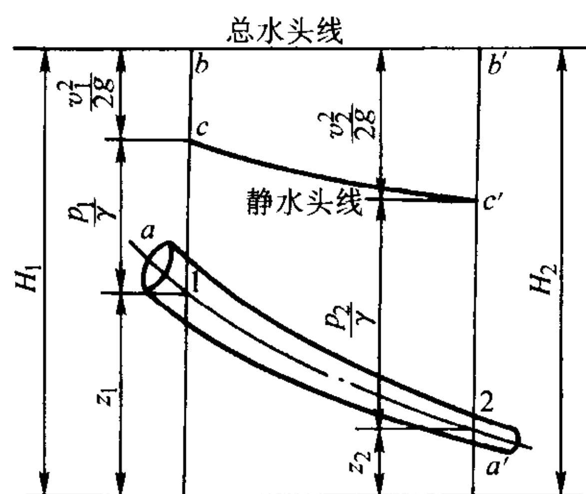


图 2.1-4 理想流体微元流束伯努利方程图解

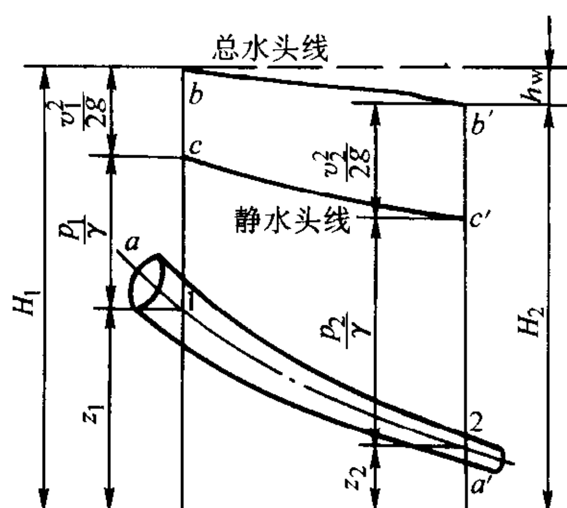


图 2.1-5 黏性流体微元流束伯努利方程图解

2）物理意义 从几何意义的讨论可以看出，方程中的每一项都具有相应的能量意义。

式（2.1-24）中，zg 可看成是单位质量流体流经该点时所具有的位能，称比位能；$\frac{p}{\rho}$ 可看成是单位质量流体流经该点时所具有的压能，称比压能；$\frac{v^2}{2}$ 是单位质量流体流经给定点时的动能，称比动能；$\frac{h'_w}{g}$ 是单位质量流体在流动过程中所损耗的机械能，称能量损失。

对于理想流体，$z_1 g + \frac{p_1}{\rho} + \frac{v_1^2}{2} = z_2 g + \frac{p_2}{\rho} + \frac{v_2^2}{2}$，表明单位质量无黏性流体沿流线自位置1流到位置2时，其各项能量可以相互转化，但它们的总和却是不变的。若用 E 表示各项比能之和，并称为总比能，则单位质量无黏性流体在整个流动过程中，总比能为一个不变的常数，即 $E=$常数。

对于黏性流体，式（2.1-24）的物理意义表明，单位质量黏性流体沿流线自位置1流到位置2时，不但各项能量可以相互转化，而且它的总机械能也是有损失的。设 E 表示总比能，ΔE 表示单位质量流体总比能的损失，则

$$E_1 = E_2 + \Delta E$$

该式表明，单位质量黏性流体在整个流动过程中，其总比能是一定有损失的。

1.2.4 动量方程

在某些工程问题上往往需要了解运动流体与固体边界面上的相互作用力，例如水在弯管中流动对管壁的冲击等。动量方程就提供了流体与固体相互作用的动力学规律。若 F 为作用于流体上所有外力的合力矢量，它等于单位时间内流出与流入的动量之差：

$$F = \rho Q\ (v_2 - v_1) \tag{2.1-25}$$

式中 Q——流量（m^3/s 或 kg/s）。

为便于计算，常将式（2.1-25）写成空间坐标的投影式，即

$$\begin{aligned} F_x &= \rho Q\ (v_{2x} - v_{1x}) \\ F_y &= \rho Q\ (v_{2y} - v_{1y}) \\ F_z &= \rho Q\ (v_{2z} - v_{1z}) \end{aligned} \tag{2.1-26}$$

1.3 流动阻力

1.3.1 流动阻力产生的原因

实际流体流动中，流体在管内流动或绕经物体流动时，都存在流动阻力，其方向与流速相反。

（1）管流的流动阻力

1）发生在整个流程上，由于黏性（内摩擦力）所引起的沿程阻力。

2）流经局部管件时，由于漩涡和速度分布的改组所引起的局部阻力。

（2）绕流的流动阻力

1）由于黏性所引起的摩擦阻力。

2）由于流束变形及漩涡等产生压力差所引起的压差阻力。

3）由于流体作非稳定流动或物体在流体中作加速运动所引起的惯性阻力，等等。

1.3.2 管内流动阻力的计算

（1）水头损失及其计算公式

1）水头损失 不可压缩流体的管流，通常不要求计算阻力，而要求计算由于阻力所引起的流体机械能的损失，称其为水头损失。不可压缩流体管流的水头损失分沿程水头损失（由沿程阻力引起）和局部水头损失（由局部阻力引起）。

2）计算公式 管流水头损失用单位质量流体的总水头损失 h_s 度量。若流体从过流断面1流到过流断面2，则

$$h_s = \left(z_1 + \frac{p_1}{\gamma} + \frac{v_1^2}{2g}\right) - \left(z_2 + \frac{p_2}{\gamma} + \frac{v_2^2}{2g}\right) \tag{2.1-27}$$

总水头损失 h_s 为各沿程水头损失 h_y 和局部水头损失 h_j 之和，即

$$h_s = \sum h_y + \sum h_j \tag{2.1-28}$$

$$h_y = \lambda \frac{l}{d_s} \frac{v^2}{2g} \tag{2.1-29}$$

$$h_j = \xi \frac{v^2}{2g} \tag{2.1-30}$$

式中，λ 为沿程阻力系数（详见本章1.3.3节）；ξ 为局部阻力系数，根据接头、阀门等形状而变化，其值可查阅有关手册；v 为平均流速，计算局部水头损失时，除特别注明者外，系用局部管件后的过流断面上的平均流速；l 为计算沿程损失 h_y 的那段管长；d_s 为上述管段的水力直径。

式（2.1-29）也称作沿程损失的达西公式。

（2）流动光滑管和粗糙管

1）层流底层 管流是湍流时，在紧靠管壁处存在一薄层流体仍保持层流（图2.1-6），这层流体称为层流底层。

圆管中的层流底层厚度为

$$\delta_n \approx 30 \frac{d}{Re\sqrt{\lambda}} \tag{2.1-31}$$

式中，d 为圆管内径；Re 为以 d 为特征长度的雷诺数，$Re = \frac{vd}{\nu}$。

2）流动光滑管 层流底层厚度 δ_n 大于壁面绝对粗糙度

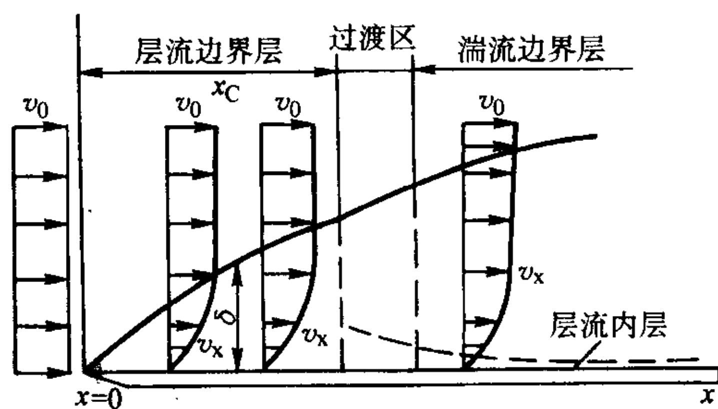
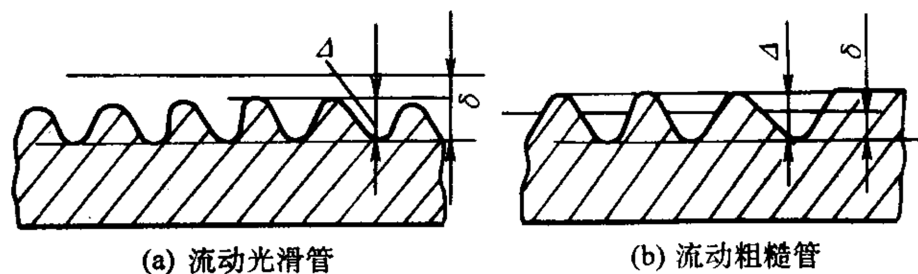


图 2.1-6　管壁附近的层流底层



(a) 流动光滑管　　(b) 流动粗糙管

图 2.1-7

Δ 的管子（图 2.1-7a），称为流动光滑管。计算流动光滑管的水头损失，可不考虑管壁粗糙度的影响。

3）流动粗糙管　层流底层厚度 δ_n 小于管壁绝对粗糙度 Δ 的管子（图 2.1-7b），称为流动粗糙管。计算流动粗糙管的水头损失，必须考虑管壁粗糙度的影响。

层流底层厚度 δ_n 随雷诺数变化，所以对同一管子，在某一雷诺数时是流动光滑管，而在另一雷诺数时可能转变为流动粗糙管。

1.3.3　沿程阻力系数

（1）影响沿程阻力系数的因素

沿程阻力系数 λ 与雷诺数 Re、管壁相对粗糙度 Δ/d_s 及过流断面形状有关。对于过流断面形状相同的管流：

1）层流时　λ 只与 Re 有关，见表 2.1-2。

表 2.1-2　圆断面管流沿程阻力系数 λ 的计算公式

流动状态			雷诺数 $Re=\frac{vd}{\nu}$		计算公式
层　流			<2 300		$\lambda=\frac{64}{Re}$
湍流	流动光滑区		$\leqslant 27.0\left(\frac{d}{\Delta}\right)^{\frac{8}{7}}$	$3\times10^3<Re<10^5$	$\lambda=\frac{0.316\,4}{Re^{0.25}}$
湍流	流动光滑区		$\leqslant 27.0\left(\frac{d}{\Delta}\right)^{\frac{8}{7}}$	$3\times10^3<Re<4\times10^5$	$\frac{1}{\sqrt{\lambda}}=2\lg(Re\sqrt{\lambda})-0.8$
湍流	流动粗糙区	过渡区	$27.0\left(\frac{d}{\Delta}\right)^{\frac{8}{7}}<Re\leqslant\frac{191.2}{\sqrt{\lambda}}\frac{d}{\Delta}$		$\frac{1}{\sqrt{\lambda}}=-2\lg\left(\frac{\Delta}{3.71d}+\frac{2.51}{Re\sqrt{\lambda}}\right)$
湍流	流动粗糙区	阻力平方区	$>\frac{191.2}{\sqrt{\lambda}}\frac{d}{\Delta}$		$\lambda=\left(2\lg\frac{d}{2\Delta}+1.74\right)^{-2}$

2）湍流时

① 在流动光滑区（流动光滑管的情况），λ 只与 Re 有关。通常，当 $Re=\frac{vd_s}{\nu}\leqslant 27.0\left(\frac{d_s}{\Delta}\right)^{\frac{8}{7}}$ 时，按流动光滑管计算。

② 阻力平方区（自模拟区），λ 只与 $\frac{\Delta}{d_s}$ 有关。通常，当 $Re=\frac{vd_s}{\nu}>\frac{191.2}{\sqrt{\lambda}}\frac{d_s}{\Delta}$ 时，按阻力平方区计算。

③ 过渡区，λ 与 Re 和 $\frac{\Delta}{d_s}$ 都有关。通常，当 $27.0\left(\frac{d_s}{\Delta}\right)^{\frac{8}{7}}<Re\left(=\frac{vd_s}{\nu}\right)\leqslant\frac{191.2}{\sqrt{\lambda}}\frac{d_s}{\Delta}$ 时，按过渡区计算。

（2）圆断面管流的沿程阻力系数　圆断面管流的沿程阻力系数 λ，可按表 2.1-2 中的公式计算，或由图 2.1-8 直接查得。

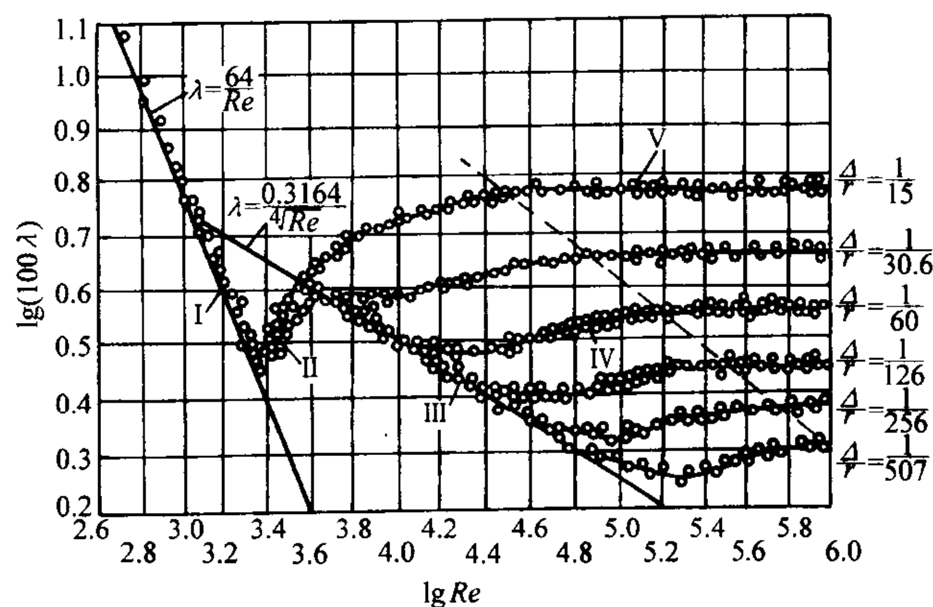


图 2.1-8　尼古拉茨阻力系数曲线

1.4　特殊流体流动

1.4.1　单个颗粒在流体中的运动

虽然在一定条件下可以通过求解纳维尔－斯托克斯方程，来计算一个形状规则的固体颗粒周围的速度分布，从而计算出作用力，但在一般情况下，更为方便的是借助于经验的阻力系数或摩擦系数，来估算作用于运动颗粒上的力。

如图 2.1-9 所示，当流体与浸没在其中的颗粒有相对运动时，除有重力及浮力的作用外，还有阻力作用于颗粒上，此阻力 F_d 可表示为

$$F_d=C_dA_p\frac{1}{2}\rho v_t^2 \qquad (2.1\text{-}32)$$

式中，C_d 为无因次阻力系数（见 1.3.3 节）；A_p 为颗粒在垂直于运动方向的平面投影面积；v_t 为流体对颗粒的相对速度。

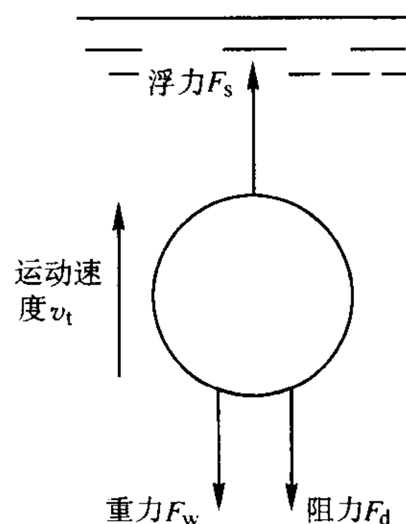


图 2.1-9　流体中作用于颗粒的力

当 $10^{-3}<Re_p<2$，颗粒处于缓慢流动或斯托克斯定律区。这一区域中，阻力系数与颗粒雷诺数成反比，可写成

$$C_d=\frac{24}{Re_p} \qquad (2.1\text{-}33)$$

将式（2.1-33）中的 C_d 代入式（2.1-32），并注意到球形颗粒的投影面积 $A_p=\pi r_p^2$，则得

$$F_d=6\pi\eta r_pv_t \qquad (2.1\text{-}34)$$

这就是著名的斯托克斯公式。

一个颗粒在流体中开始沉降（如灰尘在空气中）或上升（如夹杂物在钢液中）时，它会逐渐加速，直到引起这一运动的重力（或其它的体积力场）与阻碍这一运动的阻力相平衡时为止。如果流体是静止的，那么从达到平衡时起，颗粒的运动速度就为常数，这一速度称为沉降（或上升）终速。

令作用于颗粒上的体积力（通常为重力）与阻力相等，就可以求出沉降终速的值。参考图2.1-9，对于缓慢流动区（斯托克斯定律区），有

$$F_d=\frac{4}{3}\pi r_p^3(\rho_p-\rho)g=6\pi\eta r_p v_t \tag{2.1-35}$$

即

$$v_t=\frac{2r_p^2}{9\eta}(\rho_p-\rho)g \tag{2.1-36}$$

式中 ρ_p 是颗粒的密度。上式即为常见的沉降终速的斯托克斯公式。

当运动颗粒靠近固体表面时，作用于颗粒上的力还将有范德华力或表面张力的存在，发生力的不平衡，颗粒受到加速运动，根据力的作用情况而被表面所排斥或吸引。例如，金属的凝固前沿与金属液中的夹杂物（或陶瓷增强颗粒）发生相互作用，颗粒或被排斥而推移、或被捕捉而进入固相。这类问题归纳为流体中异相颗粒在固体表面的行为，很有实际意义。

1.4.2　悬浮液中颗粒的运动

以上的单个颗粒的沉降通常称为自由沉降。这只有颗粒很少、颗粒间相互影响可以忽略的情况下发生。当溶液中颗粒含量较多时，颗粒间相互作用不能忽略，发生“干涉沉降”，通常使沉降速度减小。颗粒沉降速度发生变化的主要原因有：①大颗粒是相对于小颗粒悬浮体进行沉降的，因此对大颗粒而言，流体有效密度和黏度增大；②因颗粒沉降而排出的流体，以相当大的速度向上运动，对其他颗粒的沉降具有阻滞作用，使沉降速度减小。

诸如颗粒增强金属基复合材料熔液这样的悬浮液，其黏度 η_C 与熔液中颗粒的体积分数 V_f 有关

$$\eta_C=\eta_M(1+2.5V_f+10.25V_f^2),\ V_f<0.25 \tag{2.1-37}$$

式中，η_M 为无颗粒时基础溶液的黏度。

悬浮液中颗粒的沉降速度 v_h 与沉降终速 v_t 有如下关系

$$v_h=v_t\ (1-V_f)^n \tag{2.1-38}$$

式中，n 为常数，n 的值只取决于管径 D 和粒径 d 之比

当 $Re_p<0.2$ 时，　$n=4.65+19.5d/D$　(2.1-39)

当 $Re_p>500$ 时，　$n=2.4$　(2.1-40)

在过渡区中，发现 n 取决于颗粒雷诺数和 d/D。

计算表明，悬浮液的颗粒浓度越大，沉降越缓慢，颗粒含量20%时的沉降速度不到5%时的1/3。研究表明，工程中颗粒含量在25%～30%时，颗粒的沉降速度已很小，不容易产生颗粒的偏析。

2　材料成形中的热量传输原理

2.1　基本概念

2.1.1　三种传热方式

热量传输简称传热，是一种极为普遍而又重要的物理现象。工件在制造工艺中的加热、冷却、熔化和凝固等均与热量的传递息息相关。热量传递有三种基本方式：导热、对流和辐射。

1）导热　物体各部分之间不发生相对位移，依靠分子、原子及自由电子等微观粒子的热运动进行的热量传递称为导热。

2）对流　流体各部分之间发生相对位移，冷热流体相互掺混所引起的热量传递方式称为对流。对流仅能发生在流体中，而且必然伴随着导热。工程上常遇到的不是单纯对流方式，而是流体流过固体表面时对流和导热联合起作用的方式。后者称为对流换热，以区别于单纯对流。

对流换热按引起流体流动的不同原因可分为自然对流与强制对流两大类。自然对流是由于液体冷、热各部分密度不同而引起的，暖气片表面附近热空气向上流动就是一个例子。如果流体的流动是由于水泵、风机或其他压差所造成的，则称为强制对流。另外，沸腾及凝结也属于对流换热，熔化及凝固则除导热外也常伴有对流换热，并且它们都是带有相变的对流换热。

3）辐射　物体通过电磁波传递能量的方式称为辐射。物体会因各种原因发出辐射能，其中因热的原因发出辐射能的现象称为热辐射。

2.1.2　傅里叶导热定律

描述导热现象的基本规律是傅里叶导热定律，它是一个经验公式。通过平板的导热如图2.1-10所示。平板的两个表面均维持各自的均匀温度。这是一维导热问题。对于 x 方向上任意一个厚度为 dx 的微薄层，根据傅里叶定律，单位时间内通过该层的热量，与该处的温度变化率及平板的截面积 A 成正比，即

$$\Phi=-\lambda A\frac{dT}{dx} \tag{2.1-41}$$

式中，λ 是个比例系数，称为热导率；负号表示热量传递的方向与温度升高的方向相反。

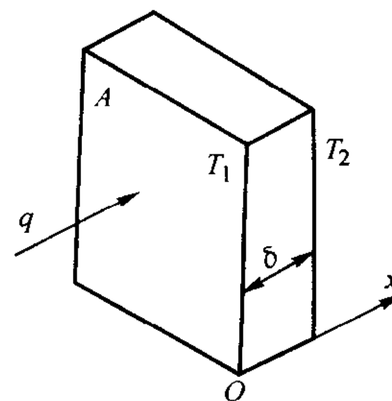


图2.1-10　通过平板的一维导热

单位时间内通过某一给定面积的热量称为热流量，记为 Φ，单位为W。单位时间内通过单位面积的热量称为热流密度（又称比热流），记为 q，单位为W/m^2。傅里叶定律用热流密度形式表示则为

$$q=\frac{\Phi}{A}=-\lambda\frac{dT}{dx} \tag{2.1-42}$$

式（2.1-41）和式（2.1-42）是一维稳态导热时傅里叶定律的数学表达式。

2.1.3　热导率与热扩散率

1）热导率　热导率的定义可由傅里叶定律表达式推出。由式（2.1-42）得到

$$\lambda=-\frac{q}{\frac{\partial T}{\partial n}} \tag{2.1-43}$$

由此可见，热导率在数值上等于温度梯度为1个单位时，物体内具有的热流密度，单位为W/（m·℃）。它反映出，在相同的温度梯度下，物体的热导率越大，导热量也越大。因此，λ 是表征物体导热能力的重要物性参数。热导率的大小取决于物质的种类和温度，可通过查有关手册获得。

2）热扩散率　傅里叶定律可以改写为

$$q=-\frac{\lambda\partial(c\rho T)}{c\rho\partial x}=-a\frac{\partial(c\rho T)}{\partial x} \tag{2.1-44}$$

式中，ρ 为物体的密度，kg/m^3；c 为物体的比热容，J/（kg·℃）；a 为热扩散率，m^2/s。

$$a=\frac{\lambda}{c\rho} \quad (2.1\text{-}45)$$

由式（2.1-45）可知，热扩散率 a 与热导率 λ 成正比，与物体的密度 ρ 和比热容 c 成反比。a 也是重要物性参数，它表征了物体内热量传输的能力。傅里叶定律的物理意义为：以物体受热升温的情况为例来分析，在升温过程中，进入物体的热量沿途不断地被吸收而使该处温度升高，此过程持续到物体内部各点温度全部相同为止。由热扩散率的定义 $a=\lambda/(\rho c)$ 可知：当其分子 λ 越大，或其分母 ρc（它是单位体积的物体升高1℃所需的热量）越小，表示导出的热量相对较高或吸收的热量相对较少，于是热量的传输就越快，物体内部温度趋于一致的能力就越大。

2.2 导热

2.2.1 导热微分方程

导热微分方程式的一般形式为：

$$\frac{\partial T}{\partial t}=\frac{\lambda}{\rho c}\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}\right)+\frac{\dot{Q}}{\rho c} \quad (2.1\text{-}46)$$

式中，$\dot{Q}$ 为单位体积内热源的热能。式（2.1-46）对稳态、非稳态，对有无内热源的问题都可适用。稳态问题以及无内热源的问题都是上述微分方程式的特例。例如，在稳态、无内热源条件下，导热微分方程式就简化成为

$$\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}=0 \quad (2.1\text{-}47)$$

一般地讲，解非稳态导热问题的定解条件有两个方面：

1）给出初始时刻的温度分布即初始条件。

2）给出物体边界上的温度或换热情况即边界条件。

运用数学上的坐标转换，式（2.1-46）可以转换成圆柱坐标或球坐标表达式。

圆柱坐标

$$\frac{\partial T}{\partial t}=a\left(\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}+\frac{1}{r^2}\frac{\partial^2 T}{\partial \theta^2}+\frac{\partial^2 T}{\partial z^2}\right)+\frac{\dot{Q}}{\rho c}$$

球坐标

$$\frac{\partial T}{\partial t}=a\left[\frac{1}{r}\frac{\partial^2(rT)}{\partial r^2}+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial T}{\partial \theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 T}{\partial \phi^2}\right]+\frac{\dot{Q}}{\rho c}$$

2.2.2 一维稳态导热

工程实践中存在着大量的稳态导热问题，有些问题在一定条件下可以简化成一维稳态导热，即温度仅沿一个空间坐标方向变化。对于一维稳态导热过程，如大平板、长圆筒、球壁等几何形态规则物体的导热问题，采用直接积分法即可获得其解析解。

（1）单层平壁的导热

单层平壁示于图2.1-11。已知平壁的两个表面分别维持均匀而恒定的温度 T_1 和 T_2，壁厚为 δ。假设壁厚远小于高度和宽度，则温度场是一维的。这时微分方程为

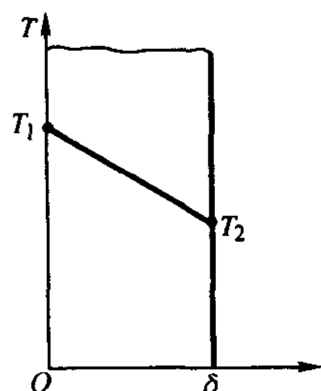


图2.1-11 单层平壁示意图

$$\frac{d^2 T}{dx^2}=0 \quad (2.1\text{-}48)$$

边界条件为：$x=0$ 时，$T=T_1$；$x=\delta$ 时，$T=T_2$。

对式（2.1-48）积分解得，通过平壁的热流密度为

$$q=\frac{\lambda\ (T_1-T_2)}{\delta}=\frac{\lambda}{\delta}\Delta T \quad (2.1\text{-}49)$$

下面引出一个在传热分析中颇为重要的热阻的概念。式（2.1-49）可改写为

$$q=\frac{\Delta T}{\delta/\lambda}=\frac{\Delta T}{R_t} \quad (2.1\text{-}50)$$

这种表达形式有助于更清楚地理解式中各项的物理意义。式中热流密度 q 为导热过程的热转移量，温差 ΔT 为导热过程的动力，而分母 δ/λ 则为导热过程的阻力。热转移过程的阻力称为热阻，记为 R_t，它与电传输过程中欧姆定律的电阻 R 相类似。

（2）圆筒壁

圆筒壁在工程上应用很广，如管道、轧机辊子等都是实例。先分析单层圆筒壁的导热。参看图2.1-12，已知内、外半径分别为 r_1、r_2 的圆筒壁的内、外表面温度分别维持均匀恒定的温度 T_1 和 T_2。

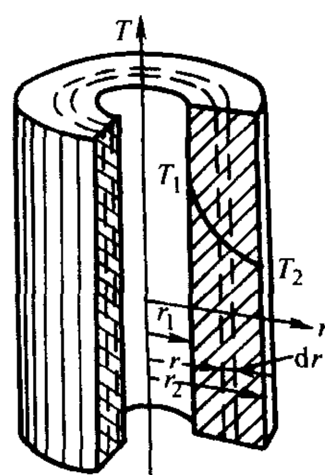


图2.1-12 单层圆筒壁

导热微分方程式简化为

$$\frac{d}{dr}\left(r\frac{dT}{dr}\right)=0 \quad (2.1\text{-}51)$$

边界条件为：$r=r_1$ 时，$T=T_1$；$r=r_2$ 时，$T=T_2$。

解方程式（2.1-51），得

$$T=T_1+\frac{T_2-T_1}{\ln\ (r_2/r_1)}\ln\ (r/r_1) \quad (2.1\text{-}52)$$

热流量计算公式为

$$\Phi=\frac{2\pi\lambda l\ (T_1-T_2)}{\ln\ (r_2/r_1)} \quad 或 \quad \Phi=\frac{2\pi\lambda l\ (T_1-T_2)}{\ln\ (d_2/d_1)} \quad (2.1\text{-}53)$$

对于圆筒壁，其总面积热阻有下列表达式：

$$R_{t,z}=\frac{\Delta T}{\Phi}=\frac{\ln\ (d_2/d_1)}{2\pi\lambda l}$$

与分析多层平壁一样，运用串联热阻叠加原则，可得通过三层圆筒壁的热流量：

$$\Phi=\frac{2\pi l(T_1-T_4)}{\ln(d_2/d_1)/\lambda_1+\ln(d_3/d_2)/\lambda_2+\ln(d_4/d_3)/\lambda_3}$$

2.2.3 二维稳态导热

在许多实际问题中，一维导热的简化分析方法不能满足工程计算的需要。这时稳态导热的温度分布将是两个或三个空间坐标的函数，称为二维或三维稳态导热。相应的导热方程是包含两个或三个自变量的偏微分方程。二维以上的多维稳态导热有多种分析求解方法，其中分离变量法是广泛采用的经典而有效的方法。

对于半无限大的平板，温度分布如图2.1-13所示。因平板相对很薄，可以认为$\frac{\partial T}{\partial z}=0$，温度场是二维的。在稳态下，它必须满足方程：

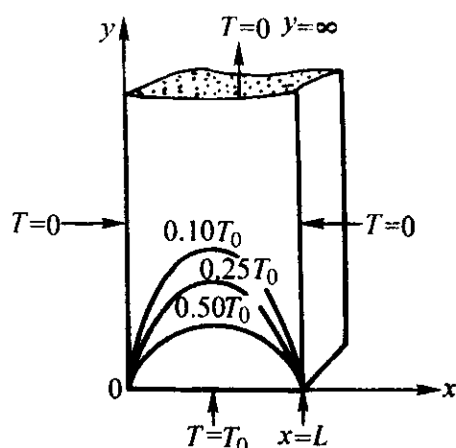


图 2.1-13　半无限大平板内的温度分布

$$\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}=0 \tag{2.1-54}$$

其边界条件为：

1）当 $x=0$ 时，$T=0$；

2）当 $x=L$ 时，$T=0$；

3）当 $y=\infty$ 时，$T=0$；

4）当 $y=0$ 时，$T=T_0$（均匀）。

利用分离变量法可解得温度分布的函数关系为

$$\frac{T}{T_0}=\sum_{n=0}^{\infty}\frac{4}{n\pi}e^{-(n\pi/L)y}\sin\frac{n\pi x}{L} \tag{2.1-55}$$

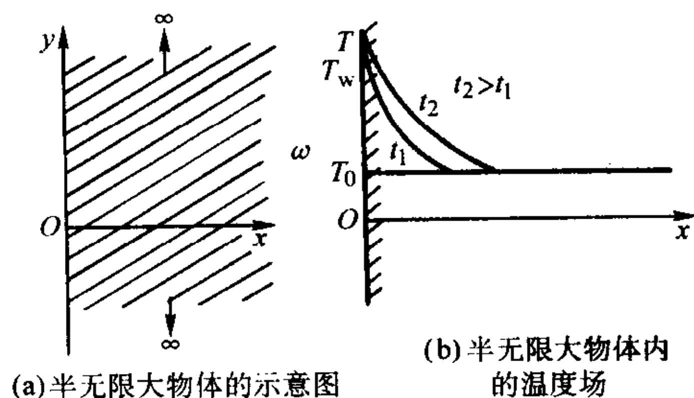
2.2.4　非稳态导热

以一块半无限大平板的导热为例，其初始温度均匀并等于室温 T_0，其表面被突然加热。参看图 2.1-14，起初表面温度 T_w 开始上升而中心温度仍为初始的温度 T_0。然后随着时间的推移，温度变化波及范围不断扩大，导致内部温度也开始上升。以上分析表明：

1）物体内温度的变化存在着部分物体不参与变化和整个物体参与变化的两个阶段。

2）不同位置达到指定温度的时间不同。这是非稳态导热问题求解的重要任务。

3）在热量传递的过程中，由于物体本身的温度变化要积蓄（或放出）热量，传热开始时这份热量较大，随着物体温度的变化，这份热量逐渐减小，在热平衡状态下降为零。即积蓄（或放出）的热量是随时间而变化的，这也是非稳态导热问题求解的任务。



(a)半无限大物体的示意图　　(b)半无限大物体内的温度场

图 2.1-14　表面温度跃升后的温度变化示意图

非稳态导热的计算有两种情况：

1）表面温度为常数的一维非稳态导热

$$\frac{\partial T}{\partial t}=a\frac{\partial^2 T}{\partial x^2} \tag{2.1-56}$$

边界条件为：$t=0$ 时，$T=T_0$ = 定值；$t>0$，$x=0$ 处，$T=T_w$ = 定值。方程的理论解为

$$\frac{T_w-T_0}{T_w-T_0}=\mathrm{erf}\left(\frac{x}{2\sqrt{at}}\right)=\mathrm{erf}\ (N) \tag{2.1-57}$$

式中，$N=x/(2\sqrt{at})$，erf（N）为高斯误差函数，它的数值可按 N 值查表得出。从傅里叶定律能解得表面的瞬时热流密度 q_w。

$$q_w=-\lambda\left.\frac{\partial T}{\partial x}\right|_{x=0}=\lambda(T_w-T_0)\frac{1}{\sqrt{\pi at}} \tag{2.1-58}$$

2）其他非稳态导热问题　在工程实际中，通常利用几个准数及线图进行计算。

2.3　对流换热

2.3.1　对流换热的准数方程式

流体流过固体表面所发生的热量传递称为对流换热；对流换热的基本计算式是牛顿冷却公式，即

$$q=h\ (T_w-T_f) \tag{2.1-59}$$

式中，h 为换热系数 W/（$m^2\cdot K$）；T_w 为固体表面温度；T_f 为流体温度。

对于面积为 A 的接触面，对流换热的热流量为

$$\Phi=hA\ (T_w-T_f) \tag{2.1-60}$$

对流换热计算的主要任务之一是求解换热系数 h。主要是利用准数方程式来计算换热系数 h 及其他物理量。与对流换热有关的准数及准数方程主要有以下一些。

1）努塞尔数 Nu

$$Nu=\frac{hl}{\lambda}=\frac{1/\lambda}{1/h}=\frac{\text{导热热阻}}{\text{对流热阻}}$$

Nu 数大说明导热热阻 $1/\lambda$ 大而对流热阻 $1/h$ 小，即对流作用强烈。由于 Nu 数中包含有对流换热系数 h，它是被决定准数，在对流换热中最为重要。

2）傅里叶数 Fo

$$Fo=\frac{at}{l^2}=\frac{\Delta T\lambda/l^2}{\Delta T\rho C_p/t}=\frac{\text{单位体积物体的导热速率}}{\text{单位体积物体的蓄热速率}}$$

它与时间因素有关，所以 Fo 是表示温度场随时间变化的不稳定导热的准数。

3）普朗特数 Pr　Pr 是流体物性的无因次组合，又称物性准数。

$$Pr=\frac{\nu}{a}=\frac{c\rho\nu}{\lambda}$$

Pr 表示流体动量传输能力与热量传输能力之比。从边界层概念出发，可以认为是动力边界层与热边界层的相对厚度指标。

4）贝克来数

$$Pe=\frac{vl}{a}=\frac{vl}{\nu}\frac{\nu}{a}=RePr=\frac{vc\rho}{\lambda/l}=\frac{\text{流体带入的热量}}{\text{流体导出的热量}}$$

它是表明温度场在空间分布的准数，Pe 越大说明进入系统的热量越大，导出的热量越少，则温度场处于非稳定状态。

5）格拉晓夫数 Gr

$$Gr=\frac{\alpha_v gL^3}{v^2}\Delta T=\frac{(l\rho g\alpha_v\Delta T)(\rho\nu)^2}{(\eta\nu/l)^2}=\frac{\text{浮力}\times\text{惯性力}}{\text{阻力}}$$

Gr 值越大，引起对流的浮力相对于阻力越大，自然对流也越强烈。

6）准数方程式　强制对流换热时，表示自然对流浮升力影响的 Gr 数可以忽略，准数方程为

$$Nu=f\ (Re,\ Pr) \tag{2.1-61}$$

自然对流时又可忽略 Re 数，而有

$$Nu=f\ (Gr,\ Pr) \tag{2.1-62}$$

在具体应用时，多表示为幂函数形式，如

$$Nu=CRe^nPr^m \tag{2.1-63}$$

$$Nu=C\ (GrPr)^n \tag{2.1-64}$$

式中的 C、n、m 通过实验求得。

2.3.2　强制对流换热

（1）外掠平板

层流区（$Re<5\times10^5$）

$$Nu = 0.664Re^{0.5}Pr^{1/3} \qquad (2.1\text{-}65)$$

湍流区（$5\times10^5 \leqslant Re < 10^7$）

$$Nu = (0.037Re^{0.8} - 871)Pr^{1/3} \qquad (2.1\text{-}66)$$

定性温度取边界平均温度 T_m

$$T_m = (T_w + T_\infty)/2$$

式中，T_w 为板面温度；T_∞ 为来流温度。特性尺度取板全长 l；Re 数中的速度取来流速度 v_∞。

（2）横掠圆柱（圆管）

流体横掠圆柱时的流动特征示于图 2.1-15。边界层的形态出现在前半圈的大部分范围，然后发生绕流脱体，在后半圈出现回流和旋涡。平均换热系数的计算式为

$$Nu = cRe^n \qquad (2.1\text{-}67)$$

式中，在不同 Re 区段内 C 和 n 具有不同的数值，见表 2.1-3。此外，定性温度采用边界层平均温度 $T_m = (T_w + T_f)/2$，特性尺度取圆柱外径 d，Re 数中的流速按来流流速计算。

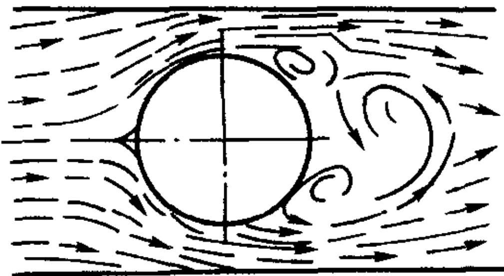


图 2.1-15　流体横向掠过圆柱管时的流动特点

表 2.1-3　C 和 n 值

Re	C	n
4 ~ 40	0.809	0.385
40 ~ 4 000	0.606	0.466
4 000 ~ 40 000	0.171	0.618
40 000 ~ 250 000	0.023 9	0.805

（3）绕流球体

流体绕流球体时，边界层的发展及分离与绕流圆管相类似。流体与球体表面间的平均换热系数可按下列准数方程计算：

对于空气

$$Nu_m = 0.37Re_m^{0.6} \qquad (2.1\text{-}68)$$

对于液体

$$Nu_m = 2.0 + 0.6Re_m^{\frac{1}{2}}Pr_m^{\frac{1}{3}} \qquad (2.1\text{-}69)$$

式（2.1-68）的适用范围是 $17 < Re_m < 70\,000$，定性温度为 T_m，定型尺寸为球体直径 d。式（2.1-69）的适用范围是 $1 < Re_m < 70\,000$，$0.6 < Pr_m < 400$，定性温度为 T_m，定型尺寸为球体直径 d。式（2.1-69）表明，$Re_m \to 0$ 时，Nu_m 趋近于 2。这一结果相当于无限滞止介质中，温度均匀的球体稳态导热时求得的 Nu_m 值。

（4）管内流动

在 $Re > 10\,000$ 的旺盛湍流区，使用最广的实验准则式为

$$Nu_f = 0.023Re_f^{0.8}Pr_f^{0.4} \qquad (2.1\text{-}70)$$

式中，定性温度取流体平均温度 T_f，习惯上 T_f 取管道进出口两截面温度的算术平均值；特性尺度取管内径。当温度差较大及流体平均温度 T_f 下的动力黏度 η_f 大于水的黏性系数两倍以上时，上式误差较大，这时可在下列实验准则式中任选一个进行计算：

$$Nu_f = 0.027Re_f^{0.8}Pr_f^{1/3}\left(\frac{\eta_f}{\eta_w}\right)^{0.14} \qquad (2.1\text{-}71)$$

$$Nu_f = 0.021Re_f^{0.8}Pr_f^{0.43}\left(\frac{Pr_f}{Pr_w}\right)^{0.25} \qquad (2.1\text{-}72)$$

两式中，除壁温 T_w 下动力黏度 η_w 或 Pr_w 取壁温 T_w 为定性温度外，其余物性仍采用流体平均温度为定性温度；管内径 d 为特性尺度。

2.3.3 自然对流换热

在自然对流换热中，代表浮力与黏性力之比、并且包括温度差 ΔT 的 Gr 准数起决定性作用。自然对流中，靠近固定表面流体的流动层就是自然对流边界层，由于其贴近固体表面处流速为零，而边界层以外静止流体的流速也为零，因而在边界层内存在一流速极大值。

自然对流换热的准数方程式为式（2.1-64），即

$$Nu = C(GrPr)^n$$

式中的 C 及 n 值与流动性质及表面朝向有关，由表 2.1-4 决定。

表 2.1-4　自然对流换热中的 C 及 n 值

表面状况	流动示意图	定型尺寸	流态及 C、n 值			
			$GrPr$ 范围	流态	C	n
垂直平板或垂直圆柱		板或柱高度 H	$10^4 \sim 10^9$	层流	0.59	1/4
			$10^9 \sim 10^{13}$	湍流	0.10	1/3
水平圆柱		外径 d	$10^4 \sim 10^9$	层流	0.53	1/4
			$10^9 \sim 10^{12}$	湍流	0.13	1/3
水平板热面向上或水平板冷面向下		矩形板取两边长平均值 L，圆板取 $0.9d$	$2\times10^4 \sim 8\times10^6$	层流	0.54	1/4
			$8\times10^6 \sim 10^{11}$	湍流	0.15	1/3
水平板热面向下或水平板冷面向上		矩形板取两边长平均值 L，圆板取 $0.9d$	$10^5 \sim 10^{11}$	层流	0.58	1/5

2.4 辐射换热

2.4.1 物体的辐射性质

1）吸收率 α、反射率 ρ 和透射率 τ　投射到物体表面上的热辐射和可见光一样，也有吸收、反射和穿透现象。若投射的热辐射能量为 Q，吸收、反射和透射的辐射能数量依次为 Q_α、Q_ρ 和 Q_τ，按能量守恒定律有：

$$\frac{Q_\alpha}{Q} + \frac{Q_\rho}{Q} + \frac{Q_\tau}{Q} = 1 \qquad (2.1\text{-}73)$$

上式左方的各能量百分比分别称为物体的吸收率、反射率和穿透率，记为 α、ρ 和 τ。于是式（2.1-73）可表示成：

$$\alpha + \rho + \tau = 1 \qquad (2.1\text{-}74)$$

通常认为固体和液体不允许热辐射穿透，即 $\tau = 0$。气体对投射来的热辐射几乎没有反射能力，可认为气体的反射率 $\rho = 0$。不同物体的吸收率、反射率和穿透率因具体条件不同而千差万别，我们把吸收率 $\alpha = 1$ 的物体叫做绝对黑体，简称黑体；把反射率 $\rho = 1$ 的物体叫做镜体；把透射率 $\tau = 1$ 的物体叫做透明体。

2）黑体的辐射力　在同一温度下，黑体的辐射力 E_b 最大。黑体的辐射力为斯蒂芬 - 玻耳兹曼表达式（又称四次方定律）：

$$E_b = \int_0^\infty E_{b\lambda}\,d\lambda = C_b\left(\frac{T}{100}\right)^4 \qquad (2.1\text{-}75)$$

式中，E_b 为单位时间内从单位面积的黑体表面上发射出去的辐射能，称为黑体的全辐射力（简称黑体的辐射力），W/(m^2·h)；T 为黑体表面的温度，°K；C_b 为黑体的辐射系数，

其值为5.67 W/（$m^2 \cdot K^4$）；$E_{b\lambda}$为黑体的单色辐射强度。

3）黑度　实际物体的辐射力与同温度下黑体辐射力之比称为该物体的发射率或称黑度 ε：

$$\varepsilon = E/E_b \tag{2.1-76}$$

4）角系数　表面1发射出的辐射能 Q_1 落到表面2上的百分数称为表面1对表面2的角系数，记为 X_{12}。它与表面1和2的几何形状、大小及相对位置有关。

5）灰体　实际物体的单色吸收率 α_λ 对不同波长的辐射具有选择性，即 α_λ 与波长 λ 有关。如果假定物体的单色吸收率与波长 λ 无关，即 α_λ = 常数，这种假定的物体称之为灰体。灰体的吸收率 α 在数值上等于灰体在同温度下的发射率，即 $\alpha = \varepsilon$。因此，在计算灰体间的辐射换热时，吸收率和发射率可以互相对换。

2.4.2　固体表面间的辐射换热

1）有效辐射　单位时间内投射到单位表面积上的总辐射能被称为投入辐射，记为 G；单位时间内离开单位表面积的总辐射能为该表面的有效辐射，记为 J。有效辐射不仅包括表面的本身辐射 E，而且还包括投入辐射 G 中被表面反射的部分 ρG。这里 ρ 为表面的反射率，可表示成 $1-\alpha$。参见图2.1-16，考察表面温度均匀、表面辐射特性为常数的表面1。根据有效辐射的定义，表面1的有效辐射 J_1 为

$$J_1 = E_1 + \rho_1 G_1 = \varepsilon_1 E_{b1} + (1-\alpha_1) G_1 \tag{2.1-77}$$

外界能感受到的表面辐射就是有效辐射。用辐射探测仪能够测量到的也是有效辐射。两个灰体间的辐射换热，可表示成与式（2.1-77）相对应的形式

$$\Phi_{12} = A_1 X_{12}\ (J_1 - J_2) \tag{2.1-78}$$

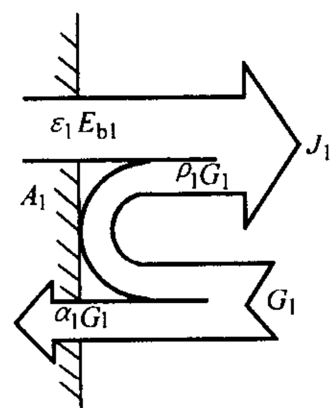
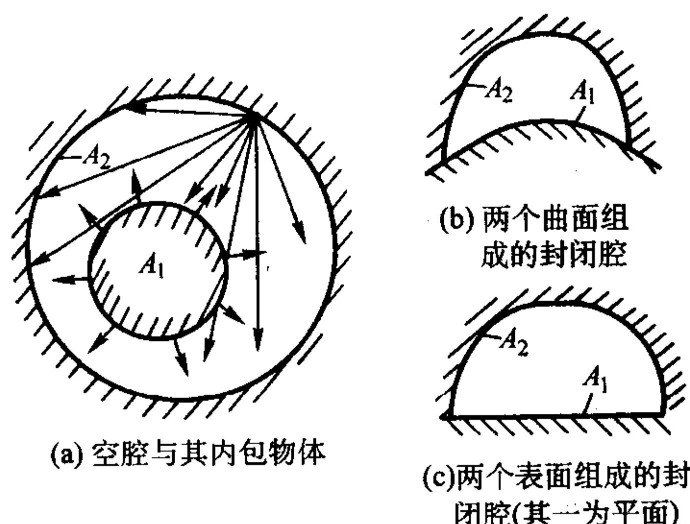


图2.1-16　有效辐射示意图

2）两个固体表面间的辐射换热　现在讨论如图2.1-17所示的仅有两个灰体参与换热的系统。图2.1-17a为空腔与其内包物体组成的换热系统。图2.1-17b、c为仅由两个表面组成的封闭腔。



(a) 空腔与其内包物体　(b) 两个曲面组成的封闭腔　(c)两个表面组成的封闭腔(其一为平面)

图2.1-17　两个物体组成的辐射换热系统

对于两个黑体和灰体表面间的辐射换热量分别为：

黑体
$$\Phi_{12} = \frac{E_{b1} - E_{b2}}{\dfrac{1}{A_1 X_{12}}} \tag{2.1-79}$$

灰体
$$\Phi_{12} = \frac{J_1 - J_2}{\dfrac{1}{A_1 X_{12}}} \tag{2.1-80}$$

2.4.3　气体辐射

在气体发射率 ε_g 和吸收率 α_g 确定之后，气体与黑体包壳之间的辐射换热计算十分简单。这时，只要把气体本身的辐射 $\varepsilon_g E_{bg}$（气体温度 T_g）减去气体所吸收的辐射 $\alpha_g E_{bw}$（包壳温度为 T_w），即可得到气体与黑体包壳间的辐射换热量

$$q = \varepsilon_g E_{bg} - \alpha_g E_{bw} = 5.67\left[\varepsilon_g\left(\frac{T_g}{100}\right)^4 - \alpha_g\left(\frac{T_w}{100}\right)^4\right] \tag{2.1-81}$$

如果包壁是发射率为 ε_w 的灰体，包壳除第一次吸收气体辐射 $\varepsilon_w \varepsilon_g E_{bg}$外，还有反射出去的辐射热量经部分吸收后反复多次返回的辐射热量。同理，气体除第一次吸收包壳本身辐射 $\varepsilon_w a_g E_{bw}$外也还吸收多次反复返回的辐射热量。总之，辐射换热量大于如下只计及第一次的吸收热量

$$q = \varepsilon_w(\varepsilon_g E_{bg} - a_g E_{bw}) \tag{2.1-82}$$

3　材料成形中的质量传输原理

3.1　基本概念和传质微分方程

在一个含有两种或两种以上组分的体系中，如果存在浓度梯度，则每一种组分都有向低浓度方向转移的趋势。物质由高浓度向低浓度方向转移的过程称为质量传输过程，简称传质。正如速度差的存在是动量传递的推动力，温度差的存在是热量传递的推动力那样，浓度差的存在是质量传递的推动力。由于物质的浓度可用多种形式表示，因而传质过程中推动力的表达也有多种形式。

质量传输的基本方式可分为分子传质（又称分子扩散）和对流传质。分子传质类似于导热现象，从本质上说，它们都是依靠分子的随机运动而引起的转移行为。不同的是前者为质量转移，后者为能量转移。分子传质在气相、液相和固相中均能发生。描述分子扩散通量或速率的基本定律为菲克第一定律。

在运动流体与固体壁面之间，或互不相溶的两种运动流体之间发生的质量传递称为对流传质。对流传质类似于对流换热，传质问题的求解也将涉及流体流动形态以及速度分布等因素。

3.1.1　菲克第一定律

在混合物中若各组分存在浓度梯度时，则发生分子扩散。对于两组分系统，通过分子扩散传递的组分 A 的质量通量密度为：

$$j_A = -D_{AB}\frac{d\rho_A}{dz} \tag{2.1-83}$$

式中，z 为组分 A 的密度发生变化的方向的坐标，m；j_A 为组分 A 的质量通量密度，表示单位时间内通过单位面积传递的组分 A 的质量，$kg \cdot m^{-2} \cdot s^{-1}$；$D_{AB}$为组分 A 在组分 B 中的扩散系数，m^2/s；$d\rho_A/dz$ 为组分 A 的质量浓度（密度）梯度，$kg \cdot m^{-3} \cdot m^{-1}$。

式中的负号表示质量通量的方向与浓度梯度的方向相反，即组分 A 朝着浓度降低的方向传递。

3.1.2　传质微分方程

根据质量守恒定律，对于任意选定的微元控制体来说，组分 A 的质量守恒原理可表述如下：

[流出控制体的组分 A 的质量速率] − [流入控制体的组分 A 的质量速度] + [控制体内由化学反应引起组分 A 的生成速率] = [控制体内组分 A 的质量累积速率]　(2.1-84)

据此可推导出双组分混合物中组分 A 的连续性方程（或称 A

的质量传输方程)

$$\frac{D(\rho w_A)}{Dt}+\nabla(\rho w_A v)=D_{AB}\nabla^2(\rho w_A)+r_A \quad (2.1\text{-}85)$$

式中，r_A 为单位控制体积内由于化学反应所引起的组分 A 的生成质量速率，$kg\cdot m^{-3}\cdot s^{-1}$；$w_A$ 为组分 A 的质量分数，%。

同理可得组分 B 的质量传输方程为

$$\frac{D(\rho w_B)}{Dt}+\nabla(\rho w_B v)=D_{AB}\nabla^2(\rho w_B)+r_B \quad (2.1\text{-}86)$$

3.1.3 菲克第二定律

对不可压缩流体总体流动可忽略不计、且没有化学反应时的非稳态传质，传质微分方程可简化为

$$\frac{\partial \rho_A}{\partial t}=D_{AB}\nabla^2\rho_A \quad (2.1\text{-}87)$$

或

$$\frac{\partial C_A}{\partial t}=D_{AB}\nabla^2 C_A \quad (2.1\text{-}88)$$

常将式（2.1-87）称为菲克第二定律。由于假定无总体流动，故它们只适用于固体、静止液体或气体所组成的等摩尔逆扩散体系。式（2.1-88）在直角坐标系中的表示式为

$$\frac{\partial C_A}{\partial t}=D_{AB}\left(\frac{\partial^2 C_A}{\partial x^2}+\frac{\partial^2 C_A}{\partial y^2}+\frac{\partial^2 C_A}{\partial z^2}\right) \quad (2.1\text{-}89)$$

3.2 分子传质

对于在不流动或停滞介质以及固体中以分子扩散方式进行的质量传递过程，由于分子扩散和导热都是由不规则分子热运动引起的，传递的机理类似，在没有总体流动的情况下，方程的形式也类似，因此求解导热问题的方法对分子扩散的求解也是适用的。但是当存在总体流动时，必须注意两者的差别。

3.2.1 一维稳态分子传质

由组分 A 和 B 组成的没有化学反应的两组分混合物，且一种组分的摩尔通量密度与另一种组分的摩尔通量密度大小相等，方向相反，即 $N_A=-N_B$，这种扩散称为等摩尔逆向扩散。计算方法与一维稳态导热具有相同的数学形式，只是意义不同。例如，对无限大平板的等摩尔逆向扩散，有

$$\frac{d^2 C}{dz^2}=0$$

及

$$N_A=-D_{AB}\frac{dC_A}{dz} \quad (2.1\text{-}90)$$

式中，C 为量浓度，mol/m^3；N_A 为组分 A 的摩尔通量密度，$mol\cdot m^{-2}\cdot s^{-1}$。

3.2.2 一维非稳态分子传质

工程中常见到忽略表面阻力的半无限大介质中的非稳态分子扩散，例如钢的表面渗碳工艺中的固相扩散过程就属于一种典型的非稳态分子扩散过程。如图 2.1-18 所示，某一初始含碳量为 C_0 的钢，在电炉中加热到某一需要的温度，在时间 $t=0$ 时暴露在含有 CO_2 和 CO 的气体混合物中，气相中的碳因浓度差而向钢的表层及内部扩散。因渗碳层比工件的断面厚度小很多，因而断面厚度方向可视为半无限大。

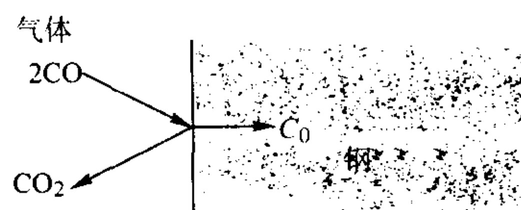


图 2.1-18 钢的表面渗碳

现考虑一初始浓度均匀分布、其值为 C_{A0} 的半无限厚介质（y、z 方向无限大，x 方向半无限大）。若一侧表面浓度突然提高到 C_{Aw}，并维持不变。随时间增加，浓度变化将逐步深入介质的内部。扩散仅沿 x 方向进行。在整个扩散过程中，介质另一侧的浓度始终维持不变。描写这一现象的微分方程为

$$\frac{\partial C_A}{\partial t}=D_{AB}\frac{\partial^2 C_A}{\partial x^2} \quad (2.1\text{-}91)$$

初始条件：$t=0$，对所有 x 值　$C_A=C_{A0}$　(2.1-91a)

边界条件：$t>0$，$x=0$　$C_A=C_{Aw}$　(2.1-91b)

$x=\infty$　$C_A=C_{A0}$　(2.1-91c)

将式（2.1-91）与式（2.1-56）相比较，可知此时的微分方程，边界条件与一维非稳态导热（以及一维非稳态流动）均类似，故可以用分离变量法或拉普拉斯变换法求解。只要将温度换成浓度，将热量扩散系数换成扩散系数，则一维非稳态导热的解式（2.1-57）就可用于一维非稳态分子扩散过程。于是，组分 A 的浓度分布为

$$\frac{C_{Aw}-C_A}{C_{Aw}-C_{A0}}=\text{erf}\left(\frac{x}{2\sqrt{D_{AB}t}}\right) \quad (2.1\text{-}92)$$

式中 erf（x）为高斯误差函数。由式（2.1-92）可以计算任一时刻的浓度分布。不同时刻的浓度分布如图 2.1-19 所示。任何时刻 t 时，在 $x=0$ 处，曲线的斜率为

$$\left.\frac{dC_A}{dx}\right|_{x=0}=\frac{C_{Aw}-C_{A0}}{\sqrt{\pi D_{AB}t}}$$

距离 $\sqrt{\pi D_{AB}t}$ 为渗透深度。

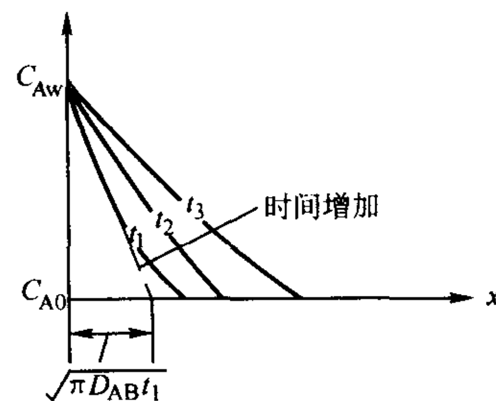


图 2.1-19 半无限大介质的非稳态扩散

3.3 对流传质

在对流传质中，不仅依靠分子扩散，而且依赖于流体各部分之间的宏观相对位移。这与对流换热十分类似。这时质量传递也将受到流体性质、流动状态（层流还是湍流）和流场的几何特性的影响。对流传质通量密度可以用类似于对流换热中牛顿冷却公式的形式来表示，即

$$N_A=k_C\Delta C_A \quad (2.1\text{-}93)$$

式中，N_A 为组分 A 的摩尔通量密度，$kmol\cdot m^{-2}\cdot s^{-1}$；$\Delta C_A$ 为组分 A 的摩尔浓度差，例如：若传质在平板上进行，ΔC_A 表示组分 A 的界面处与边界层外主流的浓度差，即 $\Delta C_A=C_{Aw}-C_{A\infty}$，单位为 $kmol/m^3$；k_C 为以 ΔC_A 为基准的对流传质系数，m/s。为便于区别，当无总体流动时，用 k_C^0 表示对流传质系数。

从式（2.1-93）出发，研究对流传质的基本目的就在于用理论分析或实验方法，来具体揭示各种场合下计算 k_C 的关系式。

下面是对流传质中常见的几个准数。

分子动量扩散率与分子质量扩散率的比值叫做施密特数 Sc

$$Sc=\frac{\nu}{D}=\frac{\eta}{\rho D} \quad (2.1\text{-}94)$$

Sc 与对流换热中的 Pr 具有类似的作用。

分子热扩散率与分子质量扩散率的比值称为路易斯数 Le

$$Le=\frac{\alpha}{D}=\frac{\lambda}{\rho c_p D} \quad (2.1\text{-}95)$$

当过程同时涉及质量和热量传输时，就要用到 Le。施密特数和路易斯数都是流体物性参数的组合，所以它们表示了扩散体系的特性。

分子扩散阻力和对流传质的阻力之比称为舍伍德数 Sh，它类似于对流换热中的努塞尔数

$$Sh=\frac{k_C d}{D} \quad (2.1\text{-}96)$$

下面是一些常见的对流传质准数方程式。

1）圆管内的层流对流传质　对于管壁处组分 A 维持恒定的传质通量的情况，有

$$Sh=\frac{k_C d}{D}=4.36 \quad (2.1\text{-}97)$$

对于管壁处组分 A 的浓度维持恒定的情况，有

$$Sh=3.66 \quad (2.1\text{-}98)$$

2）平板表面通过层流和湍流边界层的传质

$$Sh_L=0.664Re_L^{\frac{1}{2}}Sc^{\frac{1}{3}}\ （层流） \quad (2.1\text{-}99)$$

$$Sh_L=0.036Re_L^{0.8}Sc^{\frac{1}{3}}\ （湍流） \quad (2.1\text{-}100)$$

3）单个球体表面的传质　舍伍德数表示成两项，一项是由纯分子扩散而引起的传质，另一项是由强制对流而引起的传质，即：

$$Sh=2.0+cRe^m Sc^{\frac{1}{3}} \quad (2.1\text{-}101)$$

式中，c 和 m 是关联常数。当 Re 很小时，Sh 值应当接近于2.0。

4　气体－液体－固体之间的界面热力学

4.1　界面能

4.1.1　液态金属的表面能

根据吉布斯函数的基本方程

$$dG=-SdT+VdP+\sigma dA+\sum\nolimits_B \mu_B dn_B \quad (2.1\text{-}102)$$

式中，G 为自由焓；S 为熵；V 为体积；σ 为单位面积的表面能；μ_B 为 B 物质的化学位。

对于纯物质，在恒温恒压下，上式变为：

$$dG=\sigma dA$$

即

$$\sigma=\left(\frac{\partial G}{\partial A}\right)_{T,P} \quad (2.1\text{-}103)$$

σ（或用 γ 表示）即为增加单位表面积所引起的吉布斯函数的增量。也就是增加物质单位表面积所作的最大有用功。该功即等于物质单位表面积上分子所具有的比表面能，习惯上称为表面能。

显然，在恒压下，对于指定的液态金属，其 σ 应受第二相物质及温度 T 的影响。一般测定液态金属的表面能，都是处于其蒸气下测定的。

4.1.2　金属晶体的表面能

晶体表面能的概念和液体相同。但晶体与液体相比，具有不同的特性，如固体中的原子相互以价键结合，形成固定的晶形。因而晶体中的原子相对地无流动性（在室温下），不能自发地如液体那样形成平滑的球状表面。而是依据成形的条件，形成粗糙不平且具有缺陷的表面，使得表面上各处的表面能不一致。再由于在恒温恒压下晶体成形的稳定状态必须满足吉布斯函数最小原则，应具有最小的表面吉布斯函数值。因而在晶体的各个结晶面上原子排列方式不同，使得表面能具有各向异性的特点。这与液态物质表面能的特性是不同的。此外，固体成形的状态还取决于它被加工的过程。所以，这些因素给晶体表面能的研究带来复杂性，给表面能的测试增加了困难。

长期以来，很多科学工作者致力于研究金属表面能的理论计算，虽取得很大的进展，但所得结论都有其局限性。这是因为设计的理论模型很难照顾到晶体的表面畸变、点阵缺陷，以及加工应力等多变因素。致使许多固体的表面能无法直接计算。即使对于理想化晶体表面能的计算，也十分繁琐，其结果与实际情况有很大差距。因此，固体表面能的理论估算，只能作为科研工作的参考。

目前对最简单的晶体模型表面能的测试方法有解理功法、零蠕变法和经验估算法等。

晶体表面能的影响因素有多种，情况是很复杂的。影响固体表面能的因素，除了上面已讨论的晶体类型、晶面取向、表面吸附物及温度以外，还与晶体是否含有杂质、缺陷、表面粗糙程度等有关。另外，还与晶体成型的条件有关。例如，NaCl 自尿素溶液中结晶时呈八面体，而不是立方体；Ag 晶体在空气中退火时的形状与在氮气中退火时完全不同，因为前者表面有吸附氧的影响。

4.1.3　晶体的界面能

晶体界面能的研究，对于材料科学研究具有重要意义。详细讨论需要应用位错理论。在此只作简要的介绍。

一般测定界面能用热蚀法，选取一个包含有双晶粒的晶体，且已知其表面能。将它放入惰性气氛中高温加热，经一定时间后，晶体表面邻近晶界处有下陷现象产生，如图2.1-20所示。这种下陷现象的产生，使得晶界面积减小，界面能降低。晶界沟槽可由干涉显微镜测定。

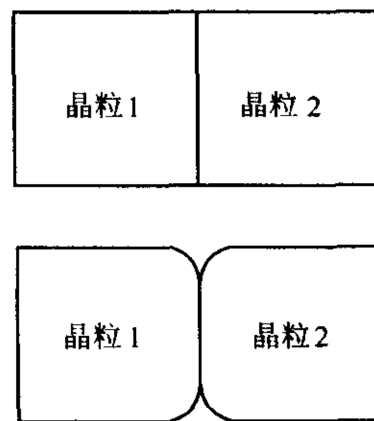


图2.1-20　热蚀法测界面张力图

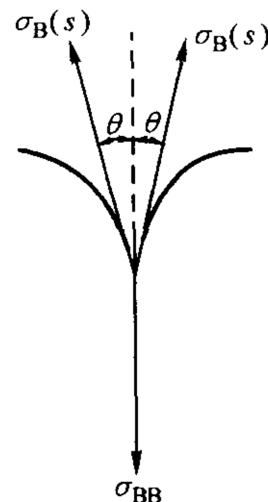


图2.1-21　界面张力与表面张力平衡图

当界面张力与表面张力呈平衡时，则力的平衡关系见图2.1-21。其关系式为：

$$\sigma_{BB}=2\sigma_B\ (s)\ \cos\theta \quad (2.1\text{-}104)$$

式中，θ 为二面角，可用干涉显微镜测定。

表2.1-5列出一些纯金属的 σ_{BB} 及 σ_B（s）的数据。从表中数据看出，σ_{BB} 与 $\Delta_s^g H_m$ 也有半定量的比例关系。凡 $\Delta_s^g H_m$ 愈高的金属，其 σ_{BB} 愈大；σ_{BB}/σ_B（s）一般在0.3左右。$\Delta_s^g H_m$（B）由量热法测定。

4.2　界面的吸附现象

液相或固相中某物质在相界面上的浓度与相内浓度产生

表 2.1-5 纯金属的 σ_B (s) 及 σ_{BB}

晶型	元素	σ_{BB} /J·m^{-2}	$\sigma_B(s)$ /J·m^{-2}	$\Delta_s^g H_m(B)$ /kJ·mol^{-1}	$\sigma_{BB}/\sigma_B(s)$
fcc	Cu	0.530	1.650±0.1	339.3	0.32
	Ag	0.400	1.140±0.1	286.2	0.35
	Au	0.400	1.400±0.1	365.3	0.29
	Pt	0.780	2.100±0.3	564.8	0.37
	Al	0.600	1.100±0.6	324.3	0.54
	γ-Fe	0.790	2.200±0.3	316.3	0.38
bcc	α-Fe	0.800	2.100±0.3	316.3	0.38
	W	1.070	2.800±0.4	836.8	0.38

差别的现象，称为吸附。吸附是进行多相化学反应的首要步骤，涉及物质的溶解、相变、金属的熔炼、焊接、热处理、腐蚀、摩擦与磨损多种过程。研究和了解吸附的规律，对于控制多相化学反应是极其重要的。

吸附按吸附物和吸附剂的相态，可分为以下3种：

1）气体在固体表面上的吸附；

2）固体在溶液中的吸附；

3）溶液表面的吸附。

结合材料专业特点，下面只讨论1）、3）两类吸附。

4.2.1 气体在固体表面上的吸附

按吸附物与吸附剂间的作用力分类，有物理吸附及化学吸附两种。物理吸附的作用力为范德瓦尔斯吸引力。吸附剂对吸附物的吸附无选择性，吸附物在吸附剂表面可以形成多分子层，吸附易于达到平衡。化学吸附的作用力为化学键力。吸附剂对吸附物的吸附具有选择性，吸附物在吸附剂表面形成单分子层，吸附不易达到平衡。

兰格缪尔（I. Langmuir，1916年）单分子层吸附理论。

其基本假定为：

1）被吸附分子在吸附剂表面上为定域吸附，而吸附剂表面是均匀的；

2）被吸附分子间无作用力；

3）吸附达到平衡时，吸附的速度等于脱附的速度；

4）吸附剂表面吸附能力是均匀的，与被遮盖的表面积多少无关；

5）吸附是单分子层。

应用经典热力学推导的兰格缪尔方程为

$$\Gamma = \Gamma_\infty \frac{bp}{1+bp} \tag{2.1-105}$$

式中，Γ 代表气相平衡压力为 p 时的吸附量（即单位质量吸附剂吸附气体的物质的量或体积）；Γ_∞ 代表单位质量吸附剂表面全部铺满单分子层时的吸附量。

令
$$\theta = \frac{\text{已被吸附气体的面积}}{\text{吸附剂的总表面积}} = \frac{\Gamma}{\Gamma_\infty} \tag{2.1-106}$$

则
$$\theta = \frac{bp}{1+bp} \tag{2.1-107}$$

式（2.1-107）中的常数 b，是反映吸附剂和吸附物特性的物理量，可通过实验求得。如将式（2.1-105）改写成直线方程为

$$\frac{1}{\Gamma} = \frac{1}{\Gamma_\infty} + \frac{1}{\Gamma_\infty bp} \tag{2.1-108}$$

式中的常数 Γ_∞ 及 b，可通过 $1/\Gamma$ 对 $1/p$ 作图求得。

4.2.2 溶液表面的吸附

（1）液体界面的结构

液体与蒸气处于平衡时，液相分子蒸发的速度等于蒸气分子凝结的速度。根据气体分子运动论可知，在室温下，气体每秒钟碰撞单位液面的分子数约为 1.2×10^{22} 个·cm^{-2}·s^{-1}。因此，在气液交界面上，在 1 cm^2 的界面上，每秒钟约有 1.2×10^{22} 个分子进进出出。

而在液相内部，多数分子的扩散系数的数量级为 10^{-5} cm^{-2}·s^{-1}。爱因斯坦（Einstein）导出了扩散系数 D 和时间 t 以及发生位移 x 的关系是

$$D = x^2/(2t)$$

若取 x 为 10^{-8} m，则 t 约为 10^{-6} s。因此，表面与邻近体相区间分子的交换也是极快的。这是布朗（Brown）运动的结果。

由此可见，从微观上观察，液体表面是处于剧烈扰动的状态。分子无规律地在气液二相之间运动。如果在适当放大的显微镜下观察，气液交界面并不是一个几何分界面，而是由液相到气相的一个过渡区，在表面化学中称为表面区。

（2）吉布斯恒温吸附式

在恒温恒压下，体系由 α 及 β 二相组成，各相皆由溶剂 A 及溶质 B 混合而成。B 物质的吸附量

$$\Gamma_B = -\frac{1}{RT}\frac{d\sigma}{d\ln a_B} = -\frac{a_B}{RT}\frac{d\sigma}{da_B} \tag{2.1-109}$$

或
$$\Gamma_B = -\frac{c_B}{RT}\frac{d\sigma}{dc_B} \tag{2.1-110}$$

式中，a_B 为 B 物质的活度；c_B 为 B 物质的量浓度。上二式即为吉布斯恒温吸附式。

当 $d\sigma/dc_B>0$ 时，$\Gamma_B<0$。则溶质在界面层的浓度小于在溶液内部的浓度。这种现象称为负吸附。这类溶质称为溶剂的非表面活性物质。

当 $d\sigma/dc_B<0$ 时，$\Gamma_B>0$。则溶质在界面层的浓度大于在溶液内部的浓度。这种现象称为正吸附。这类溶质称为溶剂的表面活性物质。

4.3 液体对固体的润湿性

在材料成形、火法冶金及粉末冶金过程中，液-固界面的润湿现象是很重要的。如非金属夹杂物质与金属液相的分离；金属液、熔盐及炉渣对炉衬和电解槽的浸蚀；粉末冶金的烧结；金属电解的沉积等等，都与液体对固体的润湿性有关。

例如，粉末冶金的烧结过程，液相对固相必须有良好的润湿性。否则，烧结时，液相如同“出汗”一样从固相中挤出，达不到液相烧结的作用。另外，液相对固相润湿性的好坏直接影响着固相的结晶长大、相分布及黏结强度。因此，在采用粉末冶金法制备硬质合金时，应选择润湿性好的黏结金属。

又例如，电解法制备镁时，如果电解液对铁阴极或钢阴极润湿性良好，就会妨碍金属镁在阴极上的停留，则阴极上生成的细小镁液滴不能聚合，阻碍生产。故生产上必须向电解槽中加入减小电解液对电极润湿性的物质，促进镁在阴极上的析出。

在将陶瓷颗粒加入液态金属中制备复合材料时，为了使颗粒能够进入液态金属，都是通过改变液相对固相的润湿性来实现的。

液体对固体润湿性的大小，直接和液-气、固-气及液-固的界面能有关。

4.3.1 纯金属的液-固界面张力

纯液态金属凝固时，产生液-固（l-s）相界面，其界面张力 σ_{ls} 对形核有很大影响，从形核的实验中，可推算出 σ_{ls}。

特赫尔（Turnhull）等应用形核实验，总结出计算 σ_{ls} 的公式：

$$\sigma_{ls} = 0.45\Delta_s^l H_m / A_s \tag{2.1-111}$$

式中，A_s 为 1 mol 金属原子铺成单原子层所具有的面积，其单位为 m^2·mol^{-1}，称为摩尔原子面积。其计算式为：

$$A_s = bL^{\frac{1}{3}} V_m^{\frac{2}{3}} \tag{2.1-112}$$

式中，b 为单层原子排列系数。对于六方紧密排列的结构

(hcp)、体心立方（bcc）及面心立方（fcc）晶体，b 分别为1.09、1.12及1.12；L 为阿伏加德罗常数；V_m 为固态金属的摩尔体积。

某些金属利用式（2.1-111）计算的 σ_{ls} 列于表2.1-6中。

表2.1-6　某些纯金属的 σ_{ls}

金　属	$\sigma_{ls}/J\cdot m^{-2}$	金　属	$\sigma_{ls}/J\cdot m^{-2}$
Hg	0.028 1	Au	0.132 0
Ga	0.056 0	Cu	0.177 0
Sn	0.059 0	Mn	0.206 0
Bi	0.054 4	Ni	0.255 0
Pb	0.033 3	Co	0.234 0
Sb	0.101 0	Fe	0.204 0
Al	0.093 0	Pd	0.209 0
Ag	0.126 0	Pt	0.240 0

4.3.2　液体对固体的润湿性

研究液体对固体的润湿性，常用润湿角来讨论，如图2.1-22所示。图2.1-22b中表示液滴在固相上铺展的形态，θ 称为接触角。接触角是从液（liquid）、气（vapor）、固（solid）三相平衡点作液-气表面的切线，切线与液-固界面的夹角。θ 的大小也反映液体对固体的润湿性，以及液滴在固相上所处的形貌。

液滴在固相上的铺展有三种情况，如图2.1-22所示。从图2.1-22a看出：$\theta=0°$，液体能完全润湿固相。液体将固相上吸附的气体全部排除，且完全铺展于其上；图2.1-22b表示：$\theta<90°$，液体部分地铺展于固体表面上，固体部分被润湿；图2.1-22c表示：$\theta>90°$，一般认为液体对固体润滑性差，$\theta=180°$时液体团聚成球状。

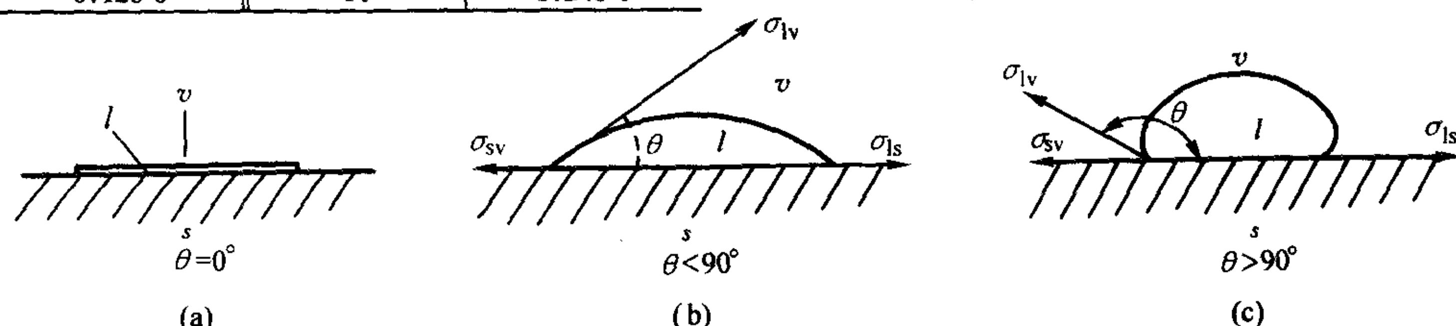


图2.1-22　接触角与润湿性

当然，润湿性大小是一个相对的概念，如 θ 等于100°的液体应比 θ 等于120°的液体的润湿性好。

液体能否在固体上产生润湿现象，取决于液体与固体和液体自身质点间的作用力。当前者的作用力大于后者时，液体能润湿固体。否则，则不能润湿固体。它们之间的差异，集中反映在接触角的大小上。

(1) 杨氏（Young）方程

由图2.1-22b及图2.1-22c可看出，当界面张力处于平衡时，σ_{ls}、σ_{sv}及 σ_{lv}之间具有以下的关系

$$\sigma_{sv}=\sigma_{ls}+\sigma_{lv}\cos\theta$$

或

$$\cos\theta=\frac{\sigma_{sv}-\sigma_{ls}}{\sigma_{lv}} \tag{2.1-113}$$

式（2.1-113）即为Young方程。由于 σ_{sv}及 σ_{ls}难以测定，限制了该方程的直接应用。但通过它仍可作出一些有意义的推论。

当液体完全润湿固体时，$\theta=0°$，$\cos\theta=1$。则 $\sigma_{sv}-\sigma_{ls}=\sigma_{lv}$；

当液体部分润湿固体时，$0°<\theta<90°$，$1>\cos\theta>0$。则 $\sigma_{sv}-\sigma_{ls}<\sigma_{lv}$或 $\sigma_{sv}<\sigma_{ls}+\sigma_{lv}$；

当液体不润湿固体时，$\theta>90°$，$\cos\theta<0$。则 $\sigma_{sv}<\sigma_{ls}$。

从以上分析看出，当 $\sigma_{ls}<\sigma_{lv}$时，液相容易润湿固相；当 $\sigma_{sv}<\sigma_{ls}$时，液相不能润湿固相。对于具体的体系，为改善其润湿性，应采取措施降低 σ_{ls}或提高 σ_{sv}。

(2) 影响润湿性的因素

液相对固相表面的润湿性，可通过Young方程进行分析。而 θ 要受到 σ_{lv}、σ_{sv}及 σ_{ls}三因素的影响。

1) 影响 σ_{sv}的因素　对于指定体系的固体，表面有无吸附膜，直接影响 σ_{sv}。

根据Young方程，方程中的 σ_{sv}，是将固体放置于蒸气中的表面能。由于固体对蒸气的吸附，致使固体表面上带有吸附膜，其 σ_{sv}不同于除去蒸气后的固体表面能 σ_{so}。也就是说，固体处于真空下的 σ_{so}不同于 σ_{sv}。

由于有蒸气吸附于固体表面上，使其表面能下降，即

$$\sigma_{so}>\sigma_{sv}$$

令

$$\sigma_{so}-\sigma_{sv}=\pi \tag{2.1-114}$$

式（2.1-114）中的 π 称为吸附膜的铺展压。因此，固体表面如有蒸气的吸附膜，对接触角是有影响的。

设固体处于真空下，液体在其上铺展的接触角为 θ_0，根据Young方程，可写成

$$\cos\theta_0=\frac{\sigma_{so}-\sigma_{ls}}{\sigma_{lv}} \tag{2.1-115}$$

显然 θ_0 小于固体处于蒸气下的 θ。式（2.1-115）表明，在真空下测定的 θ_0 小于非真空下所测定的 θ。许多液体金属在各种气氛及真空下对陶瓷材料表面的 θ 值，表现出同样的规律性。

此外，固体表面的粗糙度，也可改变液体在固体表面的润湿性。因为固体表面粗糙时，其真实表面比表观表面大 β 倍，如图2.1-23所示。当液固界面向前推进微量距离时，即增加固相表面积为 dA，同时也增加了液相的表面积 $dA\cos\theta'$（θ'为粗糙表面的接触角）。

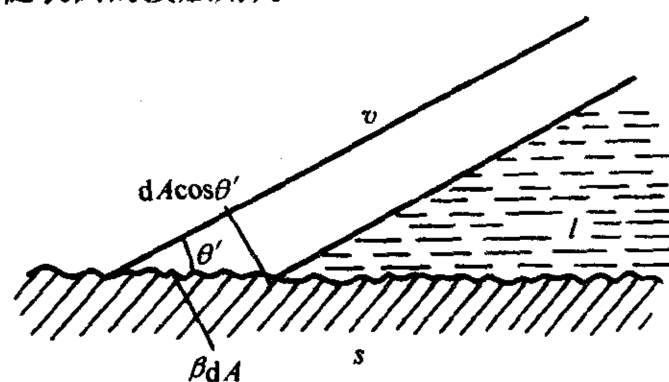


图2.1-23　表面粗糙对 θ 的影响

如图2.1-23所示，根据吉布斯函数最小原则，体系处于平衡时，应满足以下条件

$$\sigma_{ls}\beta dA+\sigma_{lv}dA\cos\theta'-\sigma_{sv}\beta dA=0$$

所以

$$\cos\theta'=\frac{\beta(\sigma_{sv}-\sigma_{ls})}{\sigma_{lv}}=\beta\cos\theta \tag{2.1-116}$$

从式（2.1-116）看出，粗糙表面 θ'的余弦值比光滑表面的 $\cos\theta$ 大 β 倍。即是说

$$\theta'<\theta$$

由此可见，当表面粗糙程度增加时，液相对固相的润湿性亦增加。

以上分析，仅从 θ 的大小来讨论液体对固体的润湿性。实际上，液体在固体表面铺展，还会遇到固体表面的摩擦阻力 ψ，阻碍着液体的铺展。所以，全面考虑，Young 方程应写成以下形式

$$\sigma_{sv}=\sigma_{ls}+\sigma_{lv}\cos\theta''+\psi$$

式中，θ''称为阻滞接触角。则

$$\cos\theta''=\frac{\sigma_{sv}-\sigma_{ls}}{\sigma_{lv}}-\frac{\psi}{\sigma_{lv}} \tag{2.1-117}$$

显然 $\theta''>0$

只有当 ψ 为零时，$\theta''=\theta$。故从有利于液相润湿固体看，固体表面愈平滑愈好。

2) 影响的 σ_{lv}因素　主要与温度及化学成分有关，此处不再赘述。

3) 影响 σ_{ls}的因素　凡是影响 σ_{sv}及 σ_{lv}的因素，也是影响 σ_{ls}的因素，情况很复杂。总的来讲有以下几个方面。

① 液相的成分。例如，在 1 475℃，液态 Ni 与固相 Al_2O_3 接触，其 σ_{ls}随 Ni 液中合金含量而变化。

② 表面吸附。有些合金元素为液相的表面活性物质，有降低液 - 固界面张力的作用。例如，1 100℃时 Cu 液在 ZrC 上的 θ 等于 135°。当 Cu 液中加入的 Ni 为 0.25%时，θ 即下降至 54°。

③ 固体在液体中的溶解。例如，液态的 Fe、Co 及 Ni 与石墨接触，由于石墨在液态金属中有一定的溶解度，则会引起 σ_{ls}的变化。

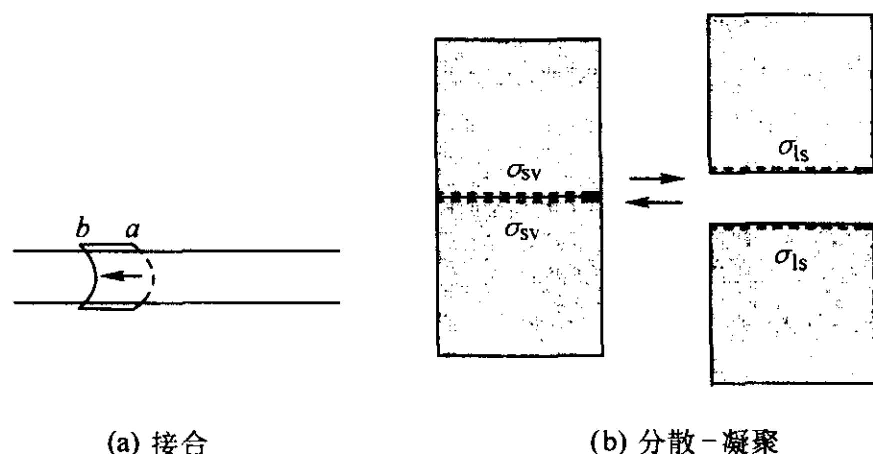
④ 液相与固体发生化学反应。例如，Al 液与 MgO 陶瓷发生反应，生成 $MgO\cdot Al_2O_3$ 化合物，接触的固相表面性质发生了变化，引起 σ_{ls}的变化。

从以上分析来看，影响润湿性的因素是很复杂的。σ_{sv}、σ_{lv}及 σ_{ls}之间存在着很复杂的依赖关系，必须通过实验进行具体的分析。

4.4 材料工程与润湿

4.4.1 附着与接合

液体与固体相接合时（液体随后变为固体的情况也包括在内）首先必须附着在固体表面上。当然，这里也包括锡焊、钎焊和其他焊接等概念。对于材料的接合或连接成形，润湿是不可欠缺的。考虑到通常的接合是在两种固体之间进行的，我们用图 2.1-24a 所示的模型进行说明。在图中由点 a 到点 b，向两个固体表面之间（也可以是固体中的孔的情况）浸入液体。这时与固 - 气界面转变为固 - 液界面相对应，界面能的变化可用（$\sigma_{sv}-\sigma_{ls}$）表示。如果 $\sigma_{sv}>\sigma_{ls}$，由于润湿使总界面能下降，液体浸入固体中。由式（2.1-113）可知，接触角 θ 必定小于 90°。由此得出，$\sigma_{sv}>\sigma_{ls}$是接合的必要条件。这样的润湿形态也叫浸渍润湿，其功称为浸渍润湿功。



(a) 接合　　(b) 分散 - 凝聚

图 2.1-24　利用润湿的接合与分散 - 凝聚模型

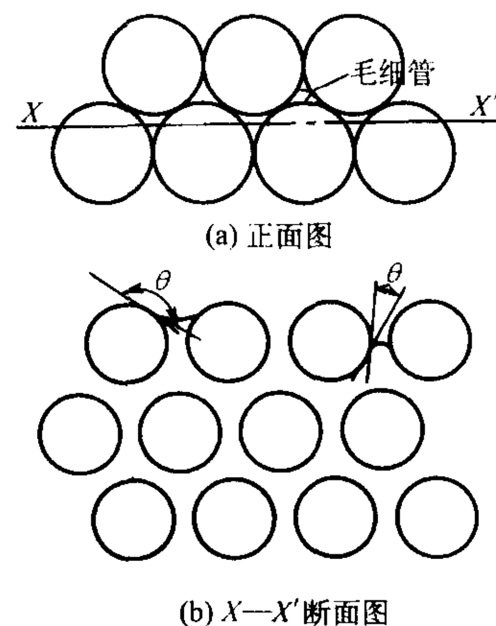
液体中微小固体（粉体）的分散 - 凝聚也是同样原理。立方体形的微小固体在液体中的凝聚以及分散的模型如图 2.1-24b 所示。在这种情况下，固体粒子间浸入液体就呈分散态。与此相同现象有，熔液中非金属夹杂物的分散 - 凝聚，或与铸型壁或坩埚壁的附着等。在这些场合，θ 在 90°以下时产生分散，在 90°以上时发生凝聚。

4.4.2 材料液态成形与润湿

(1) 砂型与熔液的润湿

铸造与润湿有很大的关系。用砂型可以铸出铸件的原因，不是由于熔融金属凝固于砂子表面，实际上是砂子与熔融金属之间不润湿，从而熔融金属不会浸入砂子而能形成铸件。如果二者之间润湿良好，则熔液浸入砂型中成为一种复合材料，产生所谓“黏砂”缺陷。

如果从砂型的表面看，砂粒间的间隙极少。但是，若从断面考虑，间隙并不少。例如，将砂粒看成相同直径的球体，可表示为图 2.1-25 所示的模型。这一排列，若看成玻璃珠子的堆积就容易理解。若表面到底部的球都是相同的排列，从图 2.1-25a 中取 $X-X'$ 断面得到图 2.1-25b。可能多少有些夸张，可看到铸型中的空间是如此之多。砂的表观容积为 0.8～0.6 cm^3/g，充填率为 54%～73%。因此，铸型的体积的 27%～46%是空气。从润湿的角度看，如果接触角 θ 在 90°以上，熔液不能够浸入砂型的间隙中（砂型的空间可看成图 2.1-25a 中虚线所示的毛细管），能够制造出良好的铸件。反之，若润湿性好，如图 2.1-25b 的右侧所示，熔液在界面能的作用下很容易地浸入砂型中。



(a) 正面图

(b) $X—X'$断面图

图 2.1-25　砂型的空隙模型

通常认为氧化铁的存在会促进砂子与液态铁的润湿。为了防止黏砂，在砂中加入碳素物质或在铸型表面涂上碳素涂料，以保持砂型中的还原性气氛，防止氧化铁的生成，以恶化熔融铁与砂的润湿性。氧化铁的作用在最近的润湿角测定的研究结果中得到证实，如图 2.1-26 所示。在静滴法测定液态铁与 SiO_2 的接触角时，导入氧气后接触角急剧减小。这一现象被认为是产生黏砂缺陷的根本原因。

(2) 夹杂物、气泡与润湿性

砂眼、夹渣等铸造缺陷也是同样，可用润湿性来说明。例如，浇注时会有很多的型砂、熔渣和空气一起被卷入型腔，而使之与熔融金属的分离变得很困难。这里，润湿与否起着相当重要的作用。在这种情况下，砂子、熔渣与熔融金属之间不润湿成为重要条件。如图 2.1-27 所示，非金属夹杂物与熔液不润湿时，夹杂物与气体接触的界面能低，更安定，到达熔液表面的夹杂物不会再移动进入熔液。同样，当向熔液中吹入气体时，夹杂物会与气泡吸附在一起，容易从熔液中去除。如图 2.1-27a 的情形，润湿性差的物质与气泡共存，容易从液态金属中上浮并分离。

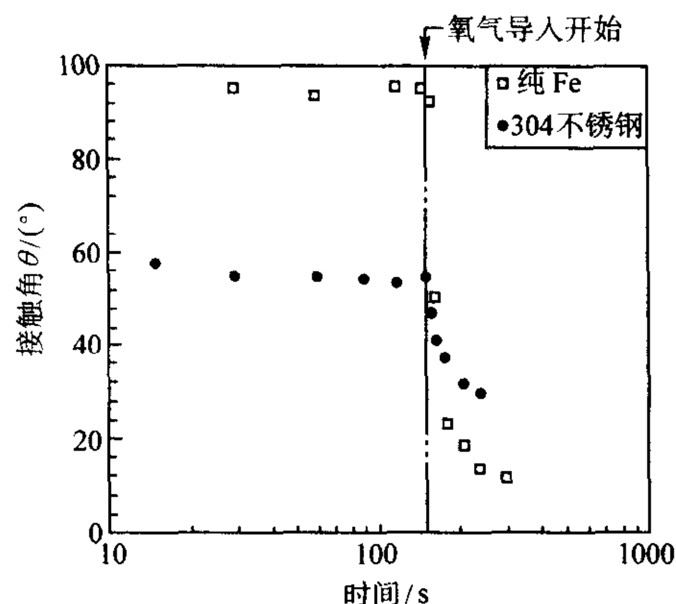


图 2.1-26 氧气对熔融铁与 SiO_2 玻璃的润湿性的影响

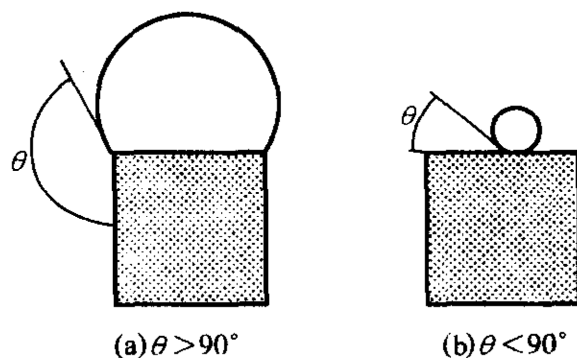


图 2.1-27 熔融金属中夹杂物的行为与润湿

球化处理及除气处理时使用的多孔塞，则需要良好的润湿性，如图 2.1-28 所示。润湿性不好的话（固 - 液界面不安定，固 - 气界面安定），吹入的气体的气泡就会在管的外部形成，而成为大泡；润湿性良好的话，气泡就会产生于管的内侧而变得较小。多孔塞的耐火物中的小孔无论多么小，如果它与熔融金属的润湿性不好，都不会形成小泡。

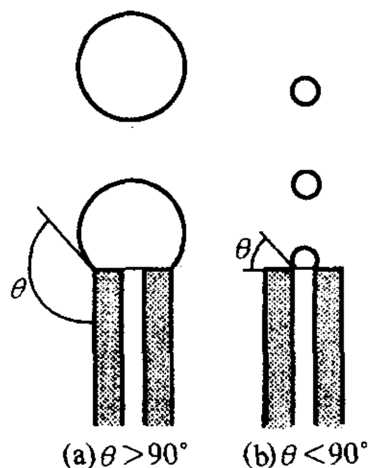


图 2.1-28 多孔塞的气泡生成与润湿

4.4.3 复合材料的制造与界面能

（1）陶瓷材料与液体金属的润湿性

以下所提到的复合材料制造法仅限于熔融金属法。通过熔融金属制造复合材料的方法大略可分成两大类。第一类是向液态金属中添加颗粒或纤维的方法（以下简称颗粒添加法）。第二类是向预先制成的陶瓷等多孔性材料做成的预制体浸渍熔融金属的方法（以下简称浸渍法）。

颗粒添加法中，颗粒与液体金属的润湿模型如图 2.1-29 所示。对于添加方法及润湿性的研究表明，如果添加的颗粒不能与熔融金属完全润湿的话，颗粒就不可能自然浸入熔融金属。这里所言及的颗粒的尺寸是相当小的，大约在 0.01 ~ 10 mm 之间。当尺寸较大时，则属其他范畴的问题了，亦即与船的沉浮（浮力原理）相当。

当 $\theta = 0°$时，液面就会润湿到球形颗粒的顶面上来，即颗粒自然进入液体中。当 $0° < \theta < 90°$时，液面可润湿到球的

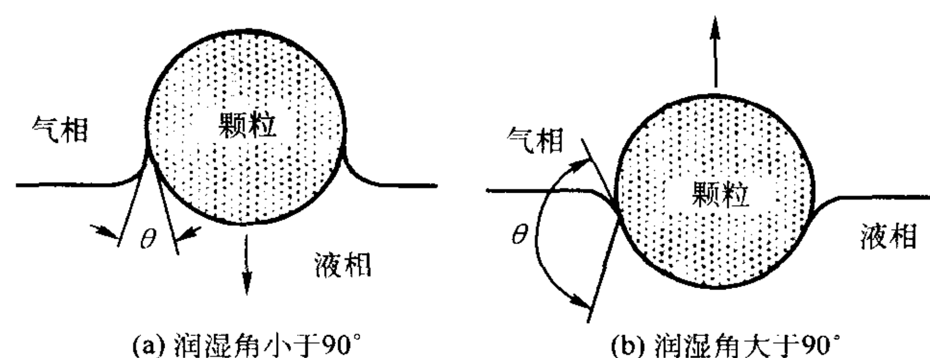


图 2.1-29 陶瓷颗粒与液态金属的润湿模型

上半球部，最后部分由于浮力增大，仍需要外力的作用才能浸入熔液中。一般情况下，陶瓷颗粒在金属液中不易润湿（$\theta > 90°$），通常需要借助外力，比如喷射法和搅拌法等来制造复合材料。

对于预制体的浸渍法，这个问题的模型如图 2.1-24a 所示。图 2.1-24a 可视为把预制体的孔洞模型化了。如果润湿不好的话，熔融金属就会停止在预制体的表面，即在图 2.1-24a 孔的入口停止了。

（2）陶瓷颗粒与凝固界面的相互作用

陶瓷颗粒增强型金属基复合材料的液态制备方法具有工艺简单，易于使成本降低及便于实现大规模工业化生产等优点。但是，液态法中虽然可通过搅拌等方法使外加的陶瓷颗粒进入液相并在液体金属中分布均匀，而在凝固后的固相中颗粒往往偏聚在晶粒边界，形成不均匀分布。增强颗粒在固相中的不均匀分布将对复合材料的强度及耐磨性等产生不利影响。

增强颗粒能否在固相中均匀分布取决于陶瓷颗粒与凝固界面的相互作用结果。此时固相 - 液相 - 颗粒三相间的界面能模型如图 2.1-30 所示。此模型与图 2.1-29 所示的模型相似，但三相的性质发生了变化。当颗粒与成长的固相之间的接触角 θ 小于 90°时，即颗粒与固相的界面能 σ_{SP} 小于颗粒与液相的界面能 σ_{PL} 时，颗粒与固相的接合更安定，颗粒将被凝固前沿所捕捉而进入固相；反之，当 θ 大于 90°时，即颗粒与固相的界面能 σ_{SP} 大于颗粒与液相的界面能 σ_{PL} 时，颗粒在液相中更安定，颗粒被凝固界面所排斥和推移。

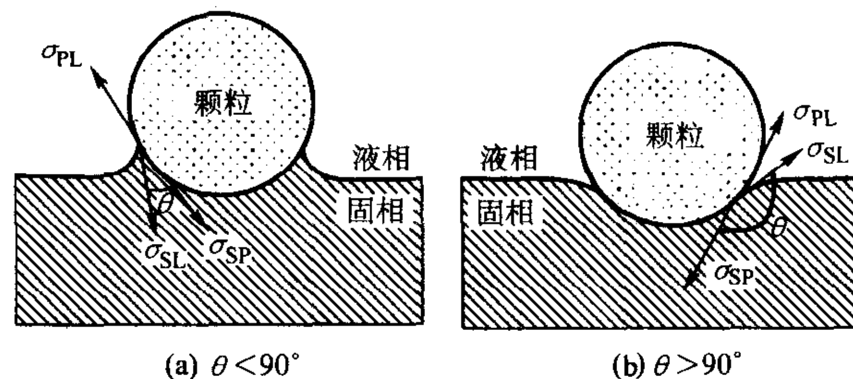


图 2.1-30 固相 - 液相 - 颗粒三相间的界面能模型

大多数的陶瓷颗粒与固体金属的界面能都较高，因此在凝固过程中都被凝固界面所排斥和推移，在凝固前沿聚集，或被生长的胞晶或枝晶机械地卷入晶粒的边界分布，如图 2.1-31 所示。此复合材料的基体为 Al - 0.2% Ce 合金，陶瓷颗粒为 20 μm 左右的 Al_2O_3 颗粒。在 Al - Si 合金中，Al_2O_3 颗粒也同样被排斥而呈不均匀分布。但是，当向 Al - Si 熔液中添加某些活性元素后，活性元素改变了 Al_2O_3 颗粒与液体和固体的界面能关系，并且 Si 相能在 Al_2O_3 颗粒的表面生长，如图 2.1-32 所示，说明 Al_2O_3 颗粒与固相的界面能小于与液体的界面能，Al_2O_3 颗粒被生长的固相所捕捉而进入固相，在固相中形成均匀分布。

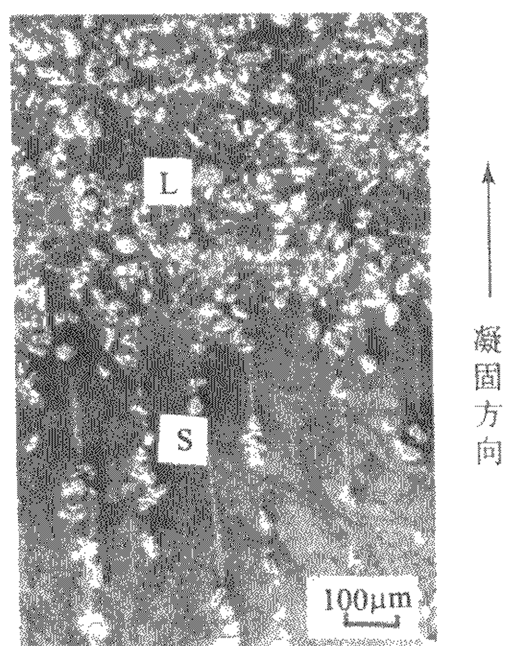


图 2.1-31 Al_2O_3 颗粒被凝固界面所排斥聚集于固液界面前沿

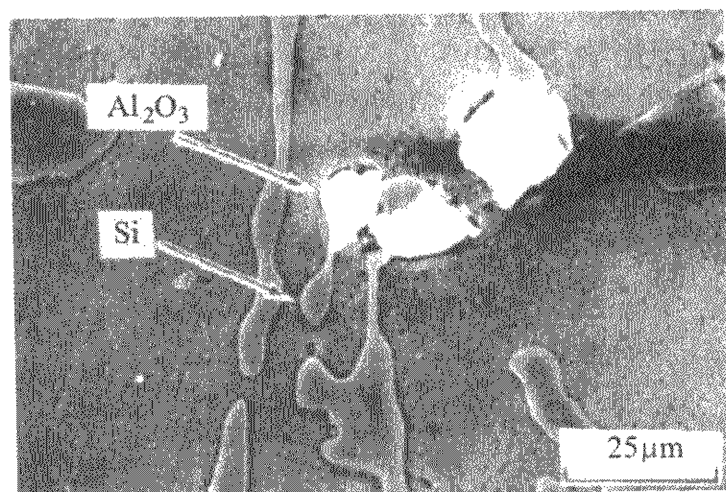


图 2.1-32 Al_2O_3 颗粒被生长的固相捕捉

5 燃烧反应的热力学及动力学

燃烧反应不仅为合金熔化等材料加工过程提供必要的高温条件，而且其产物和反应物还和加工的材料进行化学反应。因此，研究燃烧反应对研究材料加工过程具有重要意义。

目前用的燃料主要是有机燃料，如煤、石油制品、天然气等。近年来核燃料也日益被重视。

燃料的成分：固体和液体燃料中的水分 W 和灰分 A 是不能燃烧的惰性物质，除灰、水以外的物质称为燃质。燃质主要由碳、氢、氧、氮、硫构成。对煤来说，挥发分 V 也是一个重要组分。

5.1 燃料的着火

着火是燃烧的准备阶段，燃料氧化速度还不足以产生光和火焰。氧化时释出热量的一部分散失到外界，另一部分则在燃料内部积蓄起来，使过程强化。随着过程的发展，在释热超过散热时就使燃料着火。这种与热力工况有密切联系的着火称热力着火。

一些常用燃料的着火温度为：氢气 530 ~ 590℃；甲烷 645 ~ 850℃；乙炔 335 ~ 500℃；天然气 530℃；汽油 410 ~ 560℃；褐煤 450℃；焦炭 600 ~ 700℃。

5.2 气体燃料的燃烧

气体燃料燃烧使用的主要设备是混合器——烧嘴。

燃烧分预混燃烧及扩散燃烧两种。每一种又分成层流燃烧和湍流燃烧。预混燃烧广泛用在工业的无焰燃烧设备上。均匀混合物先在局部地方着火燃烧，该燃烧层的释热传递给邻近层，使着火一层一层地传播下去。火焰在火焰锋面法线方向向未燃气体传播的速度称为火焰传播速度（W），它也是单位时间内单位火焰面积上烧掉的气体体积，故也称燃烧速度。

预混可燃气体处于层流状态时，燃烧释热传递主要靠导热，这时的火焰传播速度称为正常火焰传播速度 W_n。

$$W_n \propto \sqrt{\frac{\lambda}{c_p \rho}}$$

可燃气体浓度改变时 W_n 发生变化，并存在着浓度上、下限，在上、下限以外火焰不能传播，W_n 有一最大值。若超过着火浓度极限，燃烧蒸汽与空气的混合物会发生爆炸。

工业上为强化燃烧多用湍流燃烧，湍流燃烧时热的传递靠分子微团的不规则运动，因此热交换强化，火焰传播速度 W_t 增加。实验证实，小规模湍流（$Re = 2\ 300 \sim 5\ 000$）时 $W_t / W_n \propto Re^{1/2}$。大规模湍流时已被折皱的火焰面更加凹凸不平，$W_t \propto Re$。在非常强烈的湍流情况下火焰面完全碎裂，从而大大增加了火焰表面积，使燃烧速度又增加。实验证实，在 $Re > 1 \times 10^5$ 时，对一般粗糙度管子，$W_t \approx (0.01 \sim 0.05)\ W_g$，对粗糙度大的管子 $W_t \approx (0.1 \sim 0.15)\ W_g$，$W_g$ 是气流速度。图 2.1-33 是大规模湍流时 W_t 的实验曲线。

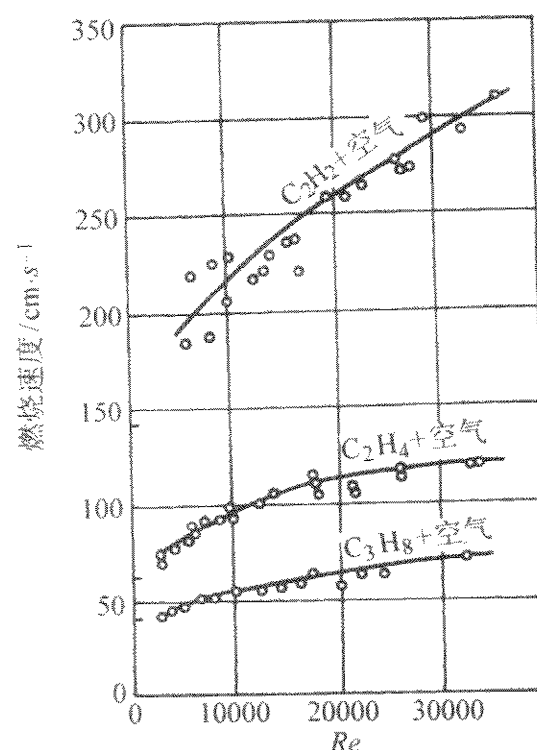


图 2.1-33 湍流时的 W_t

（纵坐标上标记"－"横的数值为层流时的 W_n）

如果燃料和氧化剂分别送入燃烧室，依靠扩散作用获得燃烧所需氧气，这种燃烧称扩散燃烧。它又分为层流扩散燃烧及湍流扩散燃烧（见图 2.1-34）。

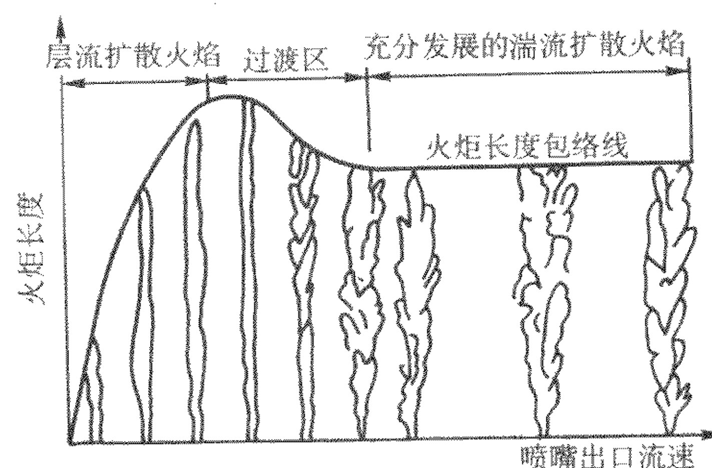


图 2.1-34 层流扩散燃烧和湍流扩散燃烧示意图

层流扩散火焰的前峰面很薄。层流扩散火炬的长度与扩散系数成反比，而流速越大、喷嘴越粗，则火炬越长。

工业上常用湍流扩散燃烧，其火炬长度与湍流扩散系数成反比，与流速关系较小，喷嘴直径加大时火炬长度增加。

5.3　液体燃料的燃烧

液体燃料燃烧使用的主要设备是雾化器——喷嘴。

液体燃料的沸点低于着火点，受热后先气化，故液体燃料的燃烧实际是单相反应。要强化燃烧应增加燃烧蒸发表面，因此工业上都使液体燃料在雾状油滴状态下燃烧。图2.1-35表示油滴燃烧的情况，燃料蒸汽向外扩散，氧气向内扩散，燃烧产物则分别向内外扩散。燃烧释热依靠辐射及燃烧产物扩散传递给油滴促使蒸发，还有部分热量传给外界，所以在油滴周围形成一球状扩散火焰，空气流速增长时油滴燃烧速度增加。影响油滴燃烧速度的主要因素是油滴半径 r，燃烧所需时间 $\tau = r^2/\psi$，燃烧速度系数 ψ 的值，汽油在800℃时为1.45。所以油滴越细，油滴尺寸在油雾炬中的分布越均匀，则燃烧越剧烈。

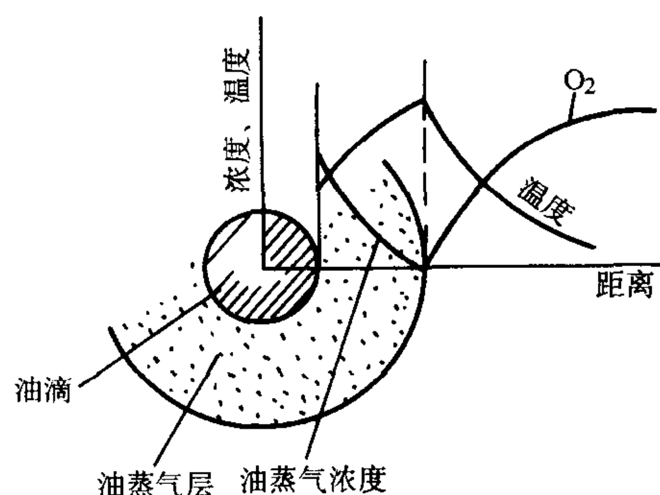


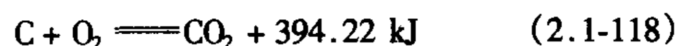
图2.1-35　油滴燃烧示意图

5.4　固体燃料的燃烧

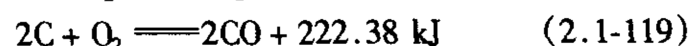
5.4.1　燃烧热力学

固体燃料的燃烧是多相燃烧，由一系列复杂的物理化学过程组成。煤粒加热时首先干燥，接着析出气态可燃物(挥发分)。挥发分析出后剩余部分称焦炭。如温度条件满足，挥发分先着火，焦炭同时被加热，最后焦炭也进入炽烈的燃烧。挥发分析出和着火、焦炭形成等阶段虽大体上有先后，然而各阶段是相互联系重叠进行的。煤的含灰量越高、挥发分越低，煤粉气流的火焰传播速度越低。故多灰分、低挥发物煤着火要困难些。

目前认为，碳的燃烧过程存在所谓一次反应和二次反应。一次反应为

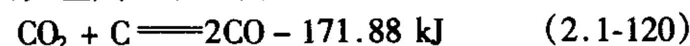


$$C + O_2 = CO_2 + 394.22\ kJ \qquad (2.1\text{-}118)$$



$$2C + O_2 = 2CO + 222.38\ kJ \qquad (2.1\text{-}119)$$

式(2.1-118)为完全燃烧反应，式(2.1-119)为不完全燃烧反应。而二次反应则可表示为



$$CO_2 + C = 2CO - 171.88\ kJ \qquad (2.1\text{-}120)$$



$$2CO + O_2 = 2CO_2 + 566.09\ kJ \qquad (2.1\text{-}121)$$

在燃烧过程中，一、二次反应重叠交错地进行，如图2.1-36，碳粒表面温度越高，CO_2 热分解的还原反应进行得越快，所产生的CO在碳粒附近又与扩散来的氧气发生反应，重新产生第二次的 CO_2。所以在 $a>1$ 和 $a<1$ 的界面附近出现一个 CO_2 浓度最高区域。然后 CO_2 再向两个方向扩散，重复原来的反应。因此在较高温度下二次反应将成为主要反应。

5.4.2　燃烧动力学

固体燃料燃烧所需时间主要取决于焦炭的燃烧时间 τ_c，它又和氧气扩散到碳表面所需时间 τ_d 及化学反应必需时间 τ_r 有关：

$$\tau_c = \tau_d + \tau_r$$

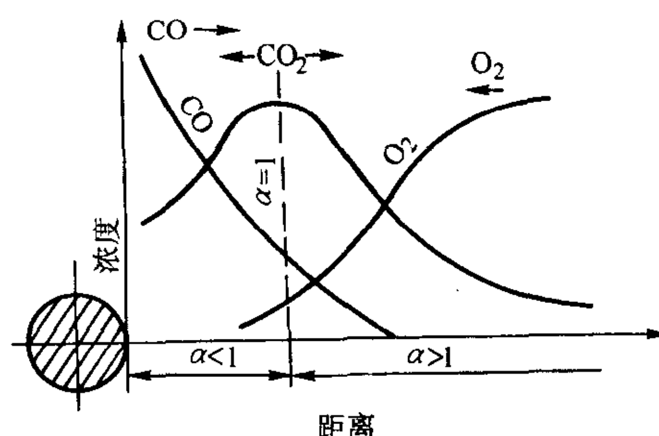


图2.1-36　高温下碳粒燃烧

依照扩散和化学反应条件的不同，固体燃料的燃烧分为三种情况。

1）扩散燃烧　化学反应能力大大超过扩散能力，燃烧速度取决于扩散能力，而与化学反应速度及温度条件关系不大。

2）动力燃烧　扩散能力大大超过化学反应能力，燃烧速度取决于化学反应。

3）过渡燃烧　介于扩散燃烧与动力燃烧之间，燃烧速度既与化学反应能力也与扩散条件有关。

固体碳燃烧动力学研究表明，影响碳燃烧速度的主要因素是温度、氧气浓度及气流速度，如图2.1-37所示。

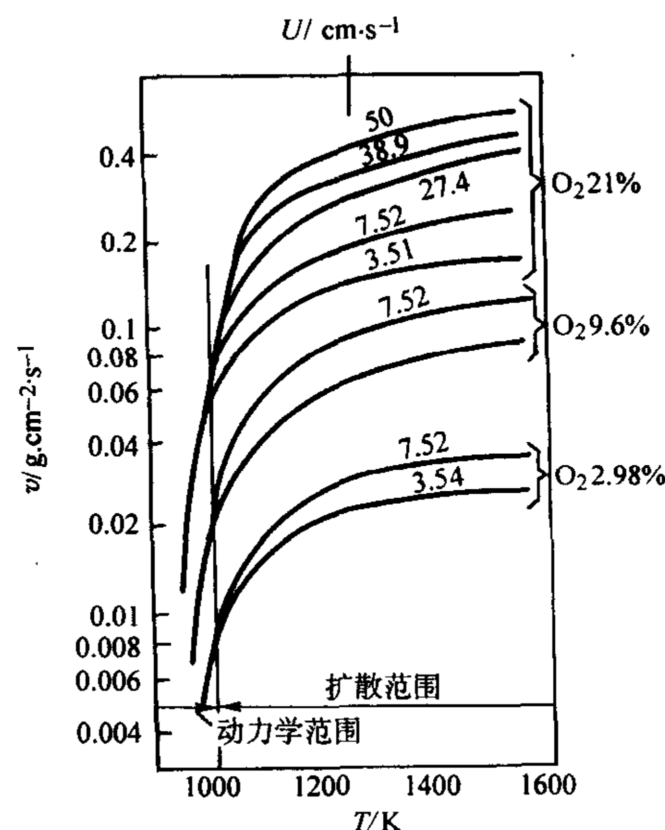


图2.1-37　球状碳燃烧速度

1）温度　温度影响传质系数或扩散系数 D，从而影响燃烧速度。D 与温度的关系如下

$$D = D_0\left(\frac{T}{273}\right)^n$$

式中 $n = 1.5 \sim 2$。$T = 273$ K时，$D_0 = D$，所以 D_0 是0℃时的扩散系数。由此可见，温度升高，O_2 扩散系数增大，故能提高燃烧速度。

2）O_2 浓度　提高气相内氧的浓度，则增大其在气膜内的浓度差，根据菲克扩散定律，就会加快 O_2 向固体碳表面扩散，提高燃烧速度。图2.1-37的实验数据同样说明这一结论。

燃烧时富氧送风，一方面提高了 O_2 浓度，同时又能提高理论燃烧温度。这两个因素都有利于提高焦炭燃烧速度。

3）气流速度　燃料表面气流的有效边界层厚度 δ_e 与雷诺数的关系为

$$\delta_e = \frac{B}{\sqrt{Re}}$$

式中，B 为常数。气流速度越大，Re 越大，δ_e 越小。因此提高气流速度，附着在固体碳表面的气膜厚度就会受到气流的冲刷而减薄，从而缩短了 O_2 向固体碳表面扩散的途径，提高了燃烧速度。从图 2.1-37 可见，在扩散范围内，燃烧速度随着气流速度的增大而提高。

煤粉在较高温度下才进入扩散燃烧区，如 0.1 mm 的碳粒在 1 700℃左右进入扩散区，而 10 mm 的碳粒在 900℃以上就进入扩散区。煤粉炉中，只有在高温的火炬中心才可能是扩散燃烧。在高温中心以外，特别是炉膛出口附近都是过渡燃烧区或动力燃烧区。故提高煤粉与气流的相对速度和提高温度均可使燃烧更完全。对于层燃炉，煤块多在扩散区燃烧，增加空气流速就能强化燃烧。

5.5 燃烧计算

对固体和液体燃料，燃烧计算是针对 1 kg 燃料，对气体燃料是针对 1 m³ 燃料。

5.5.1 理论空气量

1 kg 固体或液体燃料燃烧时所需的理论空气量 V^0（m³/kg）或 L^0（kg/kg）分别是：

$$V^0 = 0.088\,9(\%C + 0.375\%S) + 0.265\%H - 0.033\,3\%O \tag{2.1-122}$$

$$L^0 = 0.115(\%C + 0.375\%S) + 0.342\%H - 0.043\,1\%O \tag{2.1-123}$$

1 m³ 气体燃料燃烧时所需的理论空气量 V^0（m³/m³）是

$$V^0 = 0.047\,6\left[0.5\%CO + 0.5\%H_2 + 1.5\%H_2S + \sum\left(m + \frac{n}{4}\right)\%C_mH_n - \%O_2\right] \tag{2.1-124}$$

理论空气量是指燃料在完全燃烧时理论上所需的空气量。实际上要多送一些空气，使燃料能尽可能地完全燃烧，实际空气量与理论空气量之比称为过剩空气系数 a。

5.5.2 理论燃烧温度

通常所说的燃烧反应是放热反应，所放出的热量将使体系温度升高。如果没有热损失，热量全部用于升温，那么体系温度最终将达到最高。现定义：燃烧反应在绝热、恒压条件下达到平衡时，体系最终达到的温度称理论燃烧温度，以 T_m 表示之。

为了求理论燃烧温度，将燃烧反应表示成下列方框图（图 2.1-38）：

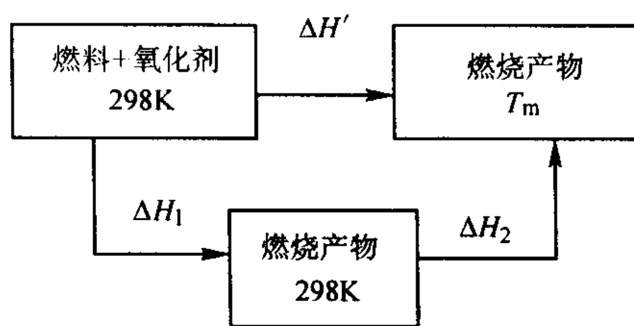


图 2.1-38　燃烧反应方框图

恒压下，$\Delta H = Q_P$，所以图中燃烧热可用焓变 ΔH 表示。根据盖斯定律，有下列关系：

$$\Delta H' = \Delta H_1 + \Delta H_2$$

绝热时，$\Delta H' = 0$，因而

$$\Delta H_2 = -\Delta H_1 \tag{2.1-125}$$

而 $\Delta H_1 = \Delta H_{298}$

$$\Delta H_2 = \sum_{i=p}\int_{298}^{T_m} n_i c_{pi}\,dT$$

工程上计算理论燃烧温度采用简化式，以避免繁琐运算，根据式（2.1-125）可得下式

$$\sum(V_i c_i)t_m = Q_1 + Q_2 + Q_3 - Q_4$$

所以

$$t_m = \frac{Q_1 + Q_2 + Q_3 - Q_4}{\sum V_i c_i} \tag{2.1-126}$$

式中，t_m 为理论燃烧温度，℃；Q_1 为燃料发热量，kJ·kg⁻¹或kJ·m⁻³；Q_2、Q_3 为燃料、氧化剂的物理热，kJ·kg⁻¹或kJ·m⁻³；Q_4 为高温离解热，kJ·kg⁻¹或kJ·m⁻³；V_i 为燃烧产物中气体的体积，m³·kg⁻¹；c_i 为燃烧产物中气体的比热容，kJ·(m³·K)⁻¹。

式（2.1-126）是按每千克或每立方米燃料来计算理论燃烧温度的。

计算表明，碳在冷风条件下完全燃烧时的理论温度

$$t_m = 2\,230℃$$

风温越高、燃烧比越大，碳的理论燃烧温度就越高。由式（2.1-126）不难预料，碳的理论燃烧温度还和空气富氧率有关。O_2 浓度越高，理论燃烧温度也越高。富氧率每提高 1%，理论燃烧温度至少提高 90℃。

6 溶液、熔渣与熔剂的性质

6.1 溶液的活度

6.1.1 活度的基本概念

有两类溶液，即理想溶液和稀溶液，其性质与组成间的关系较简单。例如，理想溶液中组元蒸气压与其浓度的关系遵守拉乌尔定律，稀溶液中溶剂也遵守拉乌尔定律，溶质则遵守亨利定律。合金熔炼过程中的合金熔体和熔融炉渣也是溶液，这些实际溶液的热力学规律却不那样简单。引入活度后，理想溶液和稀溶液的热力学公式就可以适用于实际溶液，解决实际溶液存在的问题就可以简化。由此可见，活度只是处理实际溶液的一个工具。

活度可以理解为有效浓度，这就是活度的物理意义。

既然活度 a_R 是校正后的浓度，那么溶液中某物质的实际浓度 x 乘以校正系数 γ 即得活度

$$a_R = \gamma x \tag{2.1-127}$$

现称校正系数 γ 为活度系数。

$$\gamma = \frac{a_R}{x} = \frac{p^{实}}{p^0 x} = \frac{p^{实}}{p^R} \tag{2.1-128}$$

式中，$p^{实}$ 为实际蒸气压；p^0 为纯物质的蒸气压；p^R 为按拉乌尔定律计算的理论蒸气压。$\gamma = 1$ 时，$p^{实} = p^R$，实际蒸气压和按拉乌尔定律计算的蒸气压为同一值，即实际溶液就是理想溶液。$\gamma \neq 1$ 时，$p^{实} \neq p^R$，实际溶液偏离了理想溶液。因此 γ 值可以用来衡量实际溶液偏离理想溶液的程度，这也是活度系数的物理意义。

6.1.2 溶液中组元的活度

绝大多数合金熔体和熔渣都不是理想溶液和稀溶液，因此分析这些熔体参与的反应热力学时，必须知道其中组元的活度。活度可以直接测定，有些活度也可以计算。现将合金熔液和熔渣这两类熔体中组元的活度分述如下。

(1) 合金熔体中合金元素的活度

如前所述，在 $M-i$ 二元合金溶液中，溶质 i 的活度系数将随其浓度而变化。材料加工中的许多合金，如铸造合金，实际上是多元合金，溶质的活度系数还受其他溶质的影响。令 f_i^i 表示 $M-i$ 二元系中溶质 i 的活度系数。由于第三元素 j 的加入（i 的浓度保持不变），第二元素 i 的活度系数将发生变化，因而 f_i^i 值必须乘以校正系数 f_i^j 才是 $M-i-j$ 三元系中第二元素的活度系数 f_i

$$f_i = f_i^i f_i^j \tag{2.1-129}$$

上式表明，三元系中 f_i 取决于两个因素，一是二元系中 i

自身所引起的活度系数f_i^i；另一是由于第三元素 j 加入后，给第二元素 i 的活度系数带来的影响f_i^j。

为了研究第三元素 j 对第二元素 i 的活度系数的影响，可以通过实验作 $\lg f_i^j-[\%j]$ 图，如图 2.1-39 所示。从该图可以看到各种第三元素 j 分别加入 Fe－S 二元系后，对 S 活度系数的影响。如果在 $[\%j]\rightarrow 0$ 的区域内对 $\lg f_i^j-[\%j]$曲线作切线，并令切线的斜率为 e_i^j，则

$$e_i^j=\left(\frac{\partial \lg f_i^j}{\partial [\%j]}\right)_{[\%i]} \qquad (2.1\text{-}130)$$

e_i^j 的物理意义是：在三元系中，当第二元素 i 的质量分数不变时，每加入质量分数为1%的第三元素 j 到溶液内，引起第二元素活度系数对数的变化值。因此，e_i^j 叫作活度的相互作用系数。

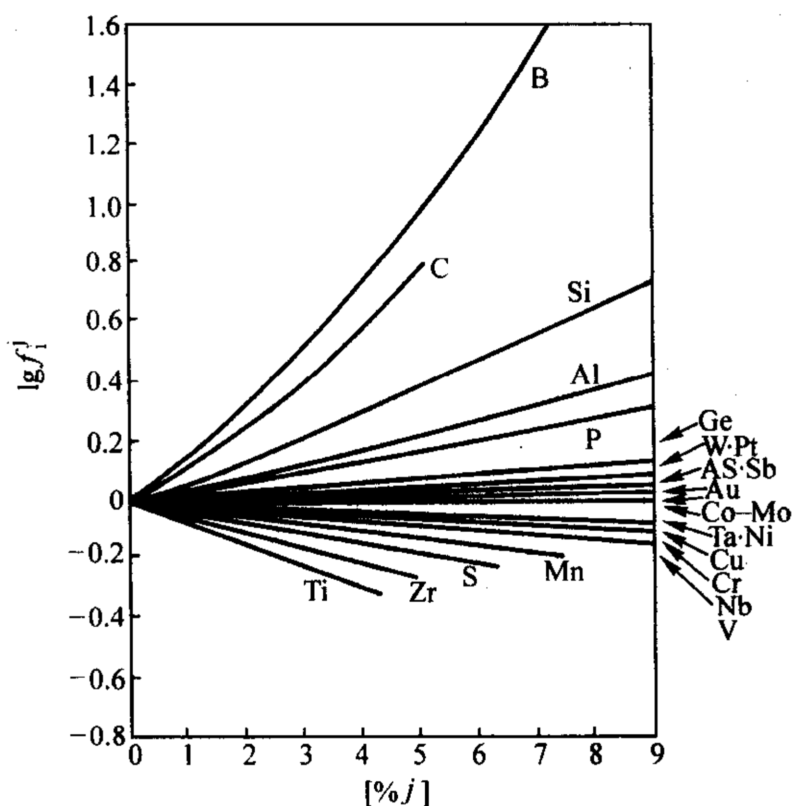


图 2.1-39　铁液内第三元素 j 对 S 活度系数的影响

由式（2.1-130）得 $\lg f_i^j=e_i^j[\%j]$。式（2.1-129）两端取对数后可得

$$\lg f_i=\lg f_i^i+\lg f_i^j=\lg f_i^i+e_i^j[\%j]=e_i^i[\%i]+e_i^j[\%j]$$

式中 $e_i^i=\dfrac{\mathrm{d}\lg f_i^i}{\mathrm{d}[\%i]}$也叫作活度的相互作用系数，它表示在 $M-i$ 二元系中，每加入质量分数 1%i 时，引起它自身活度系数对数的变化值。

推广到溶有 2、3、4…等元素的多元系，则第 2 元素的活度系数为

$$f_2=f_2^2f_2^3f_2^4\cdots$$

所以

$$\lg f_2=e_2^2[\%2]+e_2^3[\%3]+e_2^4[\%4]+\cdots$$

$$=\sum_{j=2}^{n}e_2^j[\%j] \qquad (2.1\text{-}131)$$

计算时，式中 e_2^4、e_2^5 等借用三元系的 e_i^j 值。

e_i^i 和 e_i^j 都和温度有关。例如，1 600℃时 Fe－i 系的 e_i^j 和 Fe－i－j 系的e_i^j 值可通过查表得到。1 200℃附近 Cu－i 系的e_i^i 和 Cu－i－j 系的e_i^j 值也可查表得到。

（2）熔渣中氧化物的活度

对渣中氧化物活度的研究，不如对合金溶液的研究成熟。通常用渣中某氧化物的等活度（或等活度系数）曲线图来求活度。所谓等活度曲线，就是在同一条曲线上，成分不同的渣中，某氧化物的活度相等。在知道炉渣成分的条件下，可通过查相关的等活度图资料求得炉渣中某一氧化物的活度。

6.2　熔渣的性质

6.2.1　熔渣的成分及结构

熔渣的各种物理化学性质及金属和渣间的反应都与熔渣的结构有关。高温熔渣的结构是很复杂的，由于现在还受着研究方法和研究实验手段的限制，尚难由实验直接确定液态渣以及金属熔体的结构。现有熔渣结构的理论，是根据：①凝固渣的结构分析；②熔渣高温下的某些性质，如黏度、密度、电导等性质间接推断的。

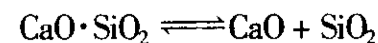
关于熔渣的结构，形成了两种理论：分子结构模型和离子结构模型。

（1）分子结构模型理论

熔渣的分子模型理论系在固体渣的化学分析、岩相分析、X 射线结构分析和状态图研究的基础上提出来的。分子结构理论的要点如下。

1）固体渣中存在着简单氧化物，如 SiO_2，CaO，MgO，Al_2O_3，FeO，Fe_2O_3，MnO 等和复杂的氧化物，如 $2CaO\cdot SiO_2$，$CaO\cdot SiO_2$，$2FeO\cdot SiO_2$，$CaO\cdot Al_2O_3$，$4CaO\cdot P_2O_5$，$CaO\cdot Al_2O_3\cdot 2SiO_2$ 等。因此，根据液态物质的结构理论，也可认为液态中存在有这些分子。

2）每种简单氧化物和其复杂氧化物之间存在着生成和分解的平衡反应，如



$$CaO\cdot SiO_2 \rightleftharpoons CaO+SiO_2$$

反应中物质的浓度由反应的离解常数确定：

$$K_D=\frac{N_{CaO}\times N_{SiO_2}}{N_{CaO\cdot SiO_2}}$$

熔渣中每种氧化物都具有两种不同的状态，以简单分子存在的叫作自由氧化物，以复杂氧化物分子存在的叫作结合氧化物。熔渣的化学性质及其反应能力主要是由这些自由氧化物的浓度决定的。

3）把熔渣看作理想溶液，因此自由氧化物的浓度，即其活度，可用摩尔分数表示。

分子理论引出的氧化物的自由浓度之概念能定性地说明金属与熔渣间的许多反应，同时它又是热力学计算的基础物性。但这种自由浓度的计算是困难的，因为复杂氧化物分子的选择和其离解常数的确定缺乏实验根据。

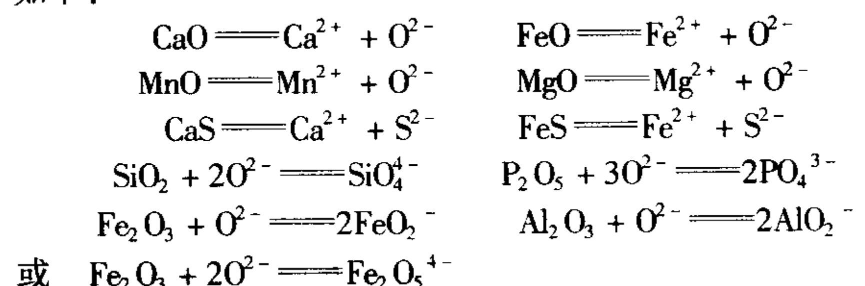
（2）离子结构模型理论

分子理论虽然能简单地说明金属和熔渣间反应的热力学规律，但不能解释熔渣的电化学性质。因此，在实验基础上提出并发展了离子结构模型。

熔渣离子结构理论是在广泛地利用 X 射线结构分析的研究、统计热力学的方法以及熔渣物理性质的研究上建立起来的。熔渣的电导、电解及电动势的测定、金属与熔渣界面上双电层的存在等都证实了熔渣是离子组成的熔体。

离子结构模型理论的要点如下。

1）虽然组成熔渣的氧化物在固态时是离子型结构，但在形成熔渣时，碱性氧化物电离，形成金属阳离子和氧阴离子。酸性氧化物则吸收熔体中的氧离子，转变为复合阴离子。因此，熔渣中离子结构单位质点的形成反应可表示如下：



$$CaO = Ca^{2+}+O^{2-} \qquad FeO = Fe^{2+}+O^{2-}$$

$$MnO = Mn^{2+}+O^{2-} \qquad MgO = Mg^{2+}+O^{2-}$$

$$CaS = Ca^{2+}+S^{2-} \qquad FeS = Fe^{2+}+S^{2-}$$

$$SiO_2+2O^{2-} = SiO_4^{4-} \qquad P_2O_5+3O^{2-} = 2PO_4^{3-}$$

$$Fe_2O_3+O^{2-} = 2FeO_2^{-} \qquad Al_2O_3+O^{2-} = 2AlO_2^{-}$$

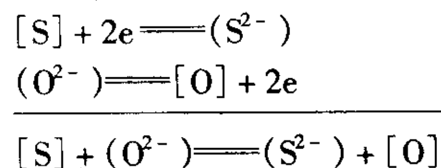
或　$$Fe_2O_3+2O^{2-} = Fe_2O_5^{4-}$$

即可认为组成熔渣的基本离子是 Ca^{2+}，Fe^{2+}，Mn^{2+}，Mg^{2+}，O^{2-}，S^{2-}，SiO_4^{4-}，PO_4^{3-}，FeO_2^-，AlO_2^- 等。

2）复杂阴离子的结构是比较复杂的，因为它随熔渣组成及温度而改变。例如熔渣中的 SiO_4^{4-} 离子可因 SiO_2 之加入形成一系列的复合硅氧离子：$Si_2O_7^{6-}$，$Si_3O_9^{6-}$，$Si_4O_{12}^{8-}$ 等。

3）离子间的作用能可近似用两离子间出现的静电力来计算。

4）关于金属与熔渣间的电化学反应式，由于熔渣是离子结构，而金属是非极性物质结构，因此两相间的反应是电化学性质的，即是由阴极与阳极两个半电池反应组成的。例如，脱硫的离子反应式为



$$[S] + 2e = (S^{2-})$$
$$(O^{2-}) = [O] + 2e$$
$$[S] + (O^{2-}) = (S^{2-}) + [O]$$

6.2.2 熔渣的性质

（1）熔渣的碱度

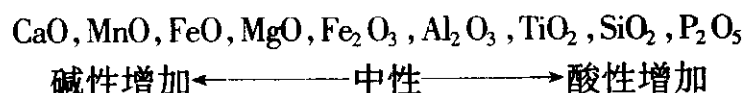
组成熔渣的氧化物可分三类：

酸性氧化物：SiO_2，P_2O_5，V_2O_5；

碱性氧化物：CaO，MgO，FeO，MnO，V_2O_5，TiO_2；

两性氧化物：Al_2O_3，V_2O_3，Fe_2O_3，(Cr_2O_3)，(TiO_2)。

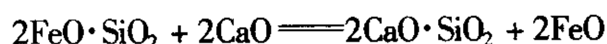
氧化物按碱性或酸性的强弱次序，可根据其静电力之大小排列如下：



CaO，MnO，FeO，MgO，Fe_2O_3，Al_2O_3，TiO_2，SiO_2，P_2O_5

碱性增加←——中性——→酸性增加

利用氧化物酸性或碱性的相对大小，可以确定反应转变的方向。例如，由于 CaO 的碱性比 FeO 的碱性强，故 CaO 能从 $2FeO \cdot SiO_2$ 中取代 FeO，而将 FeO 游离出来提高其活度



$$2FeO \cdot SiO_2 + 2CaO = 2CaO \cdot SiO_2 + 2FeO$$

即强氧化物能从复杂化合物中将弱氧化物取代出来。

熔渣的酸性或碱性取决于其中占优势的氧化物是酸性或碱性。一般用碱度表明熔渣的酸碱性。

表示碱度 B，应用最广泛的公式是碱性氧化物浓度对酸性氧化物浓度之比。常用的公式为

$$B = \frac{\%CaO}{\%SiO_2} \tag{2.1-132}$$

或 $$B = \frac{\%CaO}{\%SiO_2 + \%P_2O_5}$$ （适用于含磷高的炼钢渣）　(2.1-133)

或 $$B = \frac{\%CaO + \%MgO}{\%SiO_2 + \%Al_2O_3}$$ （适用于高炉渣）　(2.1-134)

第一式应用最普遍，第二式用于含磷高的炼钢渣，而第三式则用于含 MgO 及 Al_2O_3 量较高的高炉渣或冲天炉炉渣。这种表示法是把某些次要氧化物加以忽略，为的是应用方便，而且在理论上也是正确的，因为 FeO，MnO，MgO 放出的 O^{2-}，可为 Al_2O_3，Fe_2O_3 等形成复合阴离子所消耗。

对于材料加工用的冲天炉或电炉的炉渣，一般碱度 $B<0.8$ 时为酸性渣，$B=0.8\sim1.2$ 时为中性渣，$B>1.2$ 时为碱性渣。

（2）熔渣的氧化能力和还原能力

熔渣的氧化能力是指它能向与之接触的金属液中传递氧的能力。对于钢铁熔液，这个氧化能力是用钢与熔渣平衡的氧量 $[\%O]_S$ 和钢液中的实际氧量 $[\%O]_M$ 之差来表示

$$\Delta[\%O] = [\%O]_S - [\%O]_M \tag{2.1-135}$$

$\Delta[\%O]>0$ 时，渣中的氧能向钢液中扩散，使钢液不断受到氧化，称为氧化渣。

$\Delta[\%O]<0$ 时，钢液中的氧能向渣中扩散，即熔渣自钢液中能抽取氧，称为还原渣或脱氧渣。

$\Delta[\%O]=0$ 时，即熔渣既不具有氧化能力，也不具有还原能力，称为中性渣。

与熔渣平衡的钢液的氧量 $[\%O]_S$ 可用下式求得

$$[\%O]_S = \frac{a_{FeO}}{L_O}$$

式中，a_{FeO} 为渣中 FeO 的活度，与熔渣组成和温度有关；L_O 为氧在熔渣与钢液间的分配系数。

氧的分配系数 L_O 在数值上等于氧的溶解度 $[\%O]_{饱}$。溶解度 $[\%O]_{饱}$ 与温度有关

$$\lg[\%O]_{饱} = -\frac{6\,320}{T} + 2.73 \tag{2.1-136}$$

所以 $$[\%O]_S = [\%O]_{饱}\, a_{FeO} \tag{2.1-137}$$

（3）熔渣的熔点

熔渣为多元系，其熔点主要受熔渣的成分影响，往往通过实验的方法测定，得到等熔点曲线图。用得最多的是 $FeO-CaO-SiO_2$ 和 $CaO-SiO_2-Al_2O_3$ 二个渣系。

图 2.1-40 是 $FeO-CaO-SiO_2$ 渣系的等熔点曲线图。这个渣系在 FeO 与 SiO_2 的含量大体上相等、CaO 的含量为 10%左右时熔点最低。

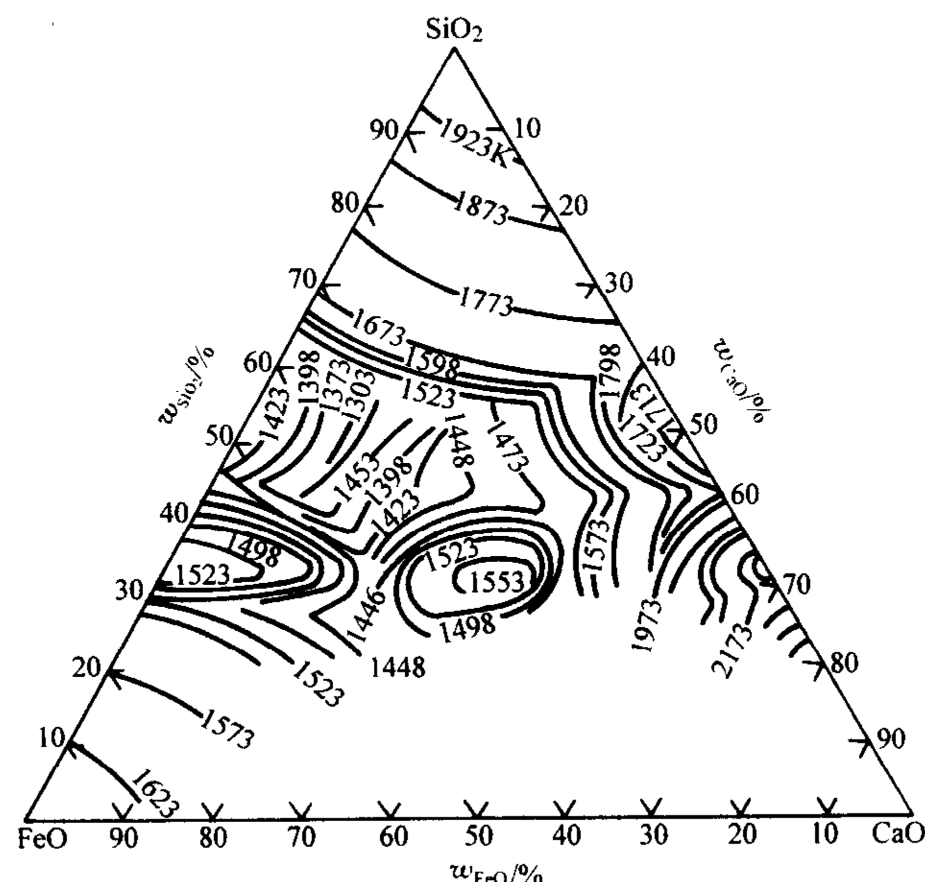


图 2.1-40　$FeO-CaO-SiO_2$ 渣系的等熔点曲线

（4）熔渣的黏度

熔渣的黏度也主要通过黏度测定方法获得。黏度主要受熔渣的组成及温度的影响。

渣中组元对黏度的影响如图 2.1-41 和图 2.1-42 所示。温度对黏度的影响较显著，一般随着温度升高，黏度降低。

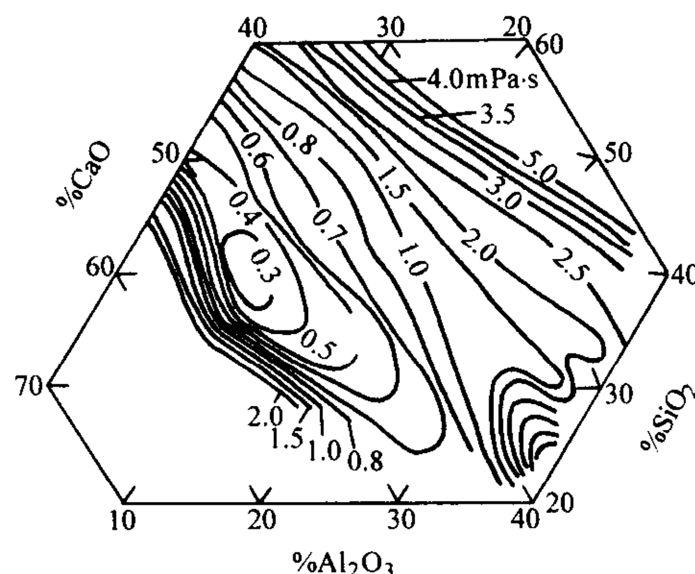


图 2.1-41　1 500℃时 $CaO-SiO_2-Al_2O_3$ 系等黏度曲线

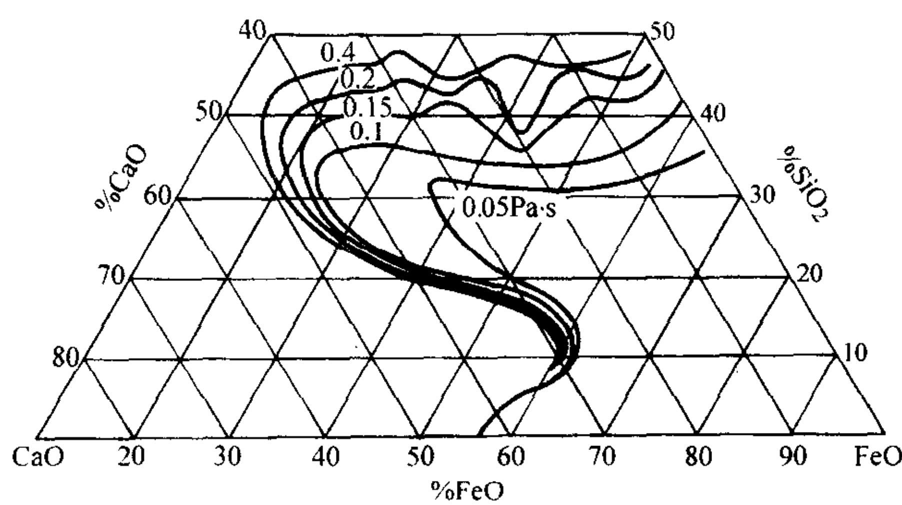


图2.1-42　1 400℃时 $FeO-CaO-SiO_2$ 系等黏度曲线

6.3　溶剂的性质

溶剂通常也叫覆盖渣，主要作用在于保护液体金属，尽可能减少金属吸气和氧化的机会，对于容易氧化的金属如铝、镁、锌等尤为重要。对于不易氧化的金属如铜、镍等，良好的覆盖也可以防止金属在熔炼中大量吸氧。此外，有的有色合金熔炼时使用溶剂进行精炼，达到除气及除渣的目的。因此，溶剂除覆盖作用外，还要求在出炉和浇注时熔渣容易与金属分离。对于溶剂或覆盖渣的要求可以归纳为下述4点。

1）熔渣的密度应小于金属液的密度，以保证其飘浮在液面。但熔渣的沸点或分解温度应尽可能高。

2）渣的熔点应低于金属的熔点，以保证其良好的流动性。

3）熔渣的表面张力应尽可能小，以保证其良好的覆盖能力与脱离能力。

4）渣与金属和炉气不发生直接的化学反应或间接的化学反应（氧的传递反应），也不相互溶解。

通常金属卤化物（特别是氯化物与氟化物）具有比氧化物更低的化学位（标准生成自由焓），因此更为稳定，不会与炉气或金属发生化学反应。卤化物的熔点也比氧化物低，密度也低于一般的金属。因此，卤化物多数可以成为良好的覆盖渣材料。但有些卤化物在高温下容易分解，有的沸点较低，故应用时应该充分注意这些特点。表2.1-7示出一些卤化物的物理性质。

通常，为了获得需要的物理性质，往往采用几种盐组成复合盐溶剂。例如，铝合金精炼时采用 NaCl 和 KCl 的复合盐。镁合金覆盖用 RJ 系列溶剂的成分为 $MgCl_2$、NaCl、KCl、$BaCl_2$ 和 CaF_2 等。

表2.1-7　一些卤化物的物理性质

物质名称	密度/$g\cdot cm^{-3}$	熔点/℃	沸点/℃	黏度/Pa·s（熔点附近）	表面张力/$mN\cdot m^{-1}$（熔点附近）
NaCl	2.16	800	1 413	0.001 49	114
KCl	1.99	790		0.001 08	88
LiCl		614	1 600	0.001 81	
$CaCl_2$	2.51	772	1 412	0.004 84	664
$MgCl_2$	2.33	718		0.004 12	
$BaCl_2$	3.87	953			
NaF		995			
KF		846			200
LiF		842			138
CaF_2		1 418			250
MgF_2		1 270			297
Na_3AlF_6	3.04			0.002 80	

7　氧化还原反应原理

7.1　氧化反应的氧位图

7.1.1　氧位定义

假定 O_2 为理想气体，O_2 在压力为 p_{O_2}、温度为 T 时的化学位为

$$\mu_{O_2}=\mu_{O_2}^{*}+RT\ln p_{O_2}$$

式中，$\mu_{O_2}^{*}$ 为 O_2 的标准化学位，即 $p_{O_2}=1$ atm 时的化学位，它只与温度有关。

T、p_{O_2} 给定后，本应得到 μ_{O_2}，但 $\mu_{O_2}^{*}$ 值无法知道，故 μ_{O_2} 值也无法得到。$\mu_{O_2}^{*}$ 移项后得

$$\Delta\mu_{O_2}=\mu_{O_2}-\mu_{O_2}^{*}=RT\ln p_{O_2}$$

结果，$\Delta\mu_{O_2}$ 可按 $RT\ln p_{O_2}$ 计算。$\Delta\mu_{O_2}$ 是 O_2 在其压力为 p_{O_2} 时与1 atm时的化学位之差，即相对于1 atm时的化学位。现定义：O_2 的相对化学位 $\Delta\mu_{O_2}$ 为氧位，其值由 $RT\ln p_{O_2}$ 衡量。目前不少文献已直接称 $RT\ln p_{O_2}$ 为氧位。之所以如此，就是由于它等于 O_2 的相对化学位。

显然，氧位高低与 O_2 分压和温度有关。

7.1.2　氧位意义

众所周知，化学位是判断物质转移方向和限度的依据，即物质总是由化学位高的一相转移到其化学位低的一相，直至它在各相的化学位相等时为止。与此相类似，氧位可以判断氧在含氧各相之间转移的方向和限度。例如，若氧化物氧位高于气相氧位，则氧化物分解，放出的 O_2 进入气相；这时气相氧位不断升高，最后与氧化物氧位相等，氧化物和气相处于平衡。如果气相氧位高于铜液的氧位，那么气相中的 O_2 溶解于铜液中，使铜液氧化；这时铜液氧位升高，直至二者相等，含 O_2 气相与含氧铜液处于平衡。由此可见，氧位标志着含氧物质的氧化能力。

因此，氧化物是否分解，合金熔体是否氧化以及其他许多重要冶金反应的方向和限度问题，都可以用氧位进行分析。这就是研究氧位的意义所在。

7.1.3　氧位图

物理化学中的氧化物的 $\Delta G^{*}-t$ 图，其实就是氧位图。在这里，氧位图应用于材料加工过程的冶金反应等化学反应。

图2.1-43为氧位图。横轴表示温度，纵轴为氧位 $RT\ln p_{O_2}$，故图中直线表示氧位与温度的关系，称氧位线。每条氧位线都对应一个冶金反应。这些反应不外乎三种类型，如表2.1-8所示。

表2.1-8　氧位线类型

类型	氧　位　线	对应的反应（举例）	说　　明
1	气相氧位线	$2CO+O_2=2CO_2$	被氧化物 M 和氧化产物 M_xO_y 均为气相
2	氧化物（气态）氧位线	$2C+O_2=2CO$	M 为纯凝聚相，M_xO_y 为气相
3	氧化物（凝聚态）氧位线	$2Fe+O_2=2FeO$	M 和 M_xO_y 均为纯凝聚相

对于类型3，下式成立，所以该种情况下纵轴又可标为反应自由焓变化。

图 2.1-43 氧位图

$$\Delta G^{\circ} = RT\ln p_{O_2} \qquad (2.1\text{-}138)$$

对于类型 2，例如表内所列反应，因

$$\Delta G^{\circ} = RT\ln p_{O_2} - 2RT\ln p_{CO}$$

故当平衡分压 $p_{CO} = 1$ atm（0.101 MPa）时，式（2.1-138）也成立。在此前提下，纵轴也可标为 ΔG°。属于同一类型的反应还有 $C + O_2 = CO_2$。

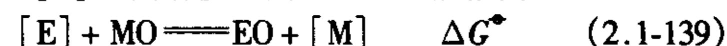
对于类型 1，可以证明：当所列反应的平衡分压比 $\frac{p_{CO}}{p_{CO_2}} = 1$ 时，式（2.1-138）也成立。在此条件下，纵轴同样可用 ΔG° 标定。反应 $2H_2 + O_2 = 2H_2O$ 属于这种类型。

在图的最左侧 $T = 0$ K 处，还有一 0 K 轴。

从氧位图可知，氧位线可概括为两类：一类是氧化物氧位线，另一类是气相氧位线。氧位线的位置越高，氧位也越高，即该氧位线所代表的氧化物或气相的氧化能力也越高。因此，氧可以从该氧化物或气相中转移到其下部氧位线所代表的体系中，使这些体系中某物质氧化。基于这个道理，氧位图在材料热加工及冶金上得到广泛应用。

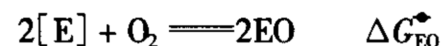
7.2 氧化还原热力学条件

合金元素 [E] 的氧化还原反应可以概括为：

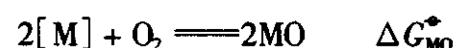


$$[E] + MO = EO + [M] \qquad \Delta G^{\circ} \qquad (2.1\text{-}139)$$

由 [E] 和 [M] 被 O_2 直接氧化



$$2[E] + O_2 = 2EO \qquad \Delta G^{\circ}_{EO}$$



$$2[M] + O_2 = 2MO \qquad \Delta G^{\circ}_{MO}$$

可得反应式（2.1-139），因而

$$\Delta G^{\circ} = \frac{1}{2}\left(\Delta G^{\circ}_{EO} - \Delta G^{\circ}_{MO}\right) \qquad (2.1\text{-}140)$$

标准态下，氧化 - 还原反应式（2.1-139）要能自发进行，必须，$\Delta G^{\circ} < 0$，即

$$-\Delta G^{\circ}_{EO} > -\Delta G^{\circ}_{MO} \qquad (2.1\text{-}141)$$

从氧化角度讨论反应（2.1-139），E 为被氧化的合金元素，MO 为氧化剂，则式（2.1-141）表示合金元素 [E] 对氧的亲和力大于氧化剂 MO 中 M 对氧的亲和力。在此条件下，[E] 就要被氧化，所以式（2.1-141）是合金元素氧化的热力学条件。该式也是 MO 作为氧化剂的条件，即 M 对氧

的亲和力必须小于合金元素 E 对氧的亲和力，MO 才能作为氧化剂。

从还原角度讨论反应式（2.1-139），则［E］为还原剂，MO 为被还原的氧化物。这时式（2.1-141）是氧化物还原的热力学条件，即还原剂对氧的亲和力必须大于被还原氧化物 MO 中 M 对氧的亲和力。式（2.1-141）也是还原剂必备的条件，不再赘述。总之，式（2.1-141）是氧化-还原反应进行的热力学条件。

再次指出，这种判断只适用标准态，其中［E］的标准态为质量分数 1%溶液，否则需借助等温方程。

此外，也可利用分解压进行判断，若

$$p_{O_2(EO)} < p_{O_2(MO)} \qquad (2.1\text{-}142)$$

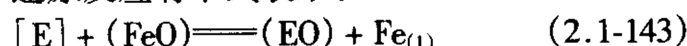
即合金元素氧化物的分解压小于氧化剂的分解压；在此条件下，合金元素就会被氧化。或者说，还原剂氧化物的分解压小于被还原氧化物的分解压时，氧化物就会被还原。所以式（2.1-142）也是合金元素氧化还原的热力学条件。

如上所述，根据式（2.1-142）判别溶解元素的氧化或还原，只限于标准态；而式（2.1-142）不受此限制，在任何条件下都适用。但是必须指出，决不能自氧位图（图 2.1-43）求得分解压，再根据式（2.1-142）判断。凝聚相不是纯物质而为溶液时，分解压还受溶液浓度的影响；这种影响有可能改变分解压的相对大小，以致改变氧化-还原的方向。

7.3　几种典型的氧化还原反应

7.3.1　钢铁溶液中合金元素被 FeO 氧化

铸铁及铸钢熔炼过程中，炉料中的合金元素主要是在熔渣形成后被渣中 FeO 氧化。由于元素氧化生成的氧化物溶于熔渣，故氧化-还原反应有下式表示：



$$[E] + (FeO) = (EO) + Fe_{(l)} \qquad (2.1\text{-}143)$$

（FeO）和（EO）都以其纯物质为标准态，因此从氧位图可以得知，FeO 氧位线以下的合金元素，例如硅、钒、锰、铬等，都能被（FeO）氧化。

式（2.1-143）的平衡常数为

$$K_E = \frac{a_{(EO)}\,a_{Fe}}{a_{(FeO)}\,a_{[E]}}$$

合金主要成分是铁，接近纯物质，因而 $a_{Fe} \approx 1$。将

$$a_{[E]} = f_{[E]}[\%E],\quad a_{(EO)} = \gamma_{(EO)}\,x_{(EO)}$$

代入上式整理后得

$$\eta_E = \frac{x_{(EO)}}{[\%E]} = K_E\,\frac{a_{(FeO)}f_{[E]}}{\gamma_{(EO)}} \qquad (2.1\text{-}144)$$

式中，η_E 为元素氧化物在渣中的浓度与元素在合金中的浓度之比，即 $\frac{x_{(EO)}}{[\%E]}$，称此为元素在渣中和合金中的分配系数，简称分配系数，表示合金元素氧化的程度。从式（2.1-144）可以分析影响元素氧化程度的因素。影响因素主要有合金元素对氧的亲和力、温度、炉渣氧化能力、炉渣碱度以及合金成分。

7.3.2　钢水脱碳反应

采用氧化法炼钢时，氧化期的主要任务之一是脱碳



$$[C] + [O] = CO \qquad (2.1\text{-}145)$$

脱碳产物是 CO 气泡，在它穿过钢水上浮过程中，有利于溶于钢中的气体（氢与氮）和非金属夹杂的排出；同时由此引起的熔池搅拌作用可以加速传质和传热，促进冶金反应进行，并使钢水成分和温度均匀。所以脱碳反应可以起到精炼的作用，它能提高钢水的质量，是氧化法炼钢过程中一个重要反应。

长期以来对脱碳反应作了大量研究工作，得到许多关于平衡常数 $K = \frac{p_{CO}}{a_{[C]}\,a_{[O]}}$ 与温度的关系式，下面是其中的一个

$$\lg K = \frac{1\,168}{T} + 2.07 \qquad (2.1\text{-}146)$$

习惯上，脱碳反应的平衡，不用平衡常数而用碳氧浓度积［%C］［%O］表示。令其为 m，并由平衡常数表达式得

$$m = [\%C][\%O] = \frac{p_{CO}}{f_{[C]}f_{[O]}K} \qquad (2.1\text{-}147)$$

如果碳和氧的活度系数积 $f_{[C]}f_{[O]} = 1$，或 $f_{[C]} = 1$、$f_{[O]} = 1$，那么当温度和 CO 分压 p_{CO} 一定时，碳氧浓度积 m 应为常数。由式（2.1-146）和式（2.1-147）求得 $t = 1\,600℃$、$p_{CO} = 1$ atm 下的 $[\%C][\%O] = 0.002$。

既然 C-O 浓度积为常数，那么［%C］和［%O］为等边双曲线关系，见图 2.1-44 上的曲线 1。所以，钢水中碳的浓度越低，则氧的浓度越高。生产上脱碳的结果，钢水含氧量必然增高。可见，C-O 浓度积揭示了铸钢熔炼的这一客观规律。

碳氧浓度积的意义还在于，可以根据脱碳要求，用它来计算钢中平衡含氧量，进而算出所需渣中 FeO 的浓度。

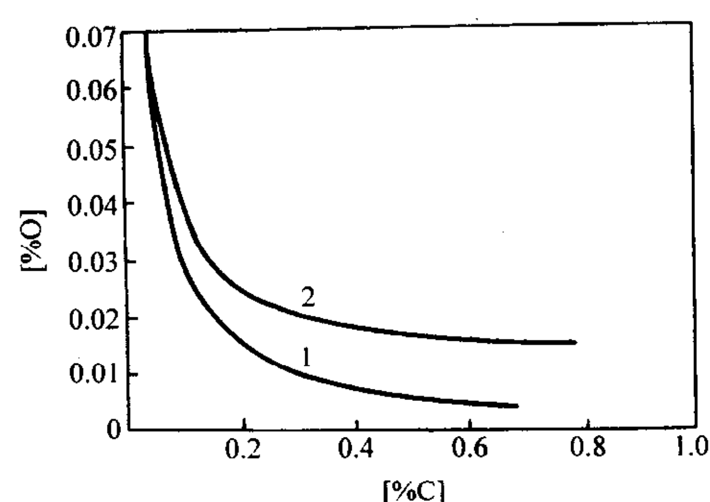


图 2.1-44　钢水中碳与氧的浓度曲线
1—电炉实际值；2—平炉实际值

7.3.3　脱氧反应

如前所述，钢液经氧化脱碳后，含氧量升高，一般将超过 0.02%～0.05%，而 1 600℃时，氧在固态 δ-Fe 中的溶解度只有 0.008 6%。这样，在钢水浇注凝固过程中，大量的氧将以 FeO 形式析出，它主要分布在晶界上，使钢铸件的力学性能下降。浇注大铸件时，由于选择结晶，在没有凝固的钢水中杂质富集，其中碳和氧的浓度将超过其平衡值，结果碳氧反应发生，放出 CO。除一部分排到大气外，还有一部分 CO 被凝固的金属包围，形成气孔。为了保证钢铸件的质量，浇注前必须脱氧。

熔炼铜合金时，在金属铜熔化过程中，一部分铜被氧化为 Cu_2O 并溶于铜液内。合金元素加入后，就要被 Cu_2O 氧化，产生氧化物夹杂。为了减少烧损和夹杂，在合金元素加入以前，铜液也必须脱氧。下面讨论钢液脱氧。

（1）脱氧反应热力学

1）脱氧常数　脱氧反应可概括为



$$x[M] + y[O] = M_xO_y$$

式中，M 为脱氧剂，M_xO_y 为脱氧产物。如果脱氧产物为纯物质 $a_{M_xO_y} = 1$。由脱氧反应平衡常数表达式可得

$$a_{[M]}^x\,a_{[O]}^y = K$$

或

$$[\%M]^x[\%O]^y = K\,\frac{1}{f_{[M]}^x f_{[O]}^y} = K' \qquad (2.1\text{-}148)$$

K 或 K' 称为脱氧常数。

脱氧常数值规定了残留在金属中脱氧剂和氧的平衡浓度，即［%M］和［%O］之间的关系。一般说来，脱氧常数值越小，则残留在金属内的氧浓度越低，该脱氧剂的脱氧

能力越强。所以，脱氧常数可用来衡量脱氧剂的脱氧能力。

表示平衡浓度［%M］和［%O］之间关系的曲线叫脱氧平衡曲线，简称脱氧曲线。图 2.1-45 为 1 600℃时钢液内一些元素的脱氧曲线，它更形象地表达了各种元素的脱氧能力。曲线的位置越低，则与等量脱氧元素平衡的残余氧量越少，因而该元素的脱氧能力越强。

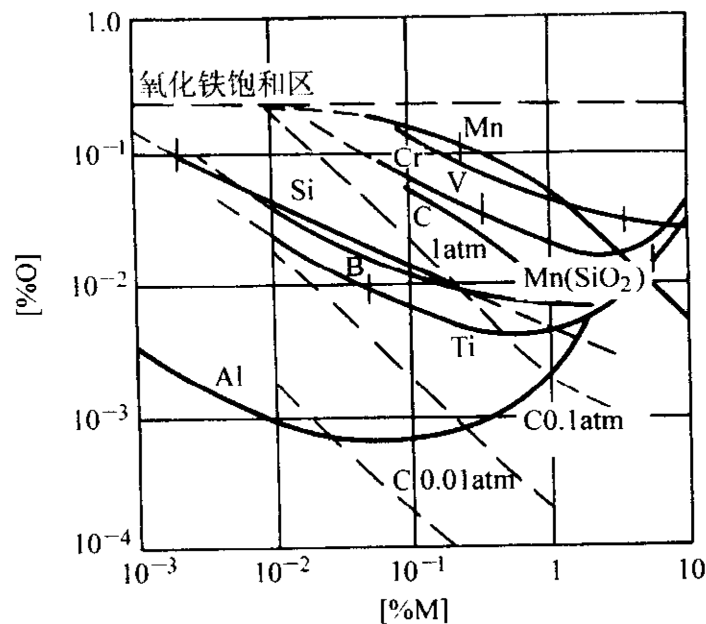


图 2.1-45　1 600℃钢液脱氧曲线

2）最低残余氧量　由图 2.1-45 可见，一些元素例如铝、钛、钒等，脱氧曲线出现最低点。换句话说，当其残余浓度［%M］增加时，残余氧量开始减少；进一步提高［%M］，残余氧量经过极小值以后，反而增大。这就启示人们，以这些元素作脱氧剂，加入量并非越多越好。另一类脱氧元素相反，它们在钢液内残余浓度越高，残余氧量越低。

（2）沉淀脱氧

生产上最常用的脱氧方法是把脱氧剂加到金属（钢液、铜液）内，使其直接和金属中的氧发生反应。生成的脱氧产物由于不溶于金属，且其密度小于金属，能自金属液内上浮排出，从而达到脱氧目的。这种脱氧方法称为沉淀脱氧。

沉淀脱氧的重要因素之一是脱氧产物能够迅速上浮排除，否则会残留在金属内凝固后成为非金属夹杂。根据斯托克斯（Stokes）上浮速度定律，若脱氧产物的颗粒大、密度差别大，将有利于上浮。

钢液采用铝、硅铁或锰铁等脱氧就属于沉淀脱氧；铜液采用磷脱氧也属于沉淀脱氧。

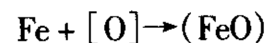
（3）扩散脱氧

沉淀脱氧由于脱氧剂直接加到钢水内部，产生脱氧产物夹杂是不可避免的。扩散脱氧能得到更纯净的钢水，电炉熔炼常用这种方法。

前面曾指出，钢、渣两相内氧的分配服从分配定律，

$$[\%O]=[\%O]_{饱}\ a_{(FeO)}$$

降低渣中 FeO 活度，氧就会从钢液内扩散进入渣中



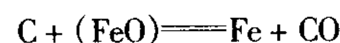
$$Fe+[O]\rightarrow(FeO)$$

使钢液含氧量下降，达到脱氧目的。

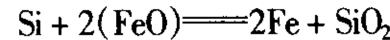
由此可见，扩散脱氧的关键在于降低渣中 FeO 活度。为此，氧化期结束后，应当扒掉氧化渣，另造还原性稀薄渣。这种渣形成时几乎不含 FeO，所以钢水中的氧可以向熔渣扩散。

随着扩散脱氧的进行，渣中 FeO 逐渐增多。为了使 $a_{(FeO)}$ 维持在低水平，必须分批地往渣中加入脱氧剂，还原渣中的 FeO，让钢中的氧不断向炉渣扩散。

作为扩散脱氧渣的脱氧剂，常采用碳粉（或部分电石）、硅铁粉等。碳、硅对氧的亲和力都比较强，能还原渣中 FeO，使其活度下降



$$C+(FeO)=\!=\!=Fe+CO$$



$$Si+2(FeO)=\!=\!=2Fe+SiO_2$$

扩散脱氧不在钢液内部进行，因而脱氧产物不会黏污钢水，这是其优点。但由于受到钢－渣界面限制及氧扩散速度限制，扩散脱氧需要较长的时间。

7.3.4　脱硫反应

通常，硫是钢铁材料的有害元素。钢铁熔液中的硫也容易与合金元素反应，硫化还原反应的原理类似于氧化还原反应。

（1）合金元素脱硫作用

硫在铁液内以 FeS 形式存在。纯 FeS 的分解压 $p_{S_2(FeS)}$ 很低，自硫位图可得：1 500℃时 $p_{S_2(FeS)}<10^{-3}$atm。铁液内 FeS 分解压将更小。所以，单靠加热使 FeS 分解达到脱硫目的是不可能的。

硫位图表明，CeS、CaS、MgS、Na_2S、MnS 等许多硫化物的分解压比 FeS 低，稳定性比 FeS 高，FeS 有可能转变为这些硫化物。另一方面，$FeS_{(1)}$ 在铁液内溶解度非常高，而上述其他硫化物中，MnS 的溶解度很小，CaS、MgS、Na_2S 等几乎不溶于铁。因此，选用一种脱硫剂，使铁液中的 FeS 转变为稳定性比 FeS 高、溶解度比 FeS 小得多的硫化物，就能把铸铁和铸钢中的硫脱掉。

下面讨论主要合金元素的脱硫能力

1）锰　锰脱硫反应如下：



$$[Mn]+[S]=\!=\!=MnS$$

MnS 熔点为 1 530℃。脱硫温度超过 1 530℃，生成液态 MnS；低于熔点，则为固态。脱硫反应生成 $MnS_{(s)}$ 时

$$\lg K_{(s)}=\lg\frac{1}{a_{[Mn]}\,a_{[S]}}=\frac{8\,295}{T}-4.96$$

生成 $MnS_{(1)}$ 时

$$\lg K_{(1)}=\lg\frac{1}{a_{[Mn]}\,a_{[S]}}=\frac{6\,951}{T}-4.22$$

前式在 1 530℃以下适用，后式适用于 1 530℃以上。

1 300℃时　　$K_{(s)}=2.06$

1 600℃时　　$K_{(1)}=0.31$

比较两个 K 值表明，低温对锰脱硫有利。

除低温外，锰脱硫还应具备以下条件：①金属中锰高；②硫的活度系数大。不具备这些条件，锰不可能脱硫。

2）稀土元素　镧和铈脱硫反应，脱硫常数与温度的关系以及 1 600℃时的脱硫常数值见表 2.1-9。从脱硫常数值可以看出，这两个稀土元素都有很强的脱硫能力。以镧为例，假定钢液内残余含量为 0.02%，那么，1 600℃时硫可降至

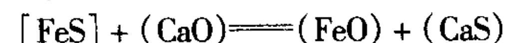
$$[\%S]=\frac{1.5\times10^{-4}}{0.02}=0.007\,5$$

表 2.1-9　镧、铈脱硫常数

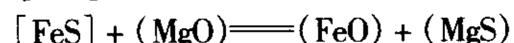
脱硫反应	脱硫常数 K_S	$\lg K_S$	1 600℃ K_S 值
$[La]+[S]=\!=\!=LaS_{(s)}$	$[\%La][\%S]$	$-\frac{26\,000}{T}+8.98$	1.5×10^{-4}
$[Ce]+[S]=\!=\!=CeS_{(s)}$	$[\%Ce][\%S]$	$-\frac{20\,600}{T}+6.39$	1.5×10^{-3}
$2[Ce]+3[S]=\!=\!=Ce_2S_{3(s)}$	$[\%Ce]^2[\%S]^3$	$-\frac{56\,123}{T}+17.13$	8×10^{-6}

（2）熔渣脱硫

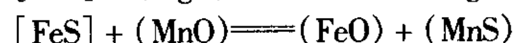
炉渣通常含有 CaO、MgO、MnO 等碱性氧化物，它们能和金属中的 FeS 发生反应



$$[FeS]+(CaO)=\!=\!=(FeO)+(CaS)$$

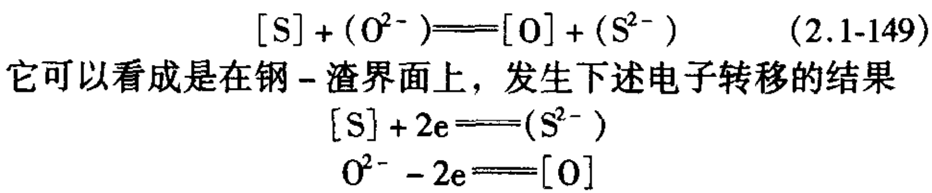


$$[FeS]+(MgO)=\!=\!=(FeO)+(MgS)$$



$$[FeS]+(MnO)=\!=\!=(FeO)+(MnS)$$

这是分子理论的脱硫反应。按照离子理论，脱硫反应为



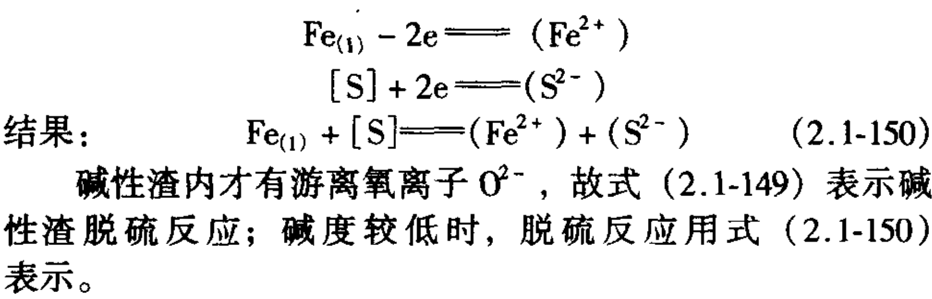
$$[S]+(O^{2-})\!=\!=\![O]+(S^{2-}) \qquad (2.1\text{-}149)$$

它可以看成是在钢－渣界面上，发生下述电子转移的结果

$$[S]+2e\!=\!=\!(S^{2-})$$

$$O^{2-}-2e\!=\!=\![O]$$

钢中硫［S］获得2个电子即进入炉渣，而放出2个电子的不一定是O^{2-}，也可以是Fe原子。它放出2个电子后变成Fe^{2+}和S^{2-}同时进入炉渣



$$Fe_{(l)}-2e\!=\!=\!(Fe^{2+})$$

$$[S]+2e\!=\!=\!(S^{2-})$$

结果：$$Fe_{(l)}+[S]\!=\!=\!(Fe^{2+})+(S^{2-}) \qquad (2.1\text{-}150)$$

碱性渣内才有游离氧离子O^{2-}，故式（2.1-149）表示碱性渣脱硫反应；碱度较低时，脱硫反应用式（2.1-150）表示。

编写：吴树森（华中科技大学）

第2章　材料成形分析的力学基础

1　连续介质力学的基本概念

连续介质力学研究物体的宏观力学行为。连续介质和质点、刚体一样，都是一种理论模型。连续介质模型认为物质连续地分布在它所占有的容积之内；这是运用数学分析工具统一研究固体、液体、气体的力学行为的基础。连续介质的质点，表示一个物质微团，它的尺度和所研究问题的宏观尺度相比是充分地小，小到在此微团内，每种物理量都可看成是均匀分布的常量，因而在数学上可以把此微团当作一个点来处理。另一方面，又要求此微团尺度和分子运动的尺度相比足够地大，大到微团中包括大量的分子，从而能对分子运动作统计平均，以得到表征宏观现象的物理量。为了运用数学分析工具，在连续介质力学中把材料当作连续的（或分段连续的），除了在某些间断面上以外，各物理量存在所需要的各阶偏导数。

连续介质力学中的规律分成两部分。

1）运动基本规律　它适用于一切物体，无论是由什么材料构成的。这些规律包括质量守恒、动量平衡、动量矩平衡和能量平衡等。

2）本构规律　即材料内力与运动的关系。

连续介质力学并不具体决定本构关系中的材料常数、函数或泛函，它们需由试验来确定，或者由材料科学来预言。

连续介质力学的任务是：

1）给定本构方程，预言材料的力学行为，即解边值问题；

2）根据材料定性的性质，提出适合于物质不变性原理的本构方程，因而提出材料常数、函数或泛函的形式。

而材料科学的任务是：

1）研究宏观力学性质与粒子物理及化学的联系；

2）设计与制造新材料；

3）定量地预言材料性质。

由此可见，连续介质力学与材料科学是互相补充的。

在材料成形分析中，我们一般地将材料作为连续介质处理。因而连续介质力学就成为我们对材料成形过程进行力学分析的最主要的理论基础之一。根据连续介质力学的理论，可以把材料成形问题抽象为一个微分方程（组）的边值问题。采用有限元法或有限差分方法对这个边值问题进行数值计算，从而求得成形过程中制件和模具中的位移（速度）场、应力场、应变场、温度场，预测成形过程中可能出现的缺陷，这是目前大多数工程问题所要求得到的结果。如果在模拟计算中引入由材料科学和其他力学分支研究得到的材料的组织、性能和状态的演化与以上场量之间的关系，则可以进一步预测成形后制件的组织性能，如晶粒大小、变形织构等，从而更有效地控制和优化成形工艺。

2　张量分析基础

张量是研究连续介质力学的基本数学工具，采用张量符号和张量运算，能使表述和推导简单明了。

2.1　张量的基本概念

张量的一个重要性质是它不依赖于具体的坐标系，因此是客观的。当然为了在数量上对张量进行表征和计算，常常需要选定参考的坐标系。在不同的坐标系下得到不同的分量值，不同坐标系下的分量之间具有确定的变换关系。

连续介质力学以及各物理学科研究的都是不依赖坐标系的物理量，因此自然要用张量这一数学工具。连续介质力学的运动规律都是用张量方程来表达的，它对各个坐标系都是成立的。

张量理论的表述可有两种不同的方式。第一种方式是完全不借助于坐标系，把张量整体作为研究的对象，对它进行运算。第二种方式是把张量看成是张量分量的集合，张量的运算就归结为张量分量的运算。两种方式各有其优点，第一种表述简单利落，第二种便于实际计算。

能在任意坐标系间变换的张量理论称为普遍张量理论；而只在笛卡儿坐标系间变换的张量理论称为笛卡儿张量理论（所谓笛卡儿坐标系指的就是右手旋转的直角坐标系）。笛卡儿张量理论远比普遍张量理论简单。以下我们主要采用笛卡儿张量，介绍连续介质力学的基本概念和基本运动规律，从而得到直角坐标系下的分量和方程组。

2.2　张量的定义与约定求和法

（1）张量的定义

若一个量有 n^N 个分量，而每个分量在 n 维空间 R^n 中的坐标变换

$$x_i = x_i\ (x'_1,\ \cdots,\ x'_n)\ (i=1,\ \cdots,\ n)$$

之下，按下面的规律变化：

$$T_{i'_1\cdots i'_m} = \frac{\partial x_{i_1}}{\partial x_{i'_1}}\cdots\frac{\partial x_{i_m}}{\partial x_{i'_m}}T_{i_1\cdots i_m} \tag{2.2-1}$$

式中，$T_{i_1\cdots i_m}$ 是 x_i 的函数；$T_{i'_1\cdots i'_m}$ 是 x'_i 的函数；$T_{i_1\cdots i_m}$（共有 n^N 个分量）为 N 阶张量。$T_{i_1\cdots i_m}$ 可以简单地记为 T_{im}。

张量概念是矢量和矩阵概念的推广，其中：标量是零阶张量，矢量是一阶张量，矩阵（方阵）是二阶张量。例如矢量 $\boldsymbol{x}$、二阶张量 $\boldsymbol{T}$ 可以分别以实体记法和分量记法表示为

$$\boldsymbol{x} = \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix},\quad \boldsymbol{T} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix}$$

通常，我们在整体记法中用黑体字表示矢量和二阶及二阶以上的张量。

（2）约定求和法

我们研究的是三维欧氏空间中的物理现象。此空间中的笛卡儿张量，其分量的每个指标都在1，2，3中取值。可以任取1或2或3的文字指标称为自由指标。为书写简便，我们约定：在同一项中，如有一个自由指标重复出现，就表示要对这个指标从1到3求和。例如

$$u_iv_i = u_1v_1 + u_2v_2 + u_3v_3$$

$$\frac{\partial u_k}{\partial x_k} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3}$$

对此约定求和法则，我们作以下规定：在一等式中，如果同一个文字指标，在其中的一项单独出现了，则它即使在其他某项内重复出现，对该项也不约定求和。例如：

$$f_i = T_{ii}$$

右边就不约定求和。

（3）常用符号

1）符号 δ_{ij}

$$\delta_{ij} = \begin{cases} 0, & 当\ i \neq j \\ 1, & 当\ i = j \end{cases} \tag{2.2-2}$$

它是一个二阶张量，可以写成一个单位矩阵

$$I=(\delta_{ij})=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

2）置换张量 e_{ijk}

$$e_{ijk}=\begin{cases}1，当(i,j,k)是(1,2,3)的偶排列\\-1，当(i,j,k)是(1,2,3)的奇排列\\0，当i,j,k中有相同者\end{cases} \tag{2.2-3}$$

它是一个三阶张量。行列式可用置换张量表示如下

$$\Delta=\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}=e_{ijk}a_{1i}a_{2j}a_{3k} \tag{2.2-4}$$

2.3　张量代数

1）张量的相等　若两个同类型的张量 $\boldsymbol{A}$，$\boldsymbol{B}$ 在同一个坐标系中的分量一一相等，则此二张量的其他一切分量均一一相等，且在任意坐标系中的一切分量均一一相等。记作 $\boldsymbol{A}=\boldsymbol{B}$。由此可知，张量方程在所有坐标系均成立。

2）加（减）法　同类型的若干个张量的对应分量相加（或相减）就得到一个新的同类型张量的分量，这种运算称为张量的加法（或减法）。例如

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{C}$$

的定义是

$$C_{i_n}=A_{i_n}+B_{i_n}$$

3）标量与张量相乘　设 k 是标量，$\boldsymbol{A}$ 是 n 阶张量，

$$k\boldsymbol{A}=\boldsymbol{A}k=\boldsymbol{B}$$

的定义是

$$B_{i_n}=kA_{i_n}$$

$\boldsymbol{B}$ 也是 n 阶张量。

4）张量乘积　两张量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的张量乘积记为 $\boldsymbol{AB}$。$\boldsymbol{A}$ 的每一分量与 $\boldsymbol{B}$ 的每一分量相乘，就得到 $\boldsymbol{AB}$ 的一个分量。设 $\boldsymbol{A}$ 是 m 阶张量，$\boldsymbol{B}$ 是 n 阶张量，则 $\boldsymbol{AB}$ 是 $m+n$ 阶的张量，

$$A_{i_m}B_{j_n}=(\boldsymbol{AB})_{i_mj_n}$$

$\boldsymbol{AB}$ 中当 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 是零阶张量时，就得到标量与张量的乘积。

以后为简单起见，用 A_{i_m} 既表示 n 阶张量 $\boldsymbol{A}$ 的任一个分量，也用 A_{i_m} 代表张量 $\boldsymbol{A}$ 本身。

5）张量的缩并　$A_{i_n}=A_{i_1}\cdots i_n$ 当 $i_1\cdots i_n$ 中有两个自由指标相同时，应用约定求和法，就得到一个有 $n-2$ 个自由指标的张量 $\boldsymbol{B}$，称 $\boldsymbol{B}$ 是张量 $\boldsymbol{A}$ 的缩并。例如当 $\boldsymbol{A}$ 中最后两个指标 $i_{n-1}=i_n=m$ 时，记

$$A_{i_1\cdots i_{n-2}mm}\equiv A_{i_{n-2}mm}=B_{i_{n-2}}$$

$\boldsymbol{B}$ 是 $n-2$ 阶张量。

6）张量的点积　张量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的张量乘积 $\boldsymbol{AB}$ 缩并一次后，就得到 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的点积。为明确起见，令 $\boldsymbol{A}$ 的最后一个指标与 $\boldsymbol{B}$ 的第一个指标相同，这样对 $\boldsymbol{AB}$ 缩并所得的张量记为 $\boldsymbol{A}\cdot\boldsymbol{B}$。设 $\boldsymbol{A}$ 是 m 阶张量，$\boldsymbol{B}$ 是 n 阶张量，则点积 $\boldsymbol{A}\cdot\boldsymbol{B}$ 是 $m+n-2$ 阶的张量。

$$A_{i_1\cdots i_{m-1}k}B_{kj_2\cdots j_n}\equiv(\boldsymbol{A}\cdot\boldsymbol{B})_{i_1\cdots i_{m-1}j_2\cdots j_n}$$

一般说来 $\boldsymbol{A}\cdot\boldsymbol{B}\neq\boldsymbol{B}\cdot\boldsymbol{A}$，即点积一般不满足交换律。而按点积定义结合律成立：

$$(\boldsymbol{A}\cdot\boldsymbol{B})\cdot\boldsymbol{C}=\boldsymbol{A}\cdot(\boldsymbol{B}\cdot\boldsymbol{C})$$

两个二阶张量的点积可用矩阵的乘法来表示。

双点积：若在两个张量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 作张量乘积之后再进行两次缩并，则称为双点积。设 $\boldsymbol{A}$ 是 m 阶张量，$\boldsymbol{B}$ 是 n 阶张量，则双点积 $\boldsymbol{A}:\boldsymbol{B}$ 是 $m+n-4$ 阶的张量

$$A_{i_1\cdots i_{m-2}kl}B_{j_1\cdots j_{n-2}kl}\equiv(\boldsymbol{A}:\boldsymbol{B})_{i_1\cdots i_{m-2}j_1\cdots j_{n-2}}$$

7）张量的商律　如 $A_{i_m}=T_{i_mj_n}B_{j_n}$ 恒成立，已知 $\boldsymbol{A}$ 是 m 阶张量，$\boldsymbol{B}$ 是 n 阶张量，则 $\boldsymbol{T}=(T_{i_mj_n})$ 必是 $m+n$ 阶的张量。这种判别张量的法则称为张量的商律。

2.4　张量的梯度、散度和奥高公式

1）张量的梯度

$$\nabla=\mathrm{grad}=\left\{\frac{\partial}{\partial x_1},\ \frac{\partial}{\partial x_2},\ \frac{\partial}{\partial x_3}\right\}$$

称为梯度算子，它是一个向量算子。

设 $\boldsymbol{A}$ 是 n 阶张量，$\boldsymbol{A}=\{A_{i_n}\}$。简记

$$\frac{\partial}{\partial x_k}A_{i_n}=A_{i_n,k}$$

张量 $\boldsymbol{A}$ 的梯度定义为

$$\nabla\boldsymbol{A}=\mathrm{grad}A=\left\{\frac{\partial}{\partial x_k}A_{i_n}\right\}=A_{i_n,k}$$

∇A 是 $n+1$ 阶张量。

2）张量的散度　设 $\boldsymbol{A}$ 为 n 阶张量，其散度定义为

$$\nabla\cdot\boldsymbol{A}=\mathrm{div}\boldsymbol{A}=\left\{\frac{\partial}{\partial x_k}A_{k,i_2\cdots i_n}\right\}$$

它是由 $\nabla\boldsymbol{A}$ 缩并所得的 $n-1$ 阶的张量。

3）奥高公式（散度定理）　对于向量 $\boldsymbol{f}$ 而言，有熟知的奥高公式：

$$\int_S\boldsymbol{n}\cdot\boldsymbol{f}\mathrm{d}S=\int_V\nabla\cdot\boldsymbol{f}\mathrm{d}V$$

式中，S 为空间区域 V 的封闭周界面；$\boldsymbol{n}$ 为 S 的单位外法向量。

奥高公式成立的条件是：曲面 S 分段光滑，$\boldsymbol{f}$ 在区域 V 内有分段连续的偏导数。

将奥高公式推广到张量的情况，设 $\boldsymbol{A}$ 为 n 阶张量，则有

$$\int_S\boldsymbol{n}\cdot\boldsymbol{A}\mathrm{d}S=\int_V\nabla\cdot\boldsymbol{A}\mathrm{d}V=\int_V\mathrm{div}\boldsymbol{A}\mathrm{d}V \tag{2.2-5}$$

奥高公式在后面的公式推导中经常要用到。

2.5　各向同性张量

一个张量，如其每一分量都是坐标系作刚体旋转变换下的不变量，则称它是各向同性张量。若 $T=\{T_{i_1\cdots i_n}\}$ 是 n 阶各向同性张量，则恒有

$$T_{i'_1\cdots i'_n}=T_{i_1\cdots i_n}$$

对于四阶以下的各向同性张量有如下结论。

1）零阶张量（标量）都是各向同性的。

2）一阶张量（向量），除零向量以外都不是各向同性的。

3）二阶各向同性张量必为 $\lambda\delta_{ij}$ 的形式，λ 为一标量。

4）三阶各向同性张量必为 λe_{ijk} 的形式，其中，λ 为一标量，e_{ijk} 为置换张量。

5）四阶各向同性张量的形式必为

$$T_{ijkl}=\lambda\delta_{ij}\delta_{kl}+\alpha\delta_{ik}\delta_{jl}+\beta\delta_{il}\delta_{jk}$$

其中 λ，α，β 都是标量。当 T_{ijkl} 对前两个指标和后两个指标均为对称时，$\alpha=\beta$。记 $\mu=\alpha=\beta$，则四阶对称的各向同性张量的一般形式为

$$T_{ijkl}=\lambda\delta_{ij}\delta_{kl}+\mu(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})$$

2.6　二阶张量

（1）二阶张量的一些简单性质

设 $\boldsymbol{A}=(A_{ij})$ 是二阶张量，有以下定义。

1）转置张量 $\boldsymbol{A}^{T}=(A_{ji})$ 是 $\boldsymbol{A}$ 的转置张量，它满足 $A_{ij}^{T}=A_{ji}$。$\boldsymbol{A}^{T}$ 亦为二阶张量。

2）正交张量 如果 $\boldsymbol{A}\cdot\boldsymbol{A}^{T}=\boldsymbol{A}^{T}\cdot\boldsymbol{A}=\boldsymbol{I}$（$\boldsymbol{I}$ 为单位张量），则称 $\boldsymbol{A}$ 是正交张量。

3）对称张量和反对称张量 若 $\boldsymbol{A}^{T}=\boldsymbol{A}$ 即 $A_{ji}=A_{ij}$ 恒成立，则称 $\boldsymbol{A}$ 是对称张量。若 $\boldsymbol{A}^{T}=-\boldsymbol{A}$ 即 $A_{ji}=-A_{ij}$ 恒成立，则称 $\boldsymbol{A}$ 是反对称张量。二阶张量的对称性和反对称性是与坐标系无关的。

二阶对称张量可用6个数来表征；二阶反对称张量的主对角分量为零，故可仅用3个数来表征。

（2）二阶张量的主值、主方向和不变量

1）主值和主方向 设 $\boldsymbol{T}$ 是二阶张量，可把它看成是向量 $\boldsymbol{a}$ 到向量 $\boldsymbol{b}$ 的线性变换

$$\boldsymbol{T}\cdot\boldsymbol{a}=\boldsymbol{b}$$

若有 $\boldsymbol{a}\neq 0$，它对应的 $\boldsymbol{b}=\lambda\boldsymbol{a}$，即有

$$\boldsymbol{T}\cdot\boldsymbol{a}=\lambda\boldsymbol{a}$$

则称此 λ 是张量 $\boldsymbol{T}$ 的特征值；而相应的非零解 $\boldsymbol{a}$ 就称为张量 $\boldsymbol{T}$ 对应于特征值 λ 的特征向量；而特征向量 $\boldsymbol{a}$ 的方向就称为张量 $\boldsymbol{T}$ 的特征方向。

由于张量 $\boldsymbol{T}$ 的特征值 λ 是两个向量 $\boldsymbol{T}\cdot\boldsymbol{a}$ 与 $\boldsymbol{a}$ 的比值，故它必是与坐标系无关的标量。

$\boldsymbol{T}^{T}$ 的特征值和特征向量称为 $\boldsymbol{T}$ 的主值和主方向。当 $\boldsymbol{T}$ 为二阶对称张量时，其特征值和特征方向就是其主值和主方向。

2）二阶张量的特征方程和不变量 求二阶张量 $\boldsymbol{T}$ 的特征值和特征方向时，应解方程 $\boldsymbol{T}\cdot\boldsymbol{a}=\lambda\boldsymbol{a}$，将此方程写作

$$(\boldsymbol{T}-\lambda\boldsymbol{I})\cdot\boldsymbol{a}=0$$

其中 $\boldsymbol{I}=(\delta_{ij})$ 是单位张量。欲使它有非零解 $\boldsymbol{a}$（即 $\boldsymbol{a}$ 的分量 a_1，a_2，a_3 不全为零），必须

$$\det(\boldsymbol{T}-\lambda\boldsymbol{I})=\begin{vmatrix} T_{11}-\lambda & T_{12} & T_{13} \\ T_{21} & T_{22}-\lambda & T_{23} \\ T_{31} & T_{32} & T_{33}-\lambda \end{vmatrix}=0$$

其中 det 表示取行列式，于是得到关于 λ 的三次方程式

$$\lambda^3-T_{kk}\lambda^2+\frac{1}{2}(T_{ii}T_{jj}-T_{ij}T_{ji})\lambda-\det\boldsymbol{T}=0 \quad (2.2\text{-}6)$$

称式（2.2-6）为张量 $\boldsymbol{T}$ 的特征方程。

由此可知，二阶张量 $\boldsymbol{T}$ 至多有三个不同的特征值，若 T_{ij} 皆实数，则 $\boldsymbol{T}$ 至少有一个实的特征值。

设方程（2.2-6）的三个根（也即 $\boldsymbol{T}$ 的三个主值）为 λ_1，λ_2，λ_3，则式（2.2-6）可写成

$$\det(\boldsymbol{T}-\lambda I)=-(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$$

由根与系数的关系，可知

$$\lambda_1+\lambda_2+\lambda_3=T_{kk}=I_1$$

$$\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3=(T_{ii}T_{jj}-T_{ij}T_{ji})/2=I_2$$

$$\lambda_1\lambda_2\lambda_3=\det\boldsymbol{T}=I_3$$

由于主值 λ_i 都是坐标变换下的不变量，因此 I_1，I_2，I_3 也都是坐标变换下的不变量。分别称 I_1，I_2，I_3 为 $\boldsymbol{T}$ 的第一、第二和第三不变量。

（3）二阶对称张量的性质

一个二阶对称张量 $\boldsymbol{T}$ 具有以下性质：

1）$\boldsymbol{T}$ 的三个主值都是实数；

2）$\boldsymbol{T}$ 对应于不同特征值的两特征向量必正交；

3）$\boldsymbol{T}$ 恒有三个互相垂直的主轴方向。

（4）二阶张量的分解

1）分解为对称张量和反对称张量之和 任一二阶张量可惟一地分解成一个对称张量和另一个反对称张量之和，如下所示

$$A_{ij}=\frac{1}{2}(A_{ij}+A_{ji})+\frac{1}{2}(A_{ij}-A_{ji})$$

右边第一项是对称张量，第二项是反对称张量。

2）分解为球张量和偏张量之和 有时将一个二阶张量分解成一个球张量和一个偏张量之和，如下所示

$$A_{ij}=A'_{ij}+\delta_{ij}A_m$$

其中

$$A_m=\frac{1}{3}A_{ii}$$

$$A'_{ij}=A_{ij}-\delta_{ij}A_m$$

A_m 称为球张量，A'_{ij} 称为偏张量。

3）极分解 设二阶张量 $\boldsymbol{F}$ 的行列式 $\det\boldsymbol{F}\neq 0$，则可对 $\boldsymbol{F}$ 惟一地进行如下极分解

$$\boldsymbol{F}=\boldsymbol{R}\cdot\boldsymbol{U}$$

$$\boldsymbol{F}=\boldsymbol{V}\cdot\boldsymbol{R} \quad (2.2\text{-}7)$$

式中，$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 是二阶正定的对称张量，$\boldsymbol{R}$ 为二阶正交张量。

3 运动与变形

3.1 变形几何学

3.1.1 运动与变形的描述

我们把物体中所有物质点瞬时位置的集合——某一瞬时物体在空间占据的区域 V 定义为该物体的构形。为了描述物体的运动和变形，需要选择某一特定时刻的构形作为参考构形，以确定每一时刻每一质点的位置。如果已知 t 时刻以前所有时刻的构形的力学特性，而要研究的是 t 时刻现时构形的变形问题，那么可以把初始构形 V_0 或 t 时刻以前的任一时刻的构形作为参考构形。

物体质点的位置，可以用质点在坐标系中的位置坐标表示。采用笛卡儿直角坐标系作为参考坐标系。在三维坐标参考系中（图2.2-1），设在 $t=t_0=0$ 时刻物体任一质点 a 的位置坐标为 X_i（$i=1,2,3$），此时的构形为初始构形，记为 V_0。此后某一时刻 t，物体运动到一个新位置，各质点间的相互位置关系发生了变化，物体产生了变形，质点 a 的位置坐标从 X_i 改变为 x_i（$i=1,2,3$）。显然 x_i 是 X_i 和时间 t 的函数，即

$$x_i=x_i(X_1,X_2,X_3,t) \quad (2.2\text{-}8)$$

变形态构形或变形后构形记为 V。

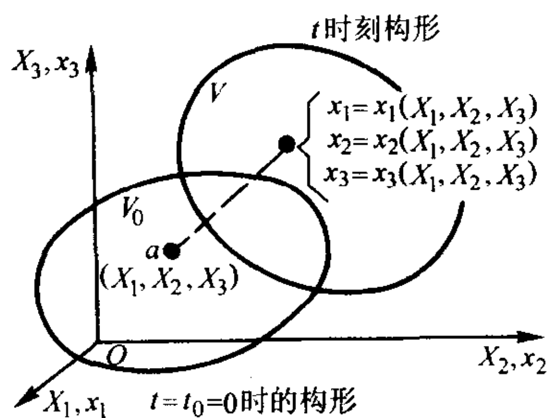


图2.2-1 物体的构形

假设物体及其运动和变形都是连续的，则 V_0 中每一质点 X_i 仅与 V 中一个质点 x_i 对应，反之亦然。于是，可认为函数 $x_i(X_1,X_2,X_3,t)$ 是单值、连续和可微的，且雅可比（Jacobi）行列式不等于零，即

$$J=\left|\frac{\partial x_i}{\partial X_j}\right|=\begin{vmatrix} \frac{\partial x_1}{\partial X_1} & \frac{\partial x_1}{\partial X_2} & \frac{\partial x_1}{\partial X_3} \\ \frac{\partial x_2}{\partial X_1} & \frac{\partial x_2}{\partial X_2} & \frac{\partial x_2}{\partial X_3} \\ \frac{\partial x_3}{\partial X_1} & \frac{\partial x_3}{\partial X_2} & \frac{\partial x_3}{\partial X_3} \end{vmatrix}\neq 0 \quad (2.2\text{-}9)$$

也有

$$X_i = X_i\ (x_1,\ x_2,\ x_3,\ t) \tag{2.2-10}$$

$$j = \left|\frac{\partial X_i}{\partial x_j}\right| \neq 0 \tag{2.2-11}$$

$$j = J^{-1} \tag{2.2-12}$$

质点 a 的初始坐标 X_i 与时间 t 无关，不同的初始坐标代表不同的质点，可以把坐标 X_i 看作是识别物质点的“标志”，所以称 X_i 为物质坐标，物质坐标也称拉格朗日（Lagrange）坐标；坐标 x_i 是识别空间质点的“标志”，同一空间点在不同时刻由不同的物质点所占据，所以 x_i 称为空间坐标，也称欧拉（Euler）坐标。

在连续介质力学中，对物体质点的运动，可以有两种描述和研究方法。一种方法是把物体质点的运动和物体的各物理量看成是物质坐标 X_i 和时间 t 的函数，并研究这些函数的变化规律，这样的描述和研究方法称为拉格朗日方法。另一种方法是把物体质点的运动和物体的各物理量看成是空间坐标 x_i 和时间 t 的函数，并研究这些函数的变化规律，这样的描述和研究方法称为欧拉方法。

变形几何学研究连续体变形的纯几何性质。在此，将 t 当作一个固定的值，研究物体由初始构形到 t 时刻构形之间所发生的运动与变形。

3.1.2　变形梯度

由于 $\boldsymbol{x}$ 是 $\boldsymbol{X}$ 的单值、连续和可微的函数，故有

$$\mathrm{d}x_i = \frac{\partial x_i}{\partial X_j}\mathrm{d}X_j\ \ (t\ 固定)$$

式中 $\partial x_i/\partial X_j = F_{ij}$ 称为变形梯度，它是一个二阶张量。于是上式可写成如下张量形式

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{X} \tag{2.2-13}$$

可见，$\boldsymbol{F}$ 将初始构形中任一线元 $\mathrm{d}\boldsymbol{X}$ 变换成现时构形中的对应线元 $\mathrm{d}\boldsymbol{x}$。它描述了运动与变形过程中任一线元长度和方向的变化。式（2.2-9）中的雅可比行列式即是 $\det\boldsymbol{F}$，可知 $\det\boldsymbol{F}\neq 0$，于是可以对 $\boldsymbol{F}$ 惟一地进行如下极分解

$$\boldsymbol{F} = \boldsymbol{R}\cdot\boldsymbol{U}$$

或

$$\boldsymbol{F} = \boldsymbol{V}\cdot\boldsymbol{R}$$

其中 $\boldsymbol{R}$ 为二阶正交张量，它代表刚体转动；$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 是二阶正定的对称张量，它们均代表纯变形。

$\boldsymbol{F} = \boldsymbol{R}\cdot\boldsymbol{U}$ 可看成先变形后转动的合成

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{x} = \boldsymbol{R}\cdot(\boldsymbol{U}\cdot\mathrm{d}\boldsymbol{X})$$

即 $\mathrm{d}\boldsymbol{X}$ 经过变形得到 $\mathrm{d}\tilde{\boldsymbol{x}} = \boldsymbol{U}\cdot\mathrm{d}\boldsymbol{x}$，再经过刚体转动 $\boldsymbol{R}$，而得到 $\mathrm{d}\boldsymbol{x} = \boldsymbol{R}\cdot\mathrm{d}\tilde{\boldsymbol{x}}$。应该指出，除了 $\boldsymbol{U}$ 的主轴方向外，$\mathrm{d}\tilde{\boldsymbol{x}}$ 与 $\mathrm{d}\boldsymbol{X}$ 的方向并不一致，这是因为剪切变形会引起方向的改变。

类似地，另一种分解 $\boldsymbol{F} = \boldsymbol{V}\cdot\boldsymbol{R}$ 可看成先转动后变形的合成。两种分解中的转动部分 $\boldsymbol{R}$ 相同，而纯变形张量之间则有 $\boldsymbol{U} = \boldsymbol{R}^{\mathrm{T}}\cdot\boldsymbol{V}\cdot\boldsymbol{R}$，如图2.2-2所示。

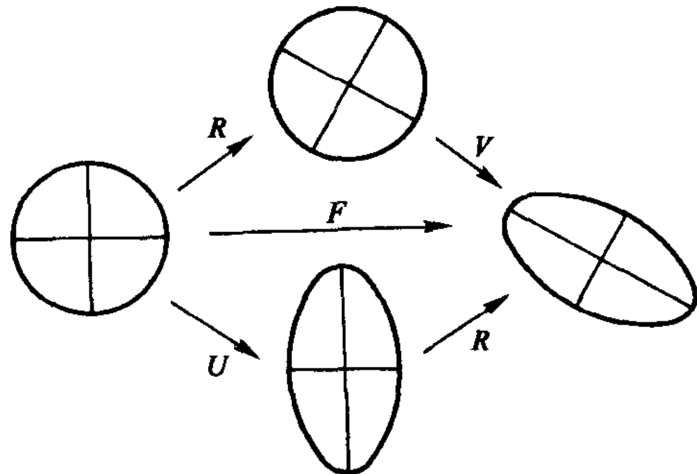


图2.2-2　变形梯度的极分解

3.1.3　变形张量

图2.2-3表示三维空间笛卡儿直角坐标系中两相邻的质点 $P\ (X_i)$ 和 $Q\ (X_i+\mathrm{d}X_i)$ 所构成的线元 $\mathrm{d}\boldsymbol{X}$ 的运动与变形。在初时刻，线元长度为 $\mathrm{d}L_0$

$$(\mathrm{d}L_0)^2 = (\mathrm{d}X_1)^2 + (\mathrm{d}X_2)^2 + (\mathrm{d}X_3)^2 = \mathrm{d}X_i\mathrm{d}X_i = \mathrm{d}\boldsymbol{X}^{\mathrm{T}}\cdot\mathrm{d}\boldsymbol{X} \tag{2.2-14}$$

在终时刻，此两质点到达新的位置 $P'\ (x_i)$ 和 $Q'\ (x_i+\mathrm{d}x_i)$，其间的距离为 $\mathrm{d}L$

$$\begin{aligned}(\mathrm{d}L)^2 &= \mathrm{d}x_i\mathrm{d}x_i \\ &= (F_{im}\mathrm{d}X_m)(F_{in}\mathrm{d}X_n) \\ &= C_{mn}\mathrm{d}X_m\mathrm{d}X_n \\ &= \mathrm{d}\boldsymbol{X}\cdot\boldsymbol{C}\cdot\mathrm{d}\boldsymbol{X}\end{aligned} \tag{2.2-15}$$

其中

$$\boldsymbol{C} = [C_{mn}] = \boldsymbol{F}^{\mathrm{T}}\cdot\boldsymbol{F}$$

称为右柯西－格林变形张量。另外

$$\boldsymbol{B} = \boldsymbol{F}\cdot\boldsymbol{F}^{\mathrm{T}}$$

称为左柯西－格林变形张量。

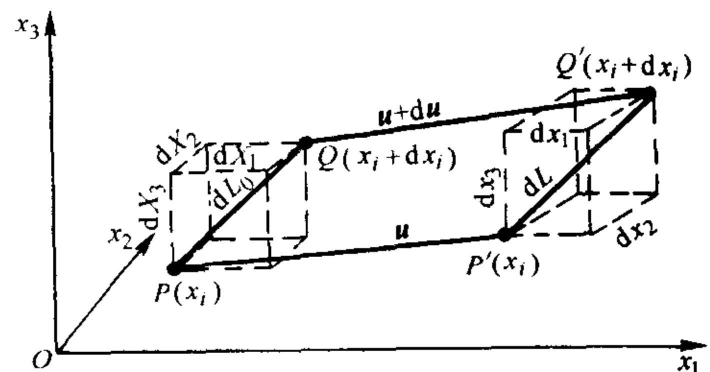


图2.2-3　线元的运动与变形

定义相对伸长为 $\Lambda = \|\mathrm{d}\boldsymbol{x}\|/\|\mathrm{d}\boldsymbol{X}\|$，这里 $\|\mathrm{d}\boldsymbol{x}\|$ 表示向量 $\mathrm{d}\boldsymbol{x}$ 的模。则 λ 可以用 $\boldsymbol{C}$ 或 $\boldsymbol{B}$ 来表征。

$$\lambda^2 = \left(\frac{\mathrm{d}L}{\mathrm{d}L_0}\right)^2 = \frac{\mathrm{d}\boldsymbol{X}\cdot\boldsymbol{C}\cdot\mathrm{d}\boldsymbol{X}}{\|\mathrm{d}\boldsymbol{X}\|^2} = \frac{\mathrm{d}\boldsymbol{X}}{\|\mathrm{d}\boldsymbol{X}\|}\cdot\boldsymbol{C}\cdot\frac{\mathrm{d}\boldsymbol{X}}{\|\mathrm{d}\boldsymbol{X}\|} = \boldsymbol{N}\cdot\boldsymbol{C}\cdot\boldsymbol{N}$$

$$\lambda^2 = \left(\frac{\mathrm{d}L}{\mathrm{d}L_0}\right)^2 = \frac{\|\mathrm{d}\boldsymbol{x}\|^2}{\mathrm{d}\boldsymbol{x}\cdot\boldsymbol{B}^{-1}\cdot\mathrm{d}\boldsymbol{x}} = \frac{1}{\dfrac{\mathrm{d}\boldsymbol{x}}{\|\mathrm{d}\boldsymbol{x}\|}\cdot\boldsymbol{B}^{-1}\cdot\dfrac{\mathrm{d}\boldsymbol{x}}{\|\mathrm{d}\boldsymbol{x}\|}} = \frac{1}{\boldsymbol{n}\cdot\boldsymbol{B}^{-1}\cdot\boldsymbol{n}}$$

其中，$\boldsymbol{N}$ 和 $\boldsymbol{n}$ 分别为沿线元 $\mathrm{d}\boldsymbol{X}$ 和 $\mathrm{d}\boldsymbol{x}$ 的单位矢量。由于 $\boldsymbol{C}$、$\boldsymbol{B}$ 均为二阶对称张量，故均有三个两两正交的主方向，在这三个主方向上取主值。对于 $\boldsymbol{C}$，以主方向 N_1、N_2、N_3 为基矢量而构成的主轴坐标系中，仅 C_{ii}（$i=1,\ 2,\ 3$）可不为零，且有

$$\lambda_i^2 = C_{ii}\ \ (i=1,\ 2,\ 3) \tag{2.2-16}$$

式（2.2-16）中的 λ_i 也即是式（2.2-7）中 $\boldsymbol{U}$ 的主分量。求得了 $\boldsymbol{U}$ 以后，式中的 $\boldsymbol{R}$ 可由 $\boldsymbol{R} = \boldsymbol{F}\cdot\boldsymbol{U}^{-1}$ 求得，这样就实现了 $\boldsymbol{F}$ 的极分解。

3.1.4　线、面、体元的变化

（1）线元的变化

变形前的线元 $\mathrm{d}\boldsymbol{X}$ 在变形后成为 $\mathrm{d}\boldsymbol{x}$。它们变换关系见式（2.2-13）。

如图2.2-4所示，可以通过线元定义面元和体元，通过矢量运算求得面元和体元的变换。

（2）体元的变化

考虑在初始构形中以任意三个线元 $\mathrm{d}\boldsymbol{A}$、$\mathrm{d}\boldsymbol{B}$ 和 $\mathrm{d}\boldsymbol{C}$ 为棱边所构成的体元 $\mathrm{d}V_0$，变形后这个体元变成现时构形中以 $\mathrm{d}\boldsymbol{a}$、$\mathrm{d}\boldsymbol{b}$ 和 $\mathrm{d}\boldsymbol{c}$ 为棱边所构成的体元 $\mathrm{d}V$。由矢量的混合积可分别求得 $\mathrm{d}V_0$ 和 $\mathrm{d}V$ 如下［参见式（2.2-4）］

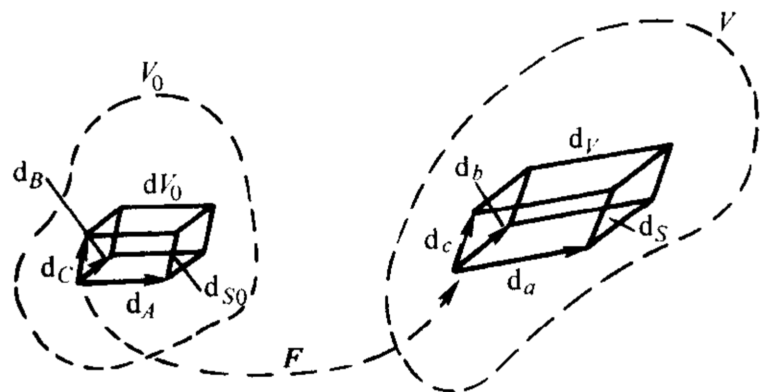


图 2.2-4　面元和体元的变化

$$dV_0 = (d\boldsymbol{A}\times d\boldsymbol{B})\cdot d\boldsymbol{C} = e_{IJK}dA_I dB_J dC_K = \begin{vmatrix} dA_1 & dA_2 & dA_3 \\ dB_1 & dB_2 & dB_3 \\ dC_1 & dC_2 & dC_3 \end{vmatrix}$$

$$dV = (d\boldsymbol{a}\times d\boldsymbol{b})\cdot d\boldsymbol{c} = [(\boldsymbol{F}\cdot d\boldsymbol{A})\times(\boldsymbol{F}\cdot d\boldsymbol{B})]\cdot(\boldsymbol{F}\cdot d\boldsymbol{C}) = e_{rst}F_{rI}dA_I F_{sJ}dB_J F_{tK}dC_K$$

由于 $e_{rst}F_{rI}F_{sJ}F_{tK}=e_{IJK}\det\boldsymbol{F}$，所以

$$dV = e_{IJK}dA_I dB_J dC_K(\det\boldsymbol{F}) = JdV_0$$

式中，为区分不同构形中的分量，对初始构形和现时构形中的分量分别用大写和小写的下标。可见，J 表示了微元体积的变化，并有

$$J=\det\boldsymbol{F}=\frac{dV}{dV_0}>0 \tag{2.2-17}$$

(3) 面元的变化

由任意两个线元所构成的面元，用这两个线元的叉积来表示。它是一个垂直于该面元的向量。考虑在初始构形中由线元 $d\boldsymbol{A}$ 和 $d\boldsymbol{B}$ 所构成的面元 dS_0，其单位法矢为 $\boldsymbol{N}$，变形后这个面元变成现时构形中由 $d\boldsymbol{a}$ 和 $d\boldsymbol{b}$ 所构成的面元 dS，其单位法矢为 $\boldsymbol{n}$，

$$\boldsymbol{N}dS_0 = d\boldsymbol{A}\times d\boldsymbol{B}$$

叉乘也可用行列式表示，展开后其分量式为

$$N_I dS_0 = e_{IJK}dA_J dB_K$$

类似地

$$\boldsymbol{n}dS = d\boldsymbol{a}\times d\boldsymbol{b} = (\boldsymbol{F}\cdot d\boldsymbol{A})\times(\boldsymbol{F}\cdot d\boldsymbol{B})$$

$$n_r dS = e_{rst}da_s db_t = e_{rst}F_{sJ}dA_J F_{tK}dB_K$$

上式两边同乘 F_{rI}，得

$$F_{rI}n_r dS = e_{rst}F_{rI}F_{sJ}F_{tK}dA_J dB_K = e_{IJK}dA_J dB_K(\det\boldsymbol{F}) = JN_I dS_0$$

最后得到

$$\boldsymbol{F}^T\cdot\boldsymbol{n}dS = J\boldsymbol{N}dS_0 \tag{2.2-18}$$

上式称为南生（Nanson）公式。

3.1.5　应变张量

(1) 格林应变张量

物体变形的基本标志是物体中质点之间的距离发生了变化。参考式（2.2-14）和式（2.2-15），采用拉格朗日描述，线元长度平方的改变量为

$$(dL)^2-(dL_0)^2 = d\boldsymbol{X}\cdot\boldsymbol{C}\cdot d\boldsymbol{X}-d\boldsymbol{X}^T\cdot d\boldsymbol{X} = d\boldsymbol{X}\cdot(\boldsymbol{C}-\boldsymbol{I})\cdot d\boldsymbol{X}$$

定义格林（Green）应变张量为

$$\boldsymbol{E}=\frac{1}{2}(\boldsymbol{C}-\boldsymbol{I}) \tag{2.2-19}$$

考虑到坐标与位移之间的如下关系

$$\boldsymbol{x}=\boldsymbol{X}+\boldsymbol{u} \tag{2.2-20}$$

可将格林应变的分量式写成

$$E_{ij}=\frac{1}{2}\left(\delta_{\alpha\beta}\frac{\partial x_\alpha}{\partial X_i}\frac{\partial x_\beta}{\partial X_j}-\delta_{ij}\right) = \frac{1}{2}\left(\frac{\partial u_i}{\partial X_j}+\frac{\partial u_j}{\partial X_i}+\frac{\partial u_k}{\partial X_i}\frac{\partial u_k}{\partial X_j}\right) \tag{2.2-21}$$

(2) 阿尔曼斯应变张量

采用欧拉描述，线元长度平方的改变量为

$$(dL)^2-(dL_0)^2 = d\boldsymbol{x}^T\cdot d\boldsymbol{x}-(\boldsymbol{F}^{-1}d\boldsymbol{x})^T\cdot(\boldsymbol{F}^{-1}d\boldsymbol{x}) = d\boldsymbol{x}\cdot(\boldsymbol{I}-\boldsymbol{B}^{-1})\cdot d\boldsymbol{x}$$

定义阿尔曼斯（Almansi）应变张量为

$$\boldsymbol{e}=\frac{1}{2}(\boldsymbol{I}-\boldsymbol{B}^{-1}) \tag{2.2-22}$$

其分量式为

$$e_{ij}=\frac{1}{2}\left(\delta_{ij}-\delta_{\alpha\beta}\frac{\partial X_\alpha}{\partial x_i}\frac{\partial X_\beta}{\partial x_j}\right) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}-\frac{\partial u_k}{\partial x_i}\frac{\partial u_k}{\partial x_j}\right) \tag{2.2-23}$$

由格林应变张量 $\boldsymbol{E}$ 和阿尔曼斯应变张量 $\boldsymbol{e}$ 的定义式，可求得它们之间的关系式

$$e_{ij}=\frac{\partial X_K}{\partial x_i}\frac{\partial X_L}{\partial x_j}E_{KL},\quad E_{KL}=\frac{\partial x_i}{\partial X_K}\frac{\partial x_j}{\partial X_L}e_{ij} \tag{2.2-24}$$

(3) 小应变张量

当 $\left|\frac{\partial u_i}{\partial x_j}\right|\ll 1$ 对所有的 i，j 都成立时，所有的 $\frac{\partial u_i}{\partial x_j}$ 都是小量，二阶小量可以忽略。这种情况下不仅变形小，而且刚体旋转也小。利用式（2.2-20），则

$$\frac{\partial u_i}{\partial X_j}=\frac{\partial u_i}{\partial x_m}\frac{\partial x_m}{\partial X_j}=\frac{\partial u_i}{\partial x_m}\frac{\partial(X_m+u_m)}{\partial X_j} = \frac{\partial u_i}{\partial x_m}\left(\delta_{mj}+\frac{\partial u_m}{\partial X_j}\right)\approx\frac{\partial u_i}{\partial x_j}$$

即 $\partial u_i/\partial X_j$ 与 $\partial u_i/\partial x_j$ 的差别可以忽略。由式（2.2-21）和式（2.2-23）忽略二阶小量后，得到

$$E_{ij}=\frac{1}{2}\left(\frac{\partial u_i}{\partial X_j}+\frac{\partial u_j}{\partial X_i}\right)$$

$$e_{ij}=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)$$

且 $E_{ij}\approx e_{ij}$

通常将此应变张量记作 ε_{ij}，称为小应变张量。

(4) 主应变和对数应变

由式（2.2-19）和关于 $\boldsymbol{C}$ 的主值的讨论可知，格林应变张量 $\boldsymbol{E}$ 的主方向与 $\boldsymbol{C}$ 的主方向一致，$\boldsymbol{E}$ 的主值为

$$\frac{1}{2}(\lambda_1^2-1),\frac{1}{2}(\lambda_2^2-1),\frac{1}{2}(\lambda_3^2-1)$$

其中，λ_i^2 即为式（2.2-16）所给出的 $\boldsymbol{C}$ 的主值。

定义对数应变张量 $\boldsymbol{H}$ 为

$$\boldsymbol{H}=\frac{1}{2}\ln\boldsymbol{C}=\frac{1}{2}\ln(\boldsymbol{I}+2\boldsymbol{E}) = \frac{1}{2}\left[2\boldsymbol{E}-\frac{1}{2}(2\boldsymbol{E})^2+\frac{1}{3}(2\boldsymbol{E})^3+\cdots\right]$$

在 $\boldsymbol{C}$ 的主轴坐标系中可求得其主值为

$$\ln\lambda_1,\ \ln\lambda_2,\ \ln\lambda_3$$

以上是针对拉格朗日描述进行的讨论，针对欧拉描述可进行类似的讨论。

当物体无变形时（只有刚体旋转），所有的应变张量的分量均为零。

(5) 其他坐标系下小应变张量的计算公式

虽然本章主要采用笛卡儿坐标系，但是在实际问题中，常常遇到具有轴对称和球对称的问题，或者可以将实际问题

简化为具有以上对称性的问题，这样能够突出问题的主要特点并大大地简化分析计算。下面给出圆柱坐标系和球系坐标系中小应变张量的计算公式。

1）圆柱坐标系中　令 u，v，w 分别为沿 ρ，θ，z 方向的位移分量，则有

$$\varepsilon_\rho=\frac{\partial u}{\partial\rho},\ \varepsilon_\theta=\frac{1}{\rho}\left(\frac{\partial v}{\partial\theta}+u\right),\ \varepsilon_z=\frac{\partial w}{\partial z}$$

$$\gamma_{\theta z}=\frac{1}{2}\left(\frac{\partial v}{\partial z}+\frac{1}{\rho}\frac{\partial w}{\partial\theta}\right),\ \gamma_{z\rho}=\frac{1}{2}\left(\frac{\partial w}{\partial\rho}+\frac{\partial u}{\partial z}\right)\quad(2.2\text{-}25)$$

$$\gamma_{\rho\theta}=\frac{1}{2}\left(\frac{\partial v}{\partial\rho}-\frac{v}{\rho}+\frac{1}{\rho}\frac{\partial u}{\partial\theta}\right)$$

轴对称状态时

$$\varepsilon_\rho=\frac{\partial u}{\partial\rho},\ \varepsilon_\theta=\frac{u}{\rho},\ \varepsilon_z=\frac{\partial w}{\partial z}$$

$$\gamma_{\theta z}=0,\gamma_{z\rho}=\frac{1}{2}\left(\frac{\partial w}{\partial\rho}+\frac{\partial u}{\partial z}\right),\gamma_{\rho\theta}=0\quad(2.2\text{-}26)$$

2）球坐标系中　令 u，v，w 分别为沿 ρ，φ，θ 方向的位移分量，则有

$$\varepsilon_\rho=\frac{\partial u}{\partial\rho},\ \varepsilon_\varphi=\frac{1}{\rho}\left(\frac{\partial v}{\partial\varphi}+u\right)$$

$$\varepsilon_\theta=\frac{1}{\rho}\left(\frac{1}{\sin\varphi}\frac{\partial w}{\partial\theta}+u+v\cot\varphi\right)$$

$$\gamma_{\varphi\theta}=\frac{1}{2\rho}\left(\frac{1}{\sin\varphi}\frac{\partial w}{\partial\theta}+\frac{\partial w}{\partial\varphi}-w\cot\varphi\right)\quad(2.2\text{-}27)$$

$$\gamma_{\rho\theta}=\frac{1}{2\rho}\left(\frac{1}{\sin\varphi}\frac{\partial u}{\partial\theta}+\rho\frac{\partial w}{\partial\rho}-w\right)$$

$$\gamma_{\rho\varphi}=\frac{1}{2\rho}\left(\rho\frac{\partial v}{\partial\rho}-v+\frac{\partial u}{\partial\varphi}\right)$$

3.2　运动学

运动学研究运动和变形与时间的关系。因此本节在物质坐标与空间坐标的相互关系式（2.2-8）和式（2.2-10）中，要考虑变化着的时间。

3.2.1　物质导数和局部导数

既然对于运动与变形，有拉格朗日描述与欧拉描述这两种不同的描述方法，相应地就有两种不同意义的对时间的导数。

在拉格朗日描述（物质描述）中，当 $\boldsymbol{X}$ 固定（即对应于一定的质点），任一物理量（标量、矢量或张量）对时间的导数，称为该物理量的物质导数（又称为随体微商），这里记作 $\frac{\mathrm{D}}{\mathrm{D}t}$。

函数 f 以 $\boldsymbol{X}$，t 为自变量时，$f=f(\boldsymbol{X},t)$，求 f 的物质导数很简单：

$$\frac{\mathrm{D}f}{\mathrm{D}t}=\frac{\partial f}{\partial t}$$

其中一个特例是质点的速度

$$\boldsymbol{v}(\boldsymbol{X},t)=\frac{\partial \boldsymbol{u}}{\partial t}=\frac{\partial(\boldsymbol{x}-\boldsymbol{X})}{\partial t}=\frac{\partial \boldsymbol{x}}{\partial t}$$

在欧拉描述（空间描述）中，当 $\boldsymbol{x}$ 固定（即对应于一定的空间位置），任一物理量（标量、向量或张量）对时间的导数，称为该物理量的局部导数，记作 $\left(\frac{\partial\cdots}{\partial t}\right)_x$，有时简记为 $\frac{\partial\cdots}{\partial t}$。

应该注意，物质导数是力学中的本质量，而局部导数则是观测量，即在空间位置处观测到的量。

当以欧拉变量为自变量时，$f=f(\boldsymbol{x},t)$，这时 f 的物质导数应为

$$\frac{\mathrm{D}f}{\mathrm{D}t}=\frac{\partial f}{\partial t}+\frac{\partial f}{\partial x_\alpha}\left.\frac{\partial x_\alpha}{\partial t}\right|_{\boldsymbol{x}不变}$$

由于 $\left.\frac{\partial x}{\partial t}\right|_{\boldsymbol{x}不变}$ 就是质点的速度 $\boldsymbol{v}$，即 $v_\alpha=\left.\frac{\partial x_\alpha}{\partial t}\right|_{\boldsymbol{x}不变}$，因此欧拉变量下计算物质导数的公式为

$$\frac{\mathrm{D}}{\mathrm{D}t}=\frac{\partial}{\partial t}+v_\alpha\frac{\partial}{\partial x_\alpha}\quad(2.2\text{-}28)$$

3.2.2　速度梯度及其分解

如图2.2-5所示，在某一瞬时的现时构形中，任一质点 P 具有速度 $\boldsymbol{v}$，在 P 的邻域内任一点 Q 具有速度 $\boldsymbol{v}+\mathrm{d}\boldsymbol{v}$。在同一时刻 t，由位置坐标不同造成的速度变化就是

$$\mathrm{d}v_i=\frac{\partial v_i}{\partial x_j}\mathrm{d}x_j$$

式中，$v_{i,j}=\partial v_i/\partial x_j$ 称为速度梯度，记为 L_{ij}，它是一个二阶张量。如果说变形梯度 $\boldsymbol{F}$ 表明一段时间中运动与变形的大小，则速度梯度 $\boldsymbol{L}$ 表明某一时刻运动与变形的快慢。

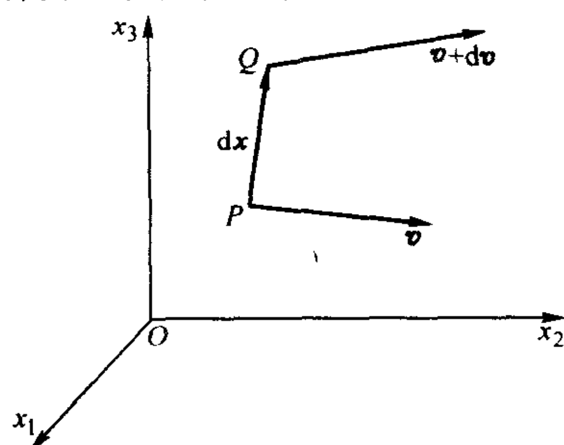


图2.2-5　质点速度在空间中的变化

将L分解为一个对称张量d和一个反对称张量ω之和，

$$\boldsymbol{d}=\frac{1}{2}(\boldsymbol{L}+\boldsymbol{L}^{\mathrm{T}}),\ \boldsymbol{\omega}=\frac{1}{2}(\boldsymbol{L}-\boldsymbol{L}^{\mathrm{T}})\quad(2.2\text{-}29)$$

式中　$\boldsymbol{d}$——应变速率；

$\boldsymbol{\omega}$——旋转速率，它表征刚体旋转的速率。

3.2.3　增量分析

涉及非线性问题时，通常是采用增量分析方法。分析中涉及到3个构形，如图2.2-6所示。其中 $t=0$ 时的构形为初始构形，$t=t$ 时的构形为经过 n 次增量计算以后得到的现时构形，而 $t=\tau$ 时的构形为下一增量步中待求的新构形。在此增量步的计算中，认为 t 为定值，τ 为变量。$t=\tau$ 时构形中质点的坐标用 $\boldsymbol{\xi}$ 表示，可以将 $\boldsymbol{\xi}$ 表示为 $\mathbf{X}$ 或 $\mathbf{x}$ 的函数，即以初始构形或现时构形作为参考构形，这两种描述方法分别称为全拉格朗日法（TL法）和更新的拉格朗日法（UL法）。

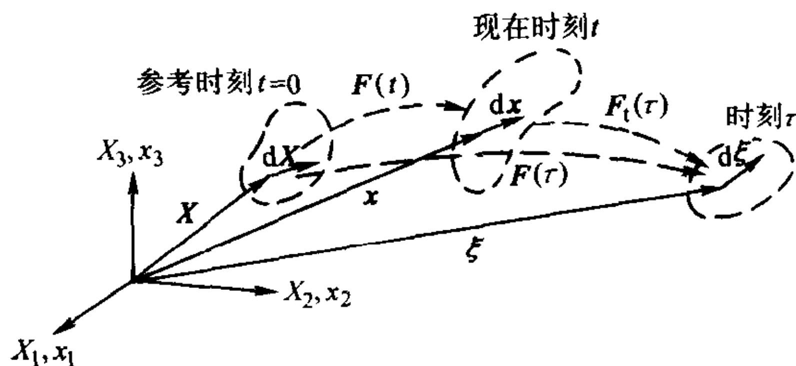


图2.2-6　运动与变形的增量描述

若采用UL法

$$\boldsymbol{\xi}=\chi_t(\boldsymbol{x},t)$$

于是

$$\frac{\partial\xi_i}{\partial X_j}=\frac{\partial\xi_i}{\partial x_k}\frac{\partial x_k}{\partial X_j}$$

即

$$\boldsymbol{F}(\tau)=\boldsymbol{F}_t(\tau)\cdot\boldsymbol{F}(t)$$

因为将 t 作为定值，则 $\boldsymbol{F}(t)$ 亦为定值，令

$$\dot{\boldsymbol{F}}(t)=\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{F}(\tau)\bigg|_{\tau=t}=\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{F}(t)$$

$$\dot{F}_t(t)=\frac{\mathrm{d}}{\mathrm{d}t}F_t|_{\tau=t}$$

则有

$$\dot{F}(t)=\dot{F}_t(t)\cdot F(t)$$

由上式求得速度梯度为

$$L(t)=\dot{F}_t(t)=\dot{F}(t)\cdot F^{-1}(t)=d+\omega$$

另一方面，对 $F_t(t)$ 进行极分解得

$$F_t(t)=R_t(t)\cdot U_t(t)$$

将上式对时间求导，并注意到当 $\tau\to t$ 时，$R_t(t)=U_t(t)=I$，得

$$L=\dot{F}_t(t)=\dot{R}_t(t)\cdot U_t(t)+R_t(t)\cdot\dot{U}_t(t)=\dot{R}_t(t)+\dot{U}(t)$$

由于 $R_t\cdot R_t^{\mathrm{T}}=I$，$\dot{R}_t+\dot{R}_t^{\mathrm{T}}=0$，可知 $\dot{R}_t$ 为反对称张量。比较以上 L 的两种分解方式知

$$d=\dot{U}_t(t),\quad \omega=\dot{R}_t(t)$$

3.2.4 面元和体元的变化率

体积变化率为

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{V}{V_0}\right)=\frac{\mathrm{d}J}{\mathrm{d}t}=J\mathrm{tr}L \tag{2.2-30}$$

式中，$\mathrm{tr}L=L_{ii}=v_{i,i}$，称为 L 的迹。

面元的变化率为

$$\frac{\mathrm{d}}{\mathrm{d}t}(n_i\mathrm{d}S)=(-v_{j,i}+v_{k,k}\delta_{ij})n_j\mathrm{d}S \tag{2.2-31}$$

单位法矢的变化率为

$$\dot{n}=[n\cdot(L\cdot n)]n-L^{\mathrm{T}}\cdot n \tag{2.2-32}$$

4 应力

4.1 体力和面力

作用在连续介质全部质点上的有距离力（即不需要和物体接触，隔一段距离仍能起作用的力），称为体力。其大小正比于物体的质量，故又称质量力，如重力、惯性力等。将单位质量所受的体力记为 b。

作用于面（表面或内部面）上的质点，只当物体直接接触才起作用的力，称为面力。面力的大小和方向都和作用面的方面有关。方向一定的面元上面力的大小又和面元的面积成正比。如压力、摩擦力等。单位面积的面力称为应力矢量，其定义为

$$p=\lim_{\Delta S\to 0}\frac{\Delta f}{\Delta S}=\frac{\mathrm{d}f}{\mathrm{d}S}$$

其中 Δf 是作用于面元 ΔS 的面力，如图 2.2-7 所示。应力矢量 p 可进一步分解为沿面元 ΔS 的法矢 n 方向的正应力 σ 和位于切平面内的剪切应力 τ。

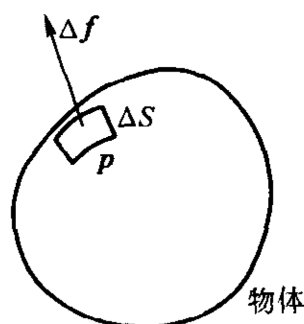


图 2.2-7　应力矢量的定义

4.2 柯西应力张量

物体受外力作用产生变形，其内部各部分之间将产生内力。用假想的截面将连续体截开来分析，则内力也是一种面力。考虑图 2.2-8 中四面体微元的平衡，其中 dS 为一面元，其单位外法线矢量为 n，面元 dS 上所受面力为 $\mathrm{d}f$，其分量为 $\mathrm{d}f_i$，四面体其它各面均平行于笛卡儿坐标系的各坐标平面。面元 dS 上的应力矢量 p 的各分量为

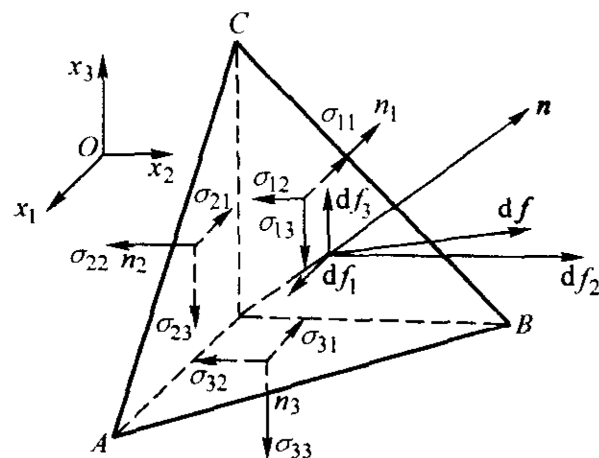


图 2.2-8　四面体微元的平衡

$$p_i=\lim_{\Delta S\to 0}\frac{\Delta f_i}{\Delta S}=\frac{\mathrm{d}f_i}{\mathrm{d}S}$$

由四面体的静力平衡和力矩平衡条件可得

$$\mathrm{d}f_i=\sigma_{ji}n_j\mathrm{d}S,\quad \sigma_{ij}=\sigma_{ji}$$

或写为

$$\mathrm{d}f=n\cdot\sigma\mathrm{d}S,\quad \sigma=\sigma^{\mathrm{T}}$$

于是得

$$p_i=\sigma_{ji}n_j,\ \text{或}\quad p=n\cdot\sigma \tag{2.2-33}$$

σ 为二阶对称张量，称为柯西应力张量。在笛卡儿坐标系中，柯西应力张量的分量就是物理应力分量。当 ΔS 为外表面上的面元时，式（2.2-33）即为该面元的应力边界条件。

4.3 其他应力张量

在有限变形情况下，要考虑参考构形和现时构形这两个构形，建立这两个构形上应力张量的表征和对应关系。因此，引入了如下两种“虚拟的”应力张量。

4.3.1 拉格朗日应力张量

图 2.2-7 中，ΔS 为现时构形中的一个面元，其单位法向量为 n，$\mathrm{d}f$ 为作用于 dS 上的面力。如果设作用于参考构形中与 dS 对应的面元 $\mathrm{d}S_0$ 上的面力仍为 $\mathrm{d}f$，则 $\mathrm{d}S_0$ 上的应力矢量为

$$p^{(\mathrm{T})}=\frac{\mathrm{d}f}{\mathrm{d}S_0} \tag{2.2-34}$$

仿照定义柯西应力张量的方法，定义拉格朗日应力张量 T，使之满足

$$\mathrm{d}f=N\cdot T\mathrm{d}S_0 \tag{2.2-35}$$

其中，N 为 $\mathrm{d}S_0$ 的单位法向量，将式（2.2-35）代入式（2.2-34）得

$$p^{(\mathrm{T})}=N\cdot T$$

拉格朗日应力张量又称为第一类 Piola-Kirchhoff 应力张量。

4.3.2 克希荷夫应力张量

在变形过程中，任一线元的长度和方向发生如下变化

$$\mathrm{d}x=F\mathrm{d}X,\quad \mathrm{d}X=F^{-1}\mathrm{d}x$$

如果将 $\mathrm{d}f$ 先进行与线元 $\mathrm{d}x$ 同样的变换

$$\mathrm{d}\tilde{f}=F^{-1}\mathrm{d}f$$

然后再作用于 $\mathrm{d}S_0$ 上，则在参考构形中定义出克希荷夫应力张量，又称为第二类 Piola-Kirchhoff 应力张量 S。

$$p^{(\mathrm{S})}=\frac{\mathrm{d}\tilde{f}}{\mathrm{d}S_0}=F^{-1}\cdot\frac{\mathrm{d}f}{\mathrm{d}S_0}=N\cdot S$$

考虑到面元变换的南生公式，即式（2.2-18），可得出柯西应力 σ，拉格朗日应力 T 和克希荷夫应力 S 之间的关系

$$T=JF^{-1}\cdot\sigma \tag{2.2-36}$$

$$S=T\cdot F^{-\mathrm{T}}=JF^{-1}\cdot\sigma\cdot F^{-\mathrm{T}} \tag{2.2-37}$$

σ 是对称的。因此，由式（2.2-36）知 T 是不对称的，而由式（2.2-37）知 S 是对称的。在有限变形分析的本构方

程中，多采用克希荷夫应力张量 S，因为其对称性会带来许多方便。

4.4 应力速率

4.4.1 张量的客观性

连续介质的物理性质是独立于观察者的，即与观察者的静止或运动无关。物理规律也应是如此。如果一种性质独立于观察者的运动，就说它具有客观性，或坐标系无关性。

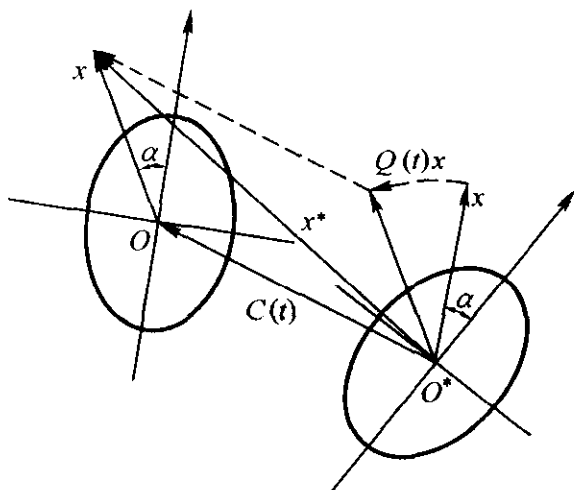


图 2.2-9　两个时空系之间的变换

考虑图 2.2-9 所示的两个作相对运动的时空系。设在 $t=0$时，它们是重合的，在 $t=t$ 时，这两个时空系之间的欧氏变换为

$$x^*=c(t)+Q(t)\cdot x$$
$$t^*=t-a$$

其中，Q 是这两个时空系之间的相对转动，为正交张量；c（t）是两时空系原点的相对平动。

如果标量 f，矢量 v 或二阶张量 S 在上述欧氏变换下保持不变，则称它们是客观的。这时，它们应满足下列关系式

$$\left.\begin{aligned}f^*&=f\\ v^*&=Q\cdot v\\ S^*&=Q\cdot S\cdot Q^{\mathrm{T}}\end{aligned}\right\}$$

下面讨论几个常用的二阶张量的客观性。先看变形梯度 F。在 $t=0$ 时，两时空系重合，因此 $\mathrm{d}X^*=\mathrm{d}X$。

$$\mathrm{d}x=F\cdot\mathrm{d}X,\quad \mathrm{d}x^*=F^*\cdot\mathrm{d}X$$

另一方面

$$\mathrm{d}x^*=Q\cdot\mathrm{d}x=Q\cdot F\cdot\mathrm{d}X=F^*\cdot\mathrm{d}X$$

于是
$$(F^*-Q\cdot F)\cdot\mathrm{d}X=0$$

由于 $\mathrm{d}X$ 是任意的，所以有

$$F^*=Q\cdot F$$

可见 F 不是具有客观性的二阶张量。

对速度梯度 L、应变速度 d 和旋转速率 ω 的分析表明

$$\left.\begin{aligned}L^*&=Q\cdot L\cdot Q^{\mathrm{T}}+\dot{Q}\cdot Q^{\mathrm{T}}\\ d^*&=Q\cdot d\cdot Q^{\mathrm{T}}\\ \omega^*&=Q\cdot\omega\cdot Q^{\mathrm{T}}+\dot{Q}\cdot Q^{\mathrm{T}}\end{aligned}\right\}\tag{2.2-38}$$

可见，d 具有客观性，而 L 和 ω 不具有客观性。

4.4.2 柯西应力的久曼速率

当本构关系中包含与材料的应变历史有关的塑性响应时，需要进行增量计算，因此要求应力的物质导数。

由不同坐标系中张量各分量的变换公式（2.2-1）知，柯西应力的变换式为

$$\sigma^*=Q\cdot\sigma\cdot Q^{\mathrm{T}}$$

即柯西应力张量本身具有客观性。但是柯西应力张量的物质导数没有客观性，因为

$$\dot{\sigma}^*=Q\cdot\dot{\sigma}\cdot Q^{\mathrm{T}}+\dot{Q}\cdot\sigma\cdot Q^{\mathrm{T}}+Q\cdot\sigma\cdot\dot{Q}^{\mathrm{T}}$$

将式（2.2-38）中第3式代入上式消去 $\dot{Q}$ 得

$$\dot{\sigma}^*+\sigma^*\cdot\omega^*-\omega^*\cdot\sigma^*=Q\cdot(\dot{\sigma}+\sigma\cdot\omega-\omega\cdot\sigma)\cdot Q^{\mathrm{T}}$$

可见

$$\hat{\sigma}=\dot{\sigma}+\sigma\cdot\omega-\omega\cdot\sigma\tag{2.2-39}$$

是具有客观性的二阶张量，这种应力速率称为柯西应力的久曼速率。

这里，我们证明了应变速率 d 和柯西应力的久曼速率 $\hat{\sigma}$ 具有客观性，所以它们适用于有限变形问题的速率（或增量）型本构方程。

4.4.3 其他应力速率

将式（2.2-37）对时间微分，得到克希荷夫应力张量的时间变化率

$$\begin{aligned}\dot{S}&=\frac{\partial S}{\partial t}=\frac{\partial}{\partial t}(JF^{-1}\sigma F^{-\mathrm{T}})\\ &=JF^{-1}(\dot{\sigma}-L\cdot\sigma-\sigma\cdot L^{\mathrm{T}}+\mathrm{tr}L\sigma)F^{-\mathrm{T}}\end{aligned}\tag{2.2-40}$$

上式中括号中的部分

$$\check{\sigma}=\dot{\sigma}-L\cdot\sigma-\sigma\cdot L^{\mathrm{T}}+\sigma\mathrm{tr}L$$

称为屈斯笛尔（Truesdell）应力速率。由于在式（2.2-39）中加上具有客观性的 $\pm d\cdot\sigma$，$\sigma\cdot d$ 等，仍然具有客观性，所以 $\check{\sigma}$ 也具有客观性。

类似地可求得拉格朗日应力张量的时间变化率

$$\dot{T}=JF^{-1}\ (\dot{\sigma}-L\cdot\sigma+\sigma\mathrm{tr}L)\tag{2.2-41}$$

5 基本方程和原理

本篇中要用到的连续介质力学的基本方程和原理包括连续性方程、运动方程、能量方程和虚功原理等。

5.1 基本方程

5.1.1 质量守恒定律和连续性方程

物体在运动和变形过程中，其质量保持不变，这就是质量守恒定律。表达质量守恒定律的数学方程都可称为连续性方程。令 M 为物体的质量，则质量守恒定律可用物质导数表示为

$$\frac{\mathrm{D}M}{\mathrm{D}t}=0$$

用 V 和 V_0 分别表示连续体在时刻 t 和 t_0 所占有的体积，ρ 和 ρ_0 分别表示对应的质量密度，则有

$$M=\int_V\rho\mathrm{d}V=\int_{V_0}\rho_0\mathrm{d}V_0$$

将 $\mathrm{d}V=J\mathrm{d}V_0$ 代入上式可得

$$\int_{V_0}(\rho J-\rho_0)\mathrm{d}V_0=0$$

假定被积函数连续，由于 V_0 的任意性得拉格朗日形式的连续性方程

$$\rho_0=\rho J\quad 或\quad \rho\mathrm{d}V=\rho_0\mathrm{d}V_0\tag{2.2-42}$$

体积元 $\mathrm{d}V$ 的质量为 $\mathrm{d}M=\rho\mathrm{d}V$，考虑如下物质导数

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V\rho A\mathrm{d}V$$

式中 ρ 为质量密度，A 可为任何一种物理量，把对体积 V 的积分换成对质量 M 的积分有

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V\rho A\mathrm{d}V=\frac{\mathrm{D}}{\mathrm{D}t}\int_M A\mathrm{d}M$$

由质量守恒定律，M 和 $\mathrm{d}M$ 都不随时间 t 改变，于是

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_M A\mathrm{d}M=\int_M\frac{\mathrm{D}A}{\mathrm{D}t}\mathrm{d}M=\int_V\rho\frac{\mathrm{D}A}{\mathrm{D}t}\mathrm{d}V$$

即

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V\rho A\mathrm{d}V=\int_V\rho\frac{\mathrm{D}A}{\mathrm{D}t}\mathrm{d}V\tag{2.2-43}$$

欧拉变量的连续性方程为

$$\frac{\partial\rho}{\partial t}+\mathrm{div}\ (\rho v)\ =0$$

5.1.2 运动方程

牛顿第二定律表明：在惯性参考系内，物体系总的动量

随时间的变化率，等于作用在物体系的外力的总和。将其推广应用于连续体内部任意区域就得到运动方程。

从物体的现时构形中任取一体积为 V 的部分，其表面为 S，表面上一点的外法线为 $\boldsymbol{n}$，该点的应力矢量为 $\boldsymbol{p}$。设单位质量的体积力为 $\boldsymbol{b}$。将牛顿第二定律用于该区域得

$$\int_V \rho \ddot{u}_i \mathrm{d}V = \int_S p_i \mathrm{d}S + \int_V \rho b_i \mathrm{d}V \tag{2.2-44}$$

其中 $\ddot{u}_i$ 为加速度分量，考虑到式（2.2-33）及奥高公式（2.2-5），右边第一项可变换为

$$\int_S p_i \mathrm{d}S = \int_S \sigma_{ij} n_j \mathrm{d}S = \int_V \frac{\partial \sigma_{ji}}{\partial x_j} \mathrm{d}V$$

将上式代入式（2.2-44）得

$$\int_V \left(\rho \ddot{u}_i - \frac{\partial \sigma_{ji}}{\partial x_j} - \rho b_i \right) \mathrm{d}V = 0$$

假定被积函数连续，由 V 的任意性得到欧拉描述法中的运动方程

$$\frac{\partial \sigma_{ji}}{\partial x_j} + \rho b_i = \rho \ddot{u}_i \tag{2.2-45}$$

当 $\ddot{\boldsymbol{u}} = 0$ 时，得到平衡方程

$$\frac{\partial \sigma_{ji}}{\partial x_j} + \rho b_i = 0 \tag{2.2-46}$$

设在初始构形中与 $\mathrm{d}S$ 对应的面元为 $\mathrm{d}S_0$，其法线为 $\boldsymbol{N}$，$\mathrm{d}S_0$ 上作用的应力矢量为 $\boldsymbol{p}^0$。由拉格朗日应力的定义，有 $\boldsymbol{p}\mathrm{d}S = \boldsymbol{p}^0 \mathrm{d}S_0$，于是

$$\int_S p_i \mathrm{d}S = \int_{S_0} p_i^0 \mathrm{d}S_0 = \int_{S_0} T_{ji} N_j \mathrm{d}S_0 = \int_{V_0} \frac{\partial T_{ji}}{\partial X_j} \mathrm{d}V_0$$

其中 T_{ji} 为拉格朗日应力。将上式代入式（2.2-44），并考虑到连续性方程式（2.2-42）得

$$\int_V \left(\frac{\partial T_{ji}}{\partial X_j} + \rho_0 b_i - \rho_0 \ddot{u}_i \right) \mathrm{d}V_0 = 0$$

由 V_0 的任意性，得到拉格朗日形式的运动方程

$$\frac{\partial T_{ji}}{\partial X_j} + \rho_0 b_i = \rho_0 \ddot{u}_i \tag{2.2-47}$$

以 $\boldsymbol{v}$ 与运动方程两边作点积，可得动能变化的方程

$$\rho \frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{v^2}{2} \right) = \rho \boldsymbol{b} \cdot \boldsymbol{v} + \mathrm{div}\ (\boldsymbol{v} \cdot \boldsymbol{\sigma}) - \boldsymbol{\sigma} \cdot \boldsymbol{d} \tag{2.2-48}$$

5.1.3 其他坐标系下的应力平衡方程

对于轴对称和球对称问题，分别在圆柱坐标系和球坐标系进行分析更方便。下面给出忽略体力时圆柱坐标系和球坐标系中的应力平衡方程。

1）圆柱坐标中

$$\begin{aligned}
&\frac{\partial \sigma_\rho}{\partial \rho} + \frac{1}{\rho}\frac{\partial \tau_{\theta\rho}}{\partial \theta} + \frac{\partial \tau_{z\rho}}{\partial z} + \frac{\sigma_\rho - \sigma_\theta}{\rho} = 0 \\
&\frac{\partial \tau_{\rho\theta}}{\partial \rho} + \frac{1}{\rho}\frac{\partial \sigma_\theta}{\partial \theta} + \frac{\partial \tau_{z\theta}}{\partial z} + \frac{2\tau_{\rho\theta}}{\rho} = 0 \\
&\frac{\partial \tau_{\rho z}}{\partial \rho} + \frac{1}{\rho}\frac{\partial \tau_{\theta z}}{\partial \theta} + \frac{\partial \sigma_z}{\partial z} + \frac{\tau_{\rho z}}{\rho} = 0
\end{aligned} \tag{2.2-49}$$

轴对称状态时

$$\frac{\partial \sigma_\rho}{\partial \rho} + \frac{\partial \tau_{z\rho}}{\partial z} + \frac{\sigma_\rho - \sigma_\theta}{\rho} = 0$$

$$\frac{\partial \tau_{\rho z}}{\partial \rho} + \frac{\partial \sigma_z}{\partial z} + \frac{\tau_{\rho z}}{\rho} = 0$$

2）球坐标系中

$$\frac{\partial \sigma_\rho}{\partial \rho} + \frac{1}{\rho \sin\varphi}\frac{\partial \tau_{\theta\rho}}{\partial \theta} + \frac{1}{\rho}\frac{\partial \tau_{\varphi\rho}}{\partial \varphi} + \frac{\partial \tau_{\theta\rho}}{\partial \theta} +$$

$$\frac{1}{\rho}(2\sigma_\rho - \sigma_\theta - \sigma_\varphi + \tau_{\varphi\rho}\cot\varphi) = 0$$

$$\frac{\partial \tau_{\rho\theta}}{\partial \rho} + \frac{1}{\rho \sin\varphi}\frac{\partial \sigma_\theta}{\partial \theta} + \frac{1}{\rho}\frac{\partial \tau_{\varphi\theta}}{\partial \varphi} + \tag{2.2-50}$$

$$\frac{1}{\rho}(3\sigma_{\rho\theta} + 2\tau_{\varphi\theta}\cot\varphi) = 0$$

$$\frac{\partial \tau_{\rho\theta}}{\partial \rho} + \frac{1}{\rho \sin\varphi}\frac{\partial \tau_{\theta\varphi}}{\partial \theta} + \frac{1}{\rho}\frac{\partial \sigma_\varphi}{\partial \varphi} +$$

$$\frac{1}{\rho}[(\sigma_\varphi - \sigma_\theta)\cot\varphi + 3\tau_{\rho\theta}] = 0$$

5.1.4 能量守恒定律和能量方程

能量守恒定律是最基本的物理定律。在材料成形中发生机械能和热能间的转换，这种情况下的能量守恒定律也就是热力学第一定律。它可以表述为

$$U = W + Q \tag{2.2-51}$$

式中 U——单位时间内物体系增加的能量；

W——单位时间内外力作功的总和；

Q——单位时间内物体系增加的热量。

下面对连续体 V 应用能量守恒定律。令 e 为单位质量物体所具有的内能，物体 V 的总内能为 $\int_V \rho e \mathrm{d}V$，总动能为 $\int_V \frac{1}{2}\rho v^2 \mathrm{d}V$，总能量为$\int_V \rho\left(\frac{v^2}{2} + e\right)\mathrm{d}V$。于是

$$U = \frac{\mathrm{D}}{\mathrm{D}t}\int_V \rho\left(\frac{v^2}{2} + e\right)\mathrm{d}V \tag{2.2-52}$$

单位时间内外力作功的总和为

$$W = \int_V \rho \boldsymbol{b} \cdot \boldsymbol{v} \mathrm{d}V + \int_S \boldsymbol{p} \cdot \boldsymbol{v} \mathrm{d}S \tag{2.2-53}$$

上式右边第一、二项分别是单位时间内体积力和表面力所作的功。

当有热辐射、相变等情况时，物体会放出或吸收热量。令 h 为单位质量的物体在单位时间内放出的热量，则物体 V 在单位时间内由于热源放出的热量为 $\int_V \rho h \mathrm{d}V$。

令 $\boldsymbol{q}$ 为热流密度矢量，$\boldsymbol{n}$ 为面元的单位法矢量，则 $q_n = \boldsymbol{q} \cdot \boldsymbol{n}$ 为单位时间内从面元负 $\boldsymbol{n}$ 方流向正 $\boldsymbol{n}$ 方的热流密度（即单位面积上流过的热量）。于是，单位时间内由表面 S 流入物体 V 的热量为 $-\int_S \boldsymbol{q} \cdot \boldsymbol{n} \mathrm{d}S$。

单位时间内物体 V 增加的热量为

$$Q = \int_V \rho h \mathrm{d}V - \int_S \boldsymbol{q} \cdot \boldsymbol{n} \mathrm{d}S \tag{2.2-54}$$

将式（2.2-52）～式（2.2-54）代入式（2.2-51），得出如下能量方程

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V \rho\left(\frac{v^2}{2} + e\right)\mathrm{d}V = \int_V (\rho \boldsymbol{b} \cdot \boldsymbol{v} + \rho h)\mathrm{d}V + \int_S (\boldsymbol{p} \cdot \boldsymbol{v} - \boldsymbol{q} \cdot \boldsymbol{n})\mathrm{d}S$$

利用式（2.2-33）、式（2.2-43）和奥高公式将上式中的面积分转换成体积分得

$$\int_V \left[\rho \frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{v^2}{2} + e\right) - \rho \boldsymbol{b} \cdot \boldsymbol{v} - \rho h - \mathrm{div}(\boldsymbol{\sigma} \cdot \boldsymbol{v}) + \mathrm{div}\boldsymbol{q} \right] \mathrm{d}V = 0$$

假定被积函数连续，由于 V 的任意性，可得微分形式的能量方程

$$\rho \frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{v^2}{2} + e\right) = \rho \boldsymbol{b} \cdot \boldsymbol{v} + \rho h + \mathrm{div}(\boldsymbol{\sigma} \cdot \boldsymbol{v}) - \mathrm{div}\boldsymbol{q} \tag{2.2-55}$$

上式与动能变化方程式（2.2-48）相减，得到内能变化的微分方程

$$\rho \frac{\mathrm{D}e}{\mathrm{D}t} = \rho h - \mathrm{div}\boldsymbol{q} + \boldsymbol{\sigma} \cdot \boldsymbol{d} \tag{2.2-56}$$

5.2 边值问题与初值问题

金属塑性成形问题可以归结为连续介质力学的边值问题和初值问题，可以用场方程、边值条件和初值条件建立其力学模型。根据具体问题的特点或为了求解方便，可以将其作为静力问题或动力问题来分析。

静力边值问题不考虑加速度的影响，运动方程成为平衡方程。为了叙述的简洁，我们只列出较为常用的欧拉描述的基本方程和边值条件。这些方程是

连续性方程 $$\frac{D\rho}{Dt}+\rho\frac{\partial v_i}{\partial x_i}=0 \tag{2.2-57}$$

平衡方程 $$\frac{\partial\sigma_{ji}}{\partial x_j}+\rho b_i=0 \tag{2.2-58}$$

几何方程 $$d_{ij}=\frac{1}{2}\left(\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i}\right) \tag{2.2-59}$$

本构方程 $$\hat{\sigma}_{ij}=C^{ep}_{ijkl}d_{kl} \tag{2.2-60}$$

其中，b 为单位质量所受的体力。以上共有16个方程；16个未知数，可以求解。求解所需的定解条件由边界条件给出。

将物体的边界分为 S_u 和 S_p 两部分。在 S_u 上给定了位移 $\bar{u}$ 或速率 $\bar{v}$，而在 S_p 上给定了面力 $\bar{p}$。我们在变量上加上"—"表示给定值，则边界条件可写为

$$\sigma_{ji}n_j=\bar{p}_i\ （在\ S_p\ 上） \tag{2.2-61}$$

$$u_i=\bar{u}_i\quad 或\quad \dot{u}_i=\bar{\dot{u}}_i\ （在\ S_u\ 上） \tag{2.2-62}$$

其中式（2.2-61）称为应力边界条件，式（2.2-62）称为运动学边界条件。在金属成形问题分析中，制件与工具之间的接触表面上的接触边界条件是一种混合边界条件，其中有一些分量是给定的或已知的，另一些是未知的，同时还给定了这些变量应满足的约束条件。

动力边值－初值问题与静力边值问题的区别在于运动方程和初值条件。动力问题中加速度的影响不能忽略，同时动力系统存在阻尼力，它引起了能量耗散。动力问题的运动方程为

$$\frac{\partial\sigma_{ji}}{\partial x_j}+\rho b_i-\rho\ddot{u}_i-\nu\dot{u}_i=0 \tag{2.2-63}$$

其中 $-\nu\dot{u}_i$ 为单位体积的阻尼力。动力问题中，位移和速度是时间的函数，需要初值条件作为定解条件。在 $t=0$ 时，位移和速度初值条件为

$$u_i|_{t=0}=\bar{u}_{i0} \tag{2.2-64}$$

$$\dot{u}_i|_{t=0}=\bar{\dot{u}}_{i0} \tag{2.2-65}$$

初值条件是对整个物体给出的。

5.3 虚功原理与虚功率原理

设一物体在变形后的现时构形中，体积为 V，表面积为 S，$S=S_p+S_u$，在 S_p 上给定表面力 $\bar{p}$，在 S_u 上给定表面位移 $\bar{u}$ 或速度 $\bar{\dot{u}}$，单位质量上作用的体积力为 $\boldsymbol{b}$。我们首先考虑静力边值问题。问题的基本方程为5.2节中的式（2.2-57）～式（2.2-60），其中式（2.2-60）在如下讨论中并不涉及，即虚功原理对于任何本构方程都是成立的。问题的边值条件为式（2.2-61）和式（2.2-62）。

5.3.1 虚功原理

在现时构形中，设想给物体内各点以任意的虚位移 $\delta\boldsymbol{u}$，它仅是坐标 $\boldsymbol{x}$ 的函数，同时在 S_u 上满足 $\delta\boldsymbol{u}=\boldsymbol{0}$。外力在虚位移上所作的虚功为

$$\delta W^*=\int_V\rho b_i\delta u_i\,dV+\int_{S_p}\bar{p}_i\delta u_i\,dS \tag{2.2-66}$$

应用边界条件式（2.2-61）和奥高公式，并考虑到在 S_u 上 $\delta\boldsymbol{u}=\boldsymbol{0}$，则有

$$\begin{aligned}\int_{S_p}\bar{p}_i\delta u_i\,dS&=\int_S\sigma_{ij}n_j\delta u_i\,dS=\int_V\frac{\partial(\sigma_{ij}\delta u_i)}{\partial x_j}dV\\&=\int_V\frac{\partial\sigma_{ij}}{\partial x_j}\delta u_i\,dV+\int_V\sigma_{ij}\frac{\partial u_i}{\partial x_j}dV\end{aligned} \tag{2.2-67}$$

定义虚应变速率为 $\delta\varepsilon_{ij}=[\partial(\delta u_i)/\partial x_j+\partial(\delta u_j)/\partial x_i]/2$，将此式及式（2.2-67）代入式（2.2-66）并考虑到 $\delta\varepsilon_{ij}$ 的对称性及平衡方程得

$$\begin{aligned}\delta W^*&=\int_V\left(\rho b_i+\frac{\partial\sigma_{ij}}{\partial x_j}\right)\delta u_i\,dV+\int_V\sigma_{ij}\delta\varepsilon_{ij}\,dV\\&=\int_V\sigma_{ij}\delta\varepsilon_{ij}\,dV\end{aligned} \tag{2.2-68}$$

综合式（2.2-66）、式（2.2-68）有

$$\int_V\rho b_i\delta u_i\,dV+\int_{S_p}\bar{p}_i\delta u_i\,ds=\int_V\sigma_{ij}\delta\varepsilon_{ij}\,dV \tag{2.2-69}$$

上式表明外力的虚功等于因虚位移引起的虚应变能。这就是现时构形中的虚功原理。它是力学中的一个普遍原理。反过来，如果设虚功原理成立，则因为

$$\begin{aligned}\delta W^*&=\int_V\sigma_{ij}\delta\varepsilon_{ij}\,dV=\int_V\frac{\partial(\sigma_{ij}\delta u_i)}{\partial x_j}dV-\int_V\frac{\partial\sigma_{ij}}{\partial x_j}\delta u_i\,dV\\&=\int_S\sigma_{ij}n_j\delta u_i\,dS-\int_V\frac{\partial\sigma_{ij}}{\partial x_j}\delta u_i\,dV\end{aligned}$$

将上式代入式（2.2-69）有

$$\int_S\sigma_{ij}n_j\delta u_i\,dS-\int_V\frac{\partial\sigma_{ij}}{\partial x_j}\delta u_i\,dV=\int_V\rho b_i\delta u_i\,dV+\int_{S_p}\bar{p}_i\delta u_i\,dS$$

即

$$\int_V\left(\frac{\partial\sigma_{ij}}{\partial x_j}+\rho b_i\right)\delta u_i\,dV+\int_{S_p}(\bar{p}_i-\sigma_{ij}n_j)\delta u_i\,dS=0$$

由于 $\delta\boldsymbol{u}$、体积 V 和相应的 S_p 的任意性，得

$$\frac{\partial\sigma_{ij}}{\partial x_j}+\rho b_i=0,\quad \sigma_{ij}n_j=\bar{p}_i$$

这就是静力平衡方程和应力边界条件。另外，在 S_u 上 $\delta\boldsymbol{u}=0$ 的前提是 $\boldsymbol{u}=\bar{\boldsymbol{u}}$。这说明虚功原理与静力平衡方程和边界条件是等价的。

设载荷的大小和方向都不变，根据连续性方程和拉格朗日应力张量的定义，有

$$\int_V\rho\boldsymbol{b}\cdot\delta\boldsymbol{u}\,dV=\int_{V_0}\rho_0\boldsymbol{b}\cdot\delta\boldsymbol{u}\,dV_0$$

$$\begin{aligned}\int_S\bar{\boldsymbol{p}}\cdot\delta\boldsymbol{u}\,dS&=\int_{S_0}\bar{\boldsymbol{p}}^{(T)}\cdot\delta\boldsymbol{u}\,dS_0\\&=\int_{S_0}\boldsymbol{N}\cdot\boldsymbol{T}\cdot\delta\boldsymbol{u}\,dS_0\\&=\int_{V_0}\mathrm{Div}\boldsymbol{T}\cdot\delta\boldsymbol{u}dV_0+\int_{V_0}\boldsymbol{T}:\delta\boldsymbol{F}^{T}dV_0\end{aligned}$$

式中，Div表示初始构形中的散度算子。将以上二式代入式（2.2-66），并运用拉格朗日形式的运动方程（令 $\ddot{\boldsymbol{u}}=\boldsymbol{0}$）可得拉格朗日形式的虚功原理

$$\int_{V_0}\rho_0\boldsymbol{b}\cdot\delta\boldsymbol{u}\,dV_0+\int_{S_0}\bar{\boldsymbol{p}}^{(T)}\cdot\delta\boldsymbol{u}\,dS_0=\int_{V_0}\boldsymbol{T}:\delta\boldsymbol{F}^{T}dV_0 \tag{2.2-70}$$

采用克希荷夫应力张量时，虚功原理可表述成

$$dW^*=\int_{V_0}\boldsymbol{S}:\delta\boldsymbol{E}\,dV_0$$

5.3.2 虚功率原理

把虚功原理中的虚位移 δu 换为虚速度 $\delta\dot{u}$，$\delta\dot{u}$ 可任意选取，但在 S_u 上必须满足 $\delta\dot{u}=0$。重复上述讨论，可得到如下虚功率原理

$$\delta\dot{W}^* = \int_V \rho b_i \delta\dot{u}_i \mathrm{d}V + \int_{S_p} \bar{p}_i \delta\dot{u}_i \mathrm{d}S$$
$$= \int_V \sigma_{ij} \frac{\partial(\delta\dot{u}_i)}{\partial x_j} \mathrm{d}V = \int_V \sigma_{ij}\delta L_{ij} \mathrm{d}V$$
$$= \int_V \sigma_{ij}\delta d_{ij} \mathrm{d}V \qquad (2.2\text{-}71)$$

其中$\delta L_{ij} = \partial(\delta\dot{u}_i)/\partial x_j$，$\delta d_{ij} = [\partial(\delta\dot{u}_i)/\partial x_j + \partial(\delta\dot{u}_j)/\partial x_i]/2$。

由 $x_i = X_i + u_i$ 可求得格林应变张量的物质导数为

$$\dot{E}_{ij} = \frac{1}{2}\left(\frac{\partial \dot{u}_k}{\partial X_i}\frac{\partial x_k}{\partial X_j} + \frac{\partial x_k}{\partial X_i}\frac{\partial \dot{u}_k}{\partial X_j}\right) = \frac{\partial x_k}{\partial X_i}\frac{\partial x_l}{\partial X_j} d_{kl}$$

克希荷夫应力张量和柯西应力张量的关系为

$$S_{ij} = J\frac{\partial X_i}{\partial x_k}\frac{\partial X_j}{\partial x_l}\sigma_{kl}$$

对于保守载荷的情况

$$\rho^0 \mathrm{d}V_0 = \rho \mathrm{d}V,\ p_i^0 \mathrm{d}S_0 = p_i \mathrm{d}S$$

将以上各式代入式 (2.2-71) 得到初始构形中的虚功率原理

$$\delta\dot{W}^* = \int_{V_0} \rho_0 b_i \delta\dot{u}_i \mathrm{d}V_0 + \int_{S_{0p}} \bar{p}_i^0 \delta\dot{u}_i \mathrm{d}S_0$$
$$= \int_{V_0} \boldsymbol{S} : \delta\dot{\boldsymbol{E}} \mathrm{d}V_0 \qquad (2.2\text{-}72)$$

其中 $\delta\dot{E}_{ij} = \frac{1}{2}\left[\frac{\partial(\delta\dot{u}_i)}{\partial X_j} + \frac{\partial(\delta\dot{u}_j)}{\partial X_i} + \frac{\partial(\delta\dot{u}_k)}{\partial X_i}\frac{\partial(\delta\dot{u}_k)}{\partial X_j}\right]$。

5.3.3 动力问题的虚功率原理

在动力问题中，可利用达朗伯原理，将惯性力 $-\rho\ddot{\boldsymbol{u}}$ 作为外力，阻尼力也是外力。在现时构形中外力的虚功率为

$$\delta\dot{W}^* = \int_V (\rho b_i - \rho\ddot{u}_i - \nu\dot{u}_i)\delta\dot{u}_i \mathrm{d}V + \int_{S_p} p_i \delta\dot{u}_i \mathrm{d}S \qquad (2.2\text{-}73)$$

重复 5.3.2 节的推导，应用奥高公式，可将上式中最后一项写为

$$\int_{S_p} \bar{p}_i \delta\dot{u}_i \mathrm{d}S = \int_V \frac{\partial\sigma_{ij}}{\partial x_j}\delta\dot{u}_i \mathrm{d}V + \int_V \sigma_{ij}\delta d_{ij} \mathrm{d}V \qquad (2.2\text{-}74)$$

将式 (2.2-74) 代入式 (2.2-73)，并应用运动方程式 (2.2-45)，得

$$\delta\dot{W}^* = \int_V \sigma_{ij}\delta d_{ij} \mathrm{d}V \qquad (2.2\text{-}75)$$

由式 (2.2-73) 和式 (2.2-75) 得

$$\int_V \rho\ddot{u}_i \delta\dot{u}_i \mathrm{d}V + \int_V \nu\dot{u}_i \delta\dot{u} \mathrm{d}V$$
$$= \int_V \rho b_i \delta\dot{u}_i \mathrm{d}V + \int_{S_p} \bar{p}_i \delta\dot{u}_i \mathrm{d}S - \int_V \sigma_{ij}\delta d_{ij} \mathrm{d}V \qquad (2.2\text{-}76)$$

这就是动力问题的虚功率原理。

编写：董湘怀（上海交通大学）

第3章　液态金属的凝固原理

1　液态金属的结构及特性

液态金属材料成形是现代工业零件毛坯的重要来源。液态金属成形过程描述的是液态金属向固态转变时的凝固过程。科学研究和生产实践使人们逐渐认识到，液态母体金属并不是杂乱无章地存在着，而有一定的结构特征，这个结构特征与凝固过程中金属体积的变化、溶质分配、热量传递，气体和夹杂物的出现有密切关系，并最终影响零件毛坯的相组织的形成、性能变化和成形效率。

1.1　液态金属的结构

1.1.1　金属原子的热运动及熔化

金属晶体的构成主要靠带正电荷的离子点阵和在正离子间运动的公有电子云之间的库仑力维系，这些离子（或原子）之间即有引力又有斥力，其相互作用之关系如图2.3-1所示。设 A、B 为两个金属离子，在一定温度下，A、B 两离子在平衡位置不停振动。当他们靠近时（距离小于 R_0），斥力增大，使两离子不易靠近。而当距离增大大于 R_0 时，斥力减少，引力迅速增加，R_0 距离恰好是斥力和引力的平衡点，离子在此处势能 W 最小。由图2.3-1还可看出，金属离子间的这种势能与离子间的距离的关系曲线极不对称，向右（距离 R 增加）势能 W 曲线是水平渐近，而向左是几乎垂直渐近，也就是说，当温度升高时，随着势能 W 的增加，离子间距离 R 会急剧增加，金属产生体积膨胀。除了离子间距的加大会造成金属膨胀外，空穴的产生也是造成金属膨胀的重要原因。在实际晶体中，离子相互作用后，将产生一定大小的能垒（$\Delta\mu$），如图2.3-2所示，由于能垒的存在，离子多在一定的节点以振动的形式活动。随温度升高，有少数离子高于能垒跑到金属的表面或原子间的间隙中，从而留下自由点阵称为空穴。温度越高，离子的能量越大，空穴数目也越多，造成金属体积的膨胀，据测定在熔点附近，空穴的数目可以达到原子总数的1%。

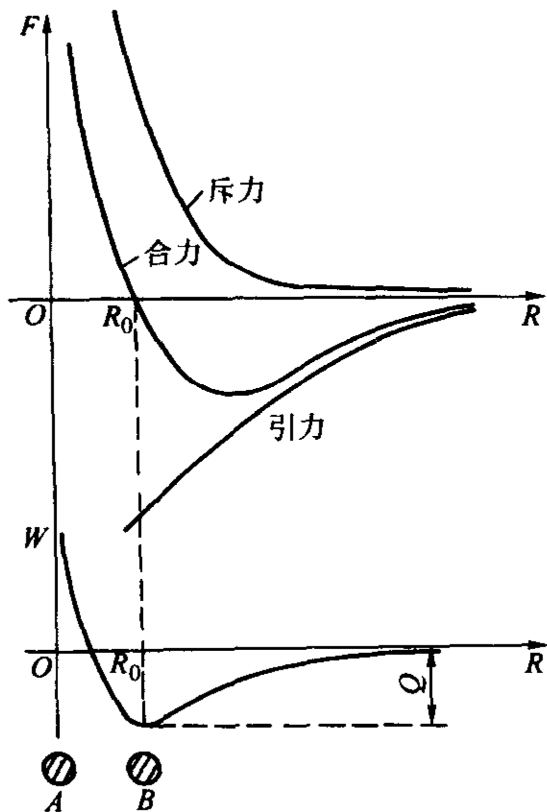


图2.3-1　离子间的作用力 F、势能 W 与距离 R 的关系

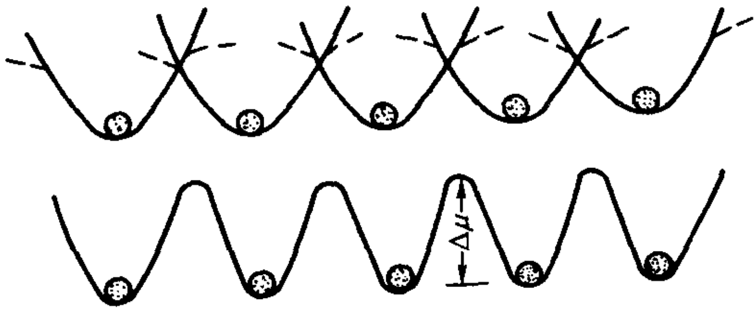


图2.3-2　离子间能垒示意图

金属的温度升高，离子间距增大，离子间的引力减弱。外界供给的能量使离子有能力越过能垒，离子的引力降低使结合健失效，规则的晶体排列突然被破坏，金属的性质发生显著变化，体积膨胀3%～5%，电阻、黏性等也有一突变，表明金属由固体进入液体状态。在金属熔点附近，金属吸收大量熔化潜热，但温度却不升高。

金属的熔化并不是原子间结合的全部消失，与气态不同，液体金属内原子的分布仍具有一定规律性。这可由金属熔化后体积仅增加3%～5%（即原子间的平均间距仅增加1%～1.5%，说明原子间仍有一定结合能）、以及熔化潜热一般只有汽化升华热的3%～7%所间接证实，即熔化时原子间的结合能仅减少了百分之几。

1.1.2　液态金属的结构特点

液态金属的特性介于气态和固态之间，而更接近于固体，特别在熔点附近表现如此。以面心立方结构的纯铝为例，铝的熔化潜热和汽化潜热对比（见表2.3-1），后者比前者大28倍，即表明将固态原子完全变为气态所需能量要比将同类原子从固态变为液态所需的能量多28倍。大多数金属熔化后，比容仅增加3%～5%。

表2.3-1　常用金属熔化时体积变化和热物理性能比较

金属	晶体结构	熔点/K	熔化热 L_m/kJ·mol^{-1}	沸点/K	汽化热 L_b/kJ·mol^{-1}	$\frac{L_b}{L_m}$	熔化时的体积变化率/%
Al	面心立方	933	10.47	2 653	291.42	27.8	+6.0
Au	面心立方	1 336	12.81	3 223	339.15	26.7	+5.1
Cu	面心立方	1 356	13.02	2 848	304.81	23.4	+4.15
Fe	面心/体心	1 812	15.20	3 343	340.40	22.4	+3.0
Zn	密排六方	693	7.20	1 180	115.14	16.0	
Cd	密排六方	594	6.41	1 038	99.65	15.6	+4.0
Mg	密排六方	923	8.71	1 376	133.98	15.4	+4.1
Sn	四方	505					+2.3
Sb	三方菱形	904					-0.95
Ga	面心斜方	303					-3.2
Bi	三方	544					-3.25
Ge	金刚石型	1 210					-5.0
Zr	密排六方	693					+4.2

通过对同种金属液固两态X射线衍射数据的比较发现（见表2.3-2），液态金属原子间距与固态相比略有增加（约1%～1.5%）。由于原子间距和空穴的增加，使原子的配位数也有增加。另由金属固液转变时熵值的变化也可以看出（见表2.3-3）在金属原子由固态转变为液态时，其熵值的增加相对于熔点前的熵值并不大，意味着金属由固态转变为液态时，原子排列的结构紊乱程度并没有明显增加。

表 2.3-2 X 射线衍射所得几种金属结构参数

金属	温度/K	液态		固态	
		原子间距离/nm	配位数	原子间距离/nm	配位数
Li	437	0.324	10	0.303	8
Na	373	0.383	8	0.372	8
Al	973	0.296	10 ~ 11	0.286	12
K	343	0.464	8	0.450	8
Zn	733	0.294	11	0.265，0.294	6 + 6
Cd	623	0.306	8	0.297，0.330	6 + 6
Sn	533	0.320	11	0.302，0.315	4 + 2
Au	1 373	0.286	11	0.288	12
Bi	613	0.332	7 ~ 8	0.309，0.346	3 + 3

表 2.3-3 一些金属的熵值变化

金属	从 25℃到熔点熵值变化 ΔS	熔点时的熵值变化 ΔS_m	$\Delta S_m/\Delta S$
Cd	4.53	2.46	0.54
Zn	5.45	2.55	0.47
Al	7.51	2.75	0.37
Mg	7.54	2.32	0.31
Cu	9.79	2.30	0.24
Au	9.78	2.21	0.23
Fe	15.50	2.00	0.13

上述研究结果说明了液态金属的结构特点：①原子间距增加不大，仍保持着较强结合能和一定固态特性；这种特性仅保持在较小范围，这个范围约为十几致几百个原子组成的集团，或称为“近程有序排列”；②液态金属内存在许多“原子集团”，集团尺寸大小与温度有关，而且内部有很强的热运动（原子扩散和空穴移动），所有原子集团处于变化之中，可变大，变小或消失。

对实际金属而言，由于存在杂质原子或第二种原子，使液态金属原子集团的结构复杂得多，即在不同原子集团间有浓度起伏和能量起伏时存在；某些化学亲和力较强的元素原子间还可能形成化合物。这些化合物可以以气、固或液态形式存在于液态金属中（多数成为非金属夹杂物）。所以实际液态金属在微观上是由成分和结构都不同的游动原子集团、空穴和许多固态、气态或液态化合物组成的混合液体。近年来，对液态铸铁金属的研究发现，存在许多“石墨未熔解的碳显微集团”，”石墨熔解了的聚合碳质集团”以及“悬浮硅酸盐化合物”等。

但液态金属和固态金属在某些方面的差异仍较显著，如液态金属转变为结晶体时，流动性有突变；金属熔化时体积膨胀 3% ~ 5%，这相当于固态金属从热力学零度到熔点温度时的总体积膨胀量；还有液态金属扩散系数 D_L 通常为 10^{-5} cm²/s 数量级，而固态合金溶质扩散系数 D_L 仅为 10^{-8} cm²/s数量级，两者相差近 1 000 倍。

1.1.3 液态金属结构的两个理论模型

（1）晶格理论模型

该理论认为，在熔点附近的液体中具有和原固相晶体相似的显微晶体，它们是大小不同、存在时间短、时聚时散的原子集团。这些原子集团中的原子呈规则排列，其结构与固态金属近似，为“近程有序”结构。在这些原子集团之间存在一定的“自由空间”或大量空穴。而大小不同的近程有序排列的原子集团的此起彼伏，构成了液态金属的动态图像。这个“原子集团”就是晶胚，在具备一定条件并大于一定尺寸后，这个晶胚就会成为晶核。

（2）密集球堆理论

该理论首先由 J.D.Beal 提出，他认为液体是原子紊乱的密集球堆，即液体是均质的、互相黏着的、本质上不规则的原子集合体。它既不包含晶体区，也没有能收容其他原子的空隙。这一理论模型和对单原子液体所进行的实验测定得出的原子径向密度相吻合。对这个球堆模型的研究发现，在紊乱紧密排列的球堆中，存在被称为“伪晶核”的高致密区，也就是“晶胚”。密集球堆模型描绘了液体的瞬时静止结构，而真实液体中，由于原子不断运动，这种结构也是时聚时散不断变化。可以这样来区别液体与固体，即液体的结构为分子（原子）的随机密堆，而固体是有序的规则密堆。

1.2 液态金属的主要特性

1.2.1 液态金属的基本物理性质

（1）密度

金属液的密度对于黏度、表面张力的测量，以及对非金属夹杂物、熔渣的上浮分离都有密切关系。同时，密度还是研究熔融状态金属结构、原子间相互作用力大小的重要依据。表 2.3-4 为部分熔融金属的密度值。合金元素的含量对密度有重要影响，但测定其值较为困难。目前只能通过一些实验分析出其趋势，如在液态铁合金中加入 Al、C、Si、Mn、P、S 时，其密度减小，加入 Cu、Co、Ni 时其密度增大。

表 2.3-4 部分纯金属液的密度

金属	Ag	Al	Au	Cu	Fe	Hg	Mg	Pb	Sn	Zn
测量温度/℃	960.5	900	1 100	1 080	1 550	200	700	400	409	419
密度/g·cm⁻³	9.3	2.315	17.42	7.93	7.0	13.115	1.536	10.51	6.834	4.92

（2）熔点

金属及合金的熔点是其内部质点相互作用的结果，即原子结合键牢固程度的外在体现。由于键的牢固程度与原子体积大小有关，所以金属的熔点也与原子体积大小有关。在元素同期表中，随着原子序数的增加，元素的原子半径和熔点均呈周期性变化。一般原子半径小的元素都有较高的熔点。

合金元素的加入减弱原子间的作用力而使熔点降低。对于除共晶成份合金之外的大部分合金，其熔点不是一个确定值，而是存在一个固 - 液两相共存的温度区间，其区间大小与合金种类及化学成份有关。两相的存在对金属的宏观和微观组织、流动性、热裂缺陷等铸造性能有重要影响。

（3）电阻率

液态金属的电阻对熔体的电磁输送、电磁搅拌以及感应加热等工艺是一个重要参数。同时，它还是研究熔融金属的电子结构和电子状态的重要物理参数。

金属熔化后，其电阻一般将会增加，导电性降低。测定结果表明，金属在液态时的自由电子数与固相比并没有多大差别，因此，金属熔化后电阻增加的原因似乎与液态金属中原子热运动振幅的增大及与结构不规则性程度的增加有关。

测量熔体的电阻率比较困难，且各测定值之间的差别也较大。液态金属电阻率的变化与固态时的晶体结构有关，表 2.3-5 表示熔点时，固态电阻率 ρ_s 与液态电阻率 ρ_L 之间比值与结晶类型的关系。

表 2.3-5 ρ_s/ρ_L 与结晶类型之间的关系

ρ_s/ρ_L	结晶类型	金属种类
1/2	稠密填充（面心立方） 密排立方	Ag、Cd、Cu、Pb、Sn、Ti、Zn
3/2	体心立方	Na、Li、K、Cs
2	疏松充填（三斜方）	Bi、Sb、Ga

(4) 比热容

一般来说，液态金属的比热容随温度升高而减小。各种金属以及化合物在液态时比热容彼此相差不大，因此，即使将液态金属的比热容看作常量，其计标结果也不会有大的差别。部分液态金属及化合物的比热容见表2.3-6。

表2.3-6　部分液态金属、化合物比热容

（熔点时的值）　$J\cdot K^{-1}\cdot mol^{-1}$

Ag	Al	Cd	Mg	Na	Sn	1/2NaCl	1/3CaF_2
34.3	29.3	30.1	33.9	31.4	30.5	33.2	31.0

(5) 导热性

液态金属的导热性对金属的散热能力影响很大。导热性好的金属，冷却速度快，其断面的温度梯度较小，在铸件内部引起的热应力也很小。当温度升高时，密度变小，原子间距增加，原子振动波的传递受到影响，导热性有所降低。导热性大小一般用热导率 λ 表示，部分液态金属的热导率见表2.3-7。

表2.3-7　部分液态金属的热导率

金属	Al	Cd	Cs	Hg	K	Li	Na	Pb	Sb	Ti	Zn
T/℃	790	380	28.5	220	400	200	400	400	700	350	600
$\lambda/W\cdot cm^{-1}\cdot K^{-1}$	121.4	43.9	18.4	12.6	39.8	37.6	71.2	15.9	20.9	28.8	56.9

(6) 扩散速度

固态金属中的扩散机理可以用晶体中的原子与空位位置交换来解释。固态晶体中的空位数约为 $10^{-3}\sim10^{-6}$。液态时空位数大大增加，大约10个原子中就有一个空位，所以扩散速度大幅度增加，液态晶体中的扩散速度比固态约大 $10^2\sim10^5$ 倍。例如在1 500℃的铁液中，Mn的扩散系数为 $6\times10^{-5}\ cm^2/s$，而在固态奥氏体中Mn的扩散系数仅为 $4\times10^{-9}\ cm^2/s$。

1.2.2　液态金属的黏度和流动性

当液体处于层流运动状态，各液层间存在防碍液体流动的内摩擦阻力，称为黏度。可以认为金属的流动性就是黏度的倒数。液态金属的黏性对金属在铸型中的流动性能，即充填能力，以及气体、非金属夹杂物、熔渣等物体的上浮以及补缩能力都有显著影响。

两层液体间的相互作用力（摩擦阻力），可由牛顿液体内摩擦定律确定，即各液层作相对移动时，产生的摩擦阻力（F）与各层接触面积（S），以及各层相对移动速度 dv/dx 成正比，并与液体本身性能有关，其关系可写为

$$F=\eta S\frac{dv}{dx}\quad 或\quad \eta=\frac{F}{S\frac{dv}{dx}}$$

其中，η 为动力黏度系数，或称动力黏度。将动力黏度除以液体密度 ρ，即得运动黏度，以 ν 表示

$$\nu=\frac{\eta}{\rho}\tag{2.3-1}$$

运动黏度适用于外力作用下的水力学运动，此时由于外力的作用，液体比重对流体流动的影响可以排除。金属液体的流动和水相差不多，因此在进行铸件浇注系统的计算时，完全可按水力学原理考虑。但是在外力作用非常小的情况，液体金属的动力黏度 η 将起主要作用，如夹杂物的上浮过程和凝固过程中的补缩等均与动力黏度系数 η 有关。表2.3-8为铸铁和铸钢的动力黏度和运动黏度。

金属的黏度与温度和成份的变化有关，温度升高黏度降低；液态金属中的固态杂质数量增多，黏度增加；合金元素的变化，也影响黏度变化，如含碳量增加，黏度降低。一般而言，共晶点附近的合金黏度最低。

表2.3-8　铸铁及铸钢的动力黏度和运动黏度

合　金	温度/℃	动力黏度 η		运动黏度 ν	
		$N\cdot s/m^2$	$gf\cdot s/cm^2$	$m^2/s\times10^{-6}$	cm^2/s
3.0%C灰铸铁	1 300	0.003 84	0.038 4	0.55	0.005 5
3.0%C灰铸铁	1 350	0.003 69	0.036 9	0.53	0.005 3
3.0%C灰铸铁	1 400	0.003 50	0.035 0	0.51	0.005 1
3.3%C灰铸铁	1 300	0.003 76	0.037 6	0.54	0.005 4
3.3%C灰铸铁	1 350	0.003 61	0.036 1	0.52	0.005 2
3.3%C灰铸铁	1 400	0.003 45	0.034 5	0.51	0.005 1
3.9%C灰铸铁	1 300	0.003 51	0.035 1	0.51	0.005 1
3.9%C灰铸铁	1 350	0.003 39	0.033 9	0.50	0.005 0
3.9%C灰铸铁	1 400	0.003 29	0.032 9	0.49	0.004 9
铸　　钢	1 500	0.002 81	0.028 1	0.40	0.004 0
铸　　钢	1 700	0.001 90	0.019 0	0.27	0.002 7

1.2.3　液态金属的表面张力

液态金属内存在着许多相与相之间的界面，如液态金属与大气、熔剂、型壁以及金属内部的气体、夹杂、晶体等接触所形成的界面，在这些界面上所发生的一些现象特别是表面张力的变化，对合金的精炼和孕育、铸型的充填、凝固结晶、气体的吸附和析出、夹杂物的形态以及铸件的补缩等有重要影响。

液态金属的表面张力主要是指液体表面之质点，由于受到周围环境质点（如气相、液相等）对它作用力的不平衡所产生的垂直于液面，且指向液体内部之力。对液气两相接触所产生的这种表面张力也可称为界面张力。表面张力的大小既与液体本身性质有关，又与它接触的相的性质有关。表2.3-9为金属在熔点时的表面张力。

表2.3-9　液态金属在其熔点时的表面张力和密度

金属	熔点/℃	表面张力/$mN\cdot m^{-1}$	测定方法	测定气氛	液态密度/$g\cdot cm^{-3}$
Li	180	398	M.B.P	氢	0.513
Be	1 277	1 390	S.D.	真空	1.69
Na	98	191	M.B.P.	氩	0.926
Mg	650	559	M.B.P.	氮	1.59
Al	660	914	M.B.P.	氮	2.38
Si	1 410	865	S.D.	氢	2.52
K	64	115	M.B.P.	氩	0.823
Ca	838	361	M.B.P.	氩	1.37
Ti	1 668	1 650	D.W.	真空	4.11
V	1 900	1 950	D.W.	真空	5.75
Cr	1 875	1 700	D.W.	氩	(6.46)
Mn	1 245	1 090	S.D.	氦	5.73
Fe	1 537	1 872	S.D.	氦	7.01
Co	1 495	1 873	M.B.P.	氢	7.67
Ni	1 453	1 778	M.B.P.	氢	7.77
Cu	1 083	1 360	M.B.P.	氦	7.99
Zn	420	782	M.D.P.	真空	6.57

续表 2.3-9

金属	熔点 /℃	表面张力 /mN·m⁻¹	测定方法	测定气氛	液态密度 /g·cm⁻³
Ga	30	718	S.D.	真空	6.09
Ge	937	621	M.B.P.	氩	5.49
Se	217	106	D.W.	空气	4.01
Rb	39	85	M.B.P.	氩	1.47
Sr	768	303	M.B.P.	氩	2.38
Zr	1 852	1 480	D.W.	真空	6.06
Nb	2 468	1 900	D.W.	真空	(7.83)
Mo	2 610	2 250	D.W.	真空	(9.35)
Ru	2 250	2 250	D.W.	真空	(10.9)
Rh	1 966	2 000	D.W.	真空	(11.1)
Pd	1 552	1 500	D.W.	真空	10.52
Ag	961	903	M.B.P.	氢	9.33
Cd	321	570	M.B.P.	氩	8.02
In	156	556	M.D.P.	真空	7.03
Sn	232	544	M.B.P.	氩	7.00
Sb	631	367	M.B.P.	氢	6.48
Te	450	180	M.D.P.	氮	5.80
Cs	29	70	M.B.P.	氩	1.838

对大多数金属和合金，随温度增加，表面张力降低，其关系如图 2.3-3 所示。

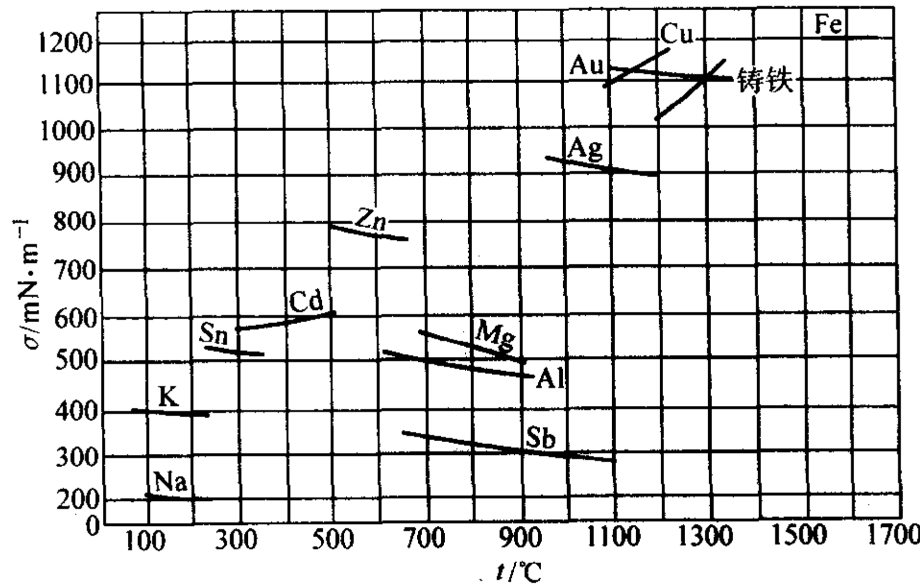


图 2.3-3　液态金属的表面张力与温度的关系

不同的溶质（包括杂质）对金属的表面张力有重要影响，这些溶质可以分为两类，一类是能使液态金属表面张力降低的溶质，称为液体金属的表面活性物质，有正吸附作用。另一类是使液态金属表面张力增加的溶质称为非表面活性物质，有负吸收作用。图 2.3-4、图 2.3-5、图 2.3-6 分别为不同合金元素加入 Al、Mg 合金及铸铁熔液所引起的表面张力变化。

1.2.4　液态金属的遗传性

所谓液态金属的遗传性，即是某种金属经重熔处理，凝固成形后的组织和性能不同程度保留有原始固态金属特征的性质。早在 20 世纪 20 年代，法国学者 Levi 就首次提出了金属材料的遗传现象，此后几十年时间不少的学者对金属的遗传现象和性质进行了大量研究。70 年代，法国冶金学家 Margerie 作了大量系统研究并明确提出“液态金属遗传性”概念，引起冶金学界的普遍关注，现在，金属材料的遗传学研究已成为现代物理冶金学研究的重要内容，金属材料遗传性质的控制，也成为液态金属成形与组织性能控制的重要手段。

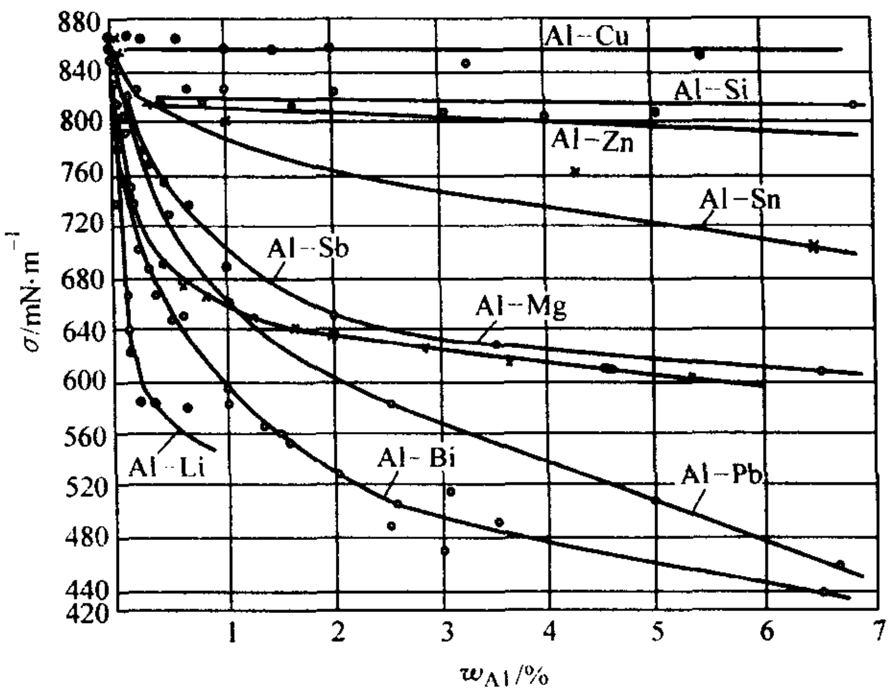


图 2.3-4　Al 中加入第二组之后表面张力的变化

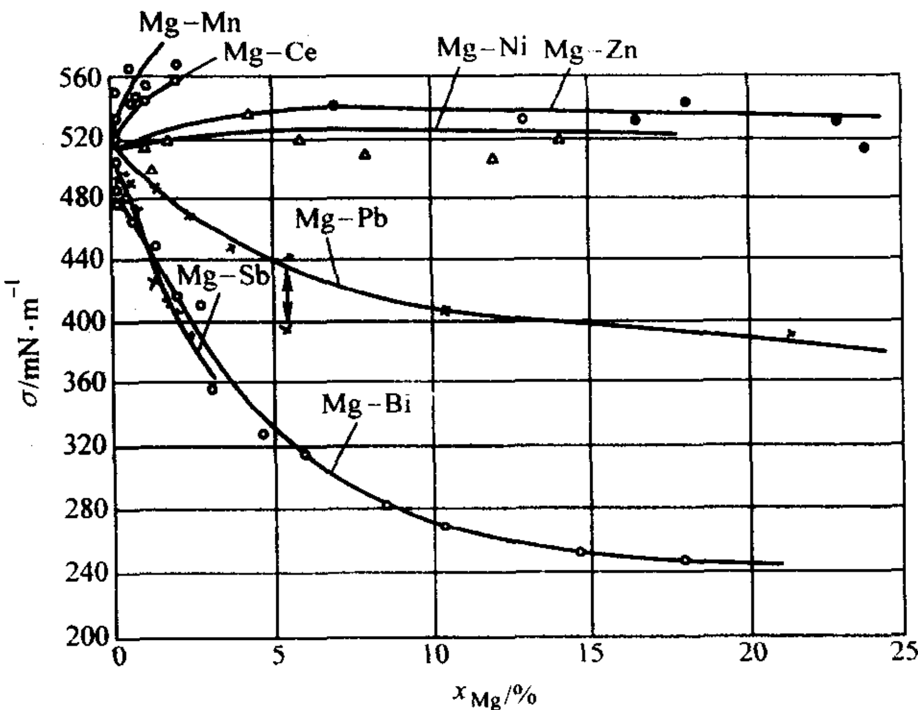
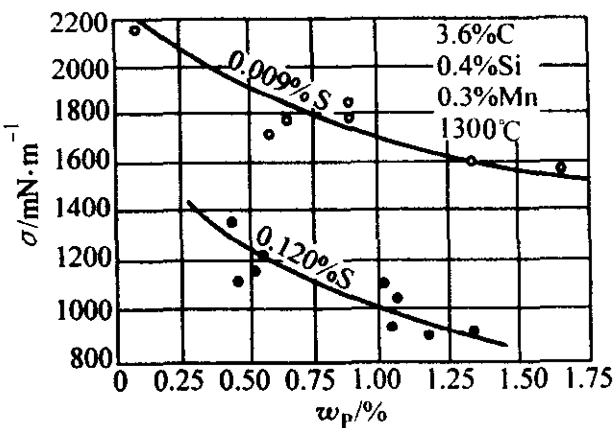
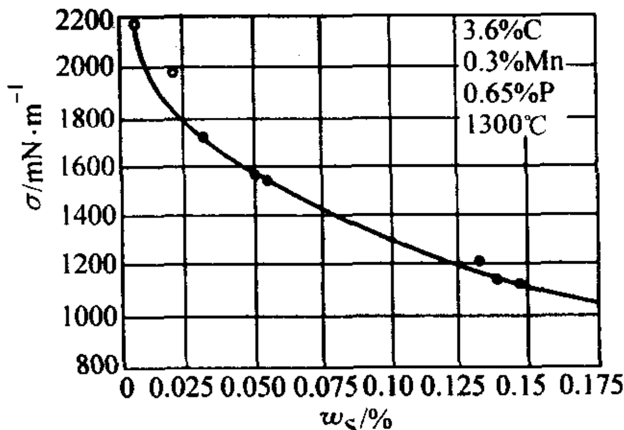


图 2.3-5　Mg 中加入第二组之后表面张力的变化





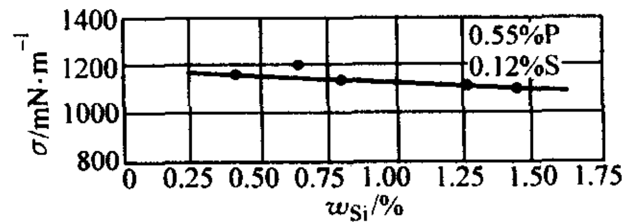


图 2.3-6　S、P、Si 对铸铁表面张力的影响

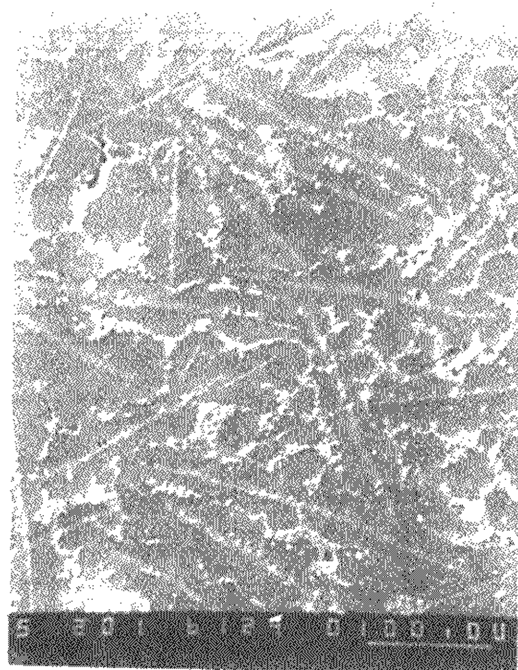
在实际工业生产中，金属的遗传特性主要反映在以下几个方面。

1）固态金属炉料的原始组织和缺陷对凝固后铸件或毛坯组织和缺陷的形成有重要影响，如采用平炉生铁、工业纯铁，或其他种类生铁配制的成分相同的合金铸铁，组织和力学性能差别很大，而原生铁中的某些缺陷如白口、热裂、收缩倾向和气孔等都会不同程度保留在重熔后的铸件中。

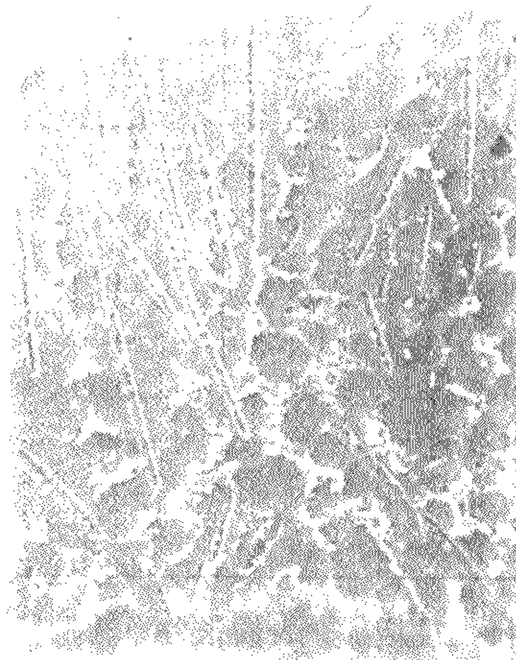
2）液态金属的结构（如过冷度、净化程度等）对凝固后铸件或毛坯的组织有重要影响。

3）液态金属加入合金，可改变金属中元素与元素之间的相互作用，进而影响凝固后铸件或毛坯的组织；值得注意的是，冶炼处理工艺的变化，包括熔化速度、液态金属保温时间和保温温度、变质处理、合金化以及浇注方式不同（重力浇注或压力铸造）等，都会不同程度改变熔体结构，从而引起遗传性能的变化。

图2.3-7显示了金属遗传现象的一个实例。该图为含有针片状 $TiAl_3$ 晶体的AlTiB中间合金组织（图2.3-7a），经重熔仍可发现不少 $TiAl_3$ 针片状结构（图2.3-7b），但变换重熔温度后（见图2.3-7c、d）由于遗传信息经处理遭到损失，针状组织数量减少，且形态发生变化，即由针状成为棒状、板状或少量块状。



(a)原始试样(F6)



(b)850℃重熔



(c)950℃重熔



(d)1000℃重熔

图2.3-7 AlTiB试样重熔前后的SEM组织

关于金属的遗传机理，有看法认为：液态金属是由一些成份和结构不同的游动着的原子集团、以及集团间无序分布的组元原子所组成。在集团内，原子的排列和结合与原有固体相似（可能由固体金属在熔解过程中所保留下来）。原子集团和无序带均在熔体中独立存在，它们由于热能起伏不断，局部地相互退化和重生，在金属熔化过程中，随着熔体温度升高，这个原子集团逐渐分裂，尺寸越来越小，无序区相应扩大。当外部条件使原子集团分离（或缩小）停止，并保留一部分较小的原子集团时，原始炉料中的一些结构信息就有可能被保留下来，传递给后来凝固成形的晶体中。

俄罗斯学者B.N.尼基金总结了金属遗传的几个规律。

1）通过一些特殊方法将组织信息储存在炉料中，就可获得组织遗传基因并控制合金的遗传特性。

2）合金遗传的倾向性大小取决于合金基体与组元之间的物理－化学作用。

3）在“炉料－熔体－铸件”遗传系统中，整体遗传效应不等于诸元素遗传效应之和，这可以解释弱遗传效应的情况。

4）熔体中的弥散质点是炉料金属组织信息的遗传因子（或载体）。

5）从炉料遗传下来的弥散质点是潜在的结晶核心，在合金制备过程中利用这些规律可有目的地控制合金的遗传性。

这些规律的认识，可有效帮助我们认识遗传特性，从而

很好地控制未来凝固金属的组织和性能。但“金属遗传学”的研究由于其复杂性及手段的缺乏，还有待进一步深入。

2 金属材料的凝固生核与长大

在金属材料的成形加工中，铸造、焊接与粉末冶金都与凝固现象有较大关系，液态金属向固相转化的凝固过程则是铸件成形过程的核心，决定了未来铸件的组织、性能，包括缺陷。

2.1 材料凝固过程中的均质形核和非均质形核

材料的凝固过程大体上有两类：一类没有固定的凝固温度，凝固成形的固体为非晶态，如高分子材料、非晶态玻璃、非晶态金属玻璃等；另一类有固定的结晶温度，凝固时会放出潜热，并成为晶体。通常条件下的金属及一些非金属材料形成晶体的凝固行为比不形成晶体的凝固要复杂得多，其凝固过程需经历形核和长大，其趋势由其体系热力学和动力学状态所决定。

2.1.1 结晶过程的热力学条件

液体中的游动原子集团逐步长大到一定尺寸后，可形成稳定的原子集团（或称固体质点），其周围液态原子可向上堆彻形成晶核，是凝固过程的起始阶段。这种晶核的生成有两种方式：①均质生核（homogeneous uncleations）：即由游动的原子集团自身长大形成晶核；②非均质生核（helerogeo-eous uncleations）：在液态金属中外来质点的表面生核。但只有当固态晶核的自由能小于液态晶核的自由能时，物质才能从液态转变为固态（如图 2.3-8 所示）。用 T_0 表示金属的熔点，又称为理论结晶温度。在 T_0 处，液固两相处于平衡，当温度 T 小于 T_0 时，固相自由能低于液相自由能，这是液体结晶的热力学条件。

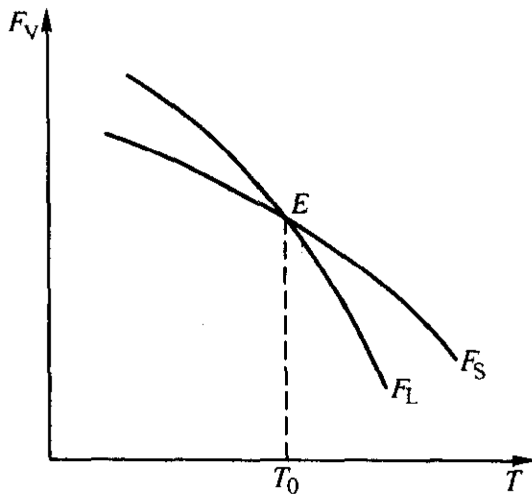


图 2.3-8 合金液态和固态自由能与温度的关系
S—固态；L—液态；F_V—体自由能

2.1.2 均质形核

当液体中出现晶核时，系统吉布斯自由能的变化由两部分组成，一部分是单位液相和固相吉布斯体积自由能差 ΔF_V，为负值，这是相变的驱动力，另一部分是由于固相出现，系统增加了新界面所产生的固相表面能 ΔF_S，为正值，它是相变的阻力。因此，系统总的吉布斯自由能变化为

$$\Delta F = \Delta F_V V + \Delta F_S = \Delta F_V V + \sigma_{LS} S \qquad (2.3\text{-}2)$$

式中，σ_{LS}为固液界面张力；V 为晶核体积；S 为晶核表面积。

假定晶核为球形，r 为球半径，则有：

$$\Delta F = \Delta F_V 4\pi r^3/3 + \sigma_{LS} 4\pi r^2 \qquad (2.3\text{-}3)$$

如果把体系总自由能 ΔF 和表面自由能增量与晶核半径 r 的关系用图形表示（见图 2.3-9），可见，当 $r < r^*$ 时，形成的新相是不稳定，它只能不断变小，才能使自由能增量降低，并导致新相重新熔化。当 $r = r^*$ 时，ΔF 有最大值，如果晶核有可能继续长大，能使自由能增量降低，这个半径 r^* 就称为

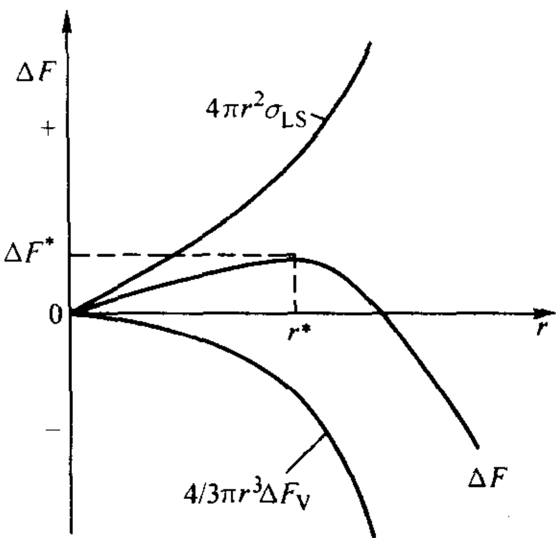


图 2.3-9 自由能的变化与晶核半径 r 的关系

“临界晶核半径”，其值可由式（2.3-3）求极值找到：

$$r^* = -\frac{2\sigma_{LS}}{\Delta F_v} \qquad (2.3\text{-}4)$$

或

$$r^* = -\frac{2\sigma_{LS} T_0}{L\Delta T} \qquad (2.3\text{-}5)$$

将式（2.3-5）代入式（2.3-3），临界晶核半径 r^* 所对应的自由能增量，称为“形核功”，其值为

$$\Delta F^* = \frac{4\pi r^{*2}}{3}\sigma_{LS} \qquad (2.3\text{-}6)$$

从数值上看它等于临界晶核界面能的 1/3。这表明，形成临界晶核时，新、旧相（液、固相）之间总的自由能变化差，只能提供形成新相（固相）所需的表面功 $4\pi r^{*2}\sigma_{LS}$ 的 2/3，其余 1/3，只有通过液相中局部的能量起优来获得。临界半径为 r^* 的晶核如果没有外来的能量补充，是不能长大的，我们只能称之为晶胚，为了使均质形核进行下去，液态金属需要一定过冷度，实验表明，为了出现均质形核所需的临界过冷度约为 $0.2T_m$（T_m 为金属熔点），在这一过冷度下，晶核的临界半径约为 10^{-7} cm,，晶核约含 200 ~ 300 个原子，表 2.3-10 为某些金属均质形核时所能达到的过冷度数值。

表 2.3-10 某些金属均质形核（小液滴）的过冷度数据

金属	熔点 T_m/K	过冷度 ΔT/℃	$\Delta T/T_m$
汞	234.3	58	0.287
镓	303	76	0.250
锡	505.7	105	0.208
铋	544	90	0.166
铅	600.7	80	0.133
锑	903	135	0.150
铝	931.7	130	0.140
锗	1 231.7	227	0.184
银	1 233.7	227	0.184
金	1 336	230	0.172
铜	1 356	236	0.174
锰	1 493	308	0.206
镍	1 725	319	0.185
钴	1 763	330	0.187
铁	1 803	295	0.164
钯	1 828	332	0.182
铂	2 043	370	0.181

晶核的形成速率 I^* 可用下式表达

$$I^* = k_v \exp\left(-\frac{\Delta F_A + \Delta F^*}{kT}\right) \tag{2.3-7}$$

式中，ΔF_A 为液态金属中原子扩散激活能；ΔF^* 为生核功；k 为玻耳兹曼常数；k_v 为系数。

ΔF_A 和 ΔF^* 都与结晶过冷度 ΔT 有关，生核速度在 ΔF_A 和 ΔF^* 相互作用下，随过冷度 ΔT 增加时，有一个最大生核速度。

2.1.3 非均质形核

由于液相中存在大量夹杂颗粒，金属氧化膜，以及铸型表面的存在等等，实际液体金属都在显著小于均质形核理论所预料的过冷度下以非均质形式形核，如大多数金属在低于熔点温度仅十度左右时就可以形核。

但并非所有外来质点的表面都能使新相生核，而需要一定的条件。有两种非均质形核机理对非均质形核发生的可能性进行了描述，一种是界面润湿理论，另一种是晶格对应理论。

按照界面润湿理论，假设生核衬底质点表面为平面，所形成的晶粒为球面（球冠），如图 2.3-10 所示。σ_{LC}表示液相与固相之间的界面张力，σ_{LS}为液相与质点之间的界面张力，σ_{CS}为固相与质点之间的界面张力，θ 称为新相与质点衬底的湿润角，那么这三个界面张力有以下平衡关系

$$\sigma_{LS} = \sigma_{CS} + \sigma_{LC}\cos\theta \tag{2.3-8}$$

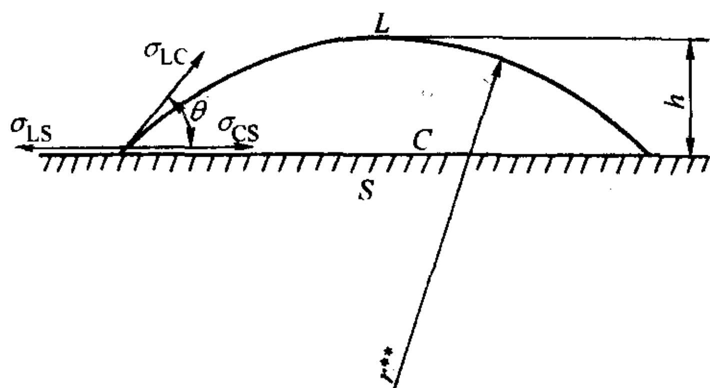


图 2.3-10 界面润湿理论非均质生核示意图

经计算可得出非均质形核功为

$$\Delta F^{**} = \frac{16}{3}\,\frac{\pi\sigma_{LC}^3 T_0^2}{L^2\Delta T^2}\left[\frac{(2+\cos\theta)(1-\cos\theta)^2}{4}\right] \tag{2.3-9}$$

将此式与均质形核相比，可得其比值为

$$\frac{\Delta F^{**}}{\Delta F^*} = f(\theta) = \frac{(2+\cos\theta)(1-\cos\theta)^2}{4} \tag{2.3-10}$$

可见，当 $\theta = 180°$时 $f(\theta) = 1$，$\Delta F^{**} = \Delta F^*$，其生核功等于均质生核的生核功，新相不能依附于外来质点的表面生核。当 $\theta = 0°$时、$f(\theta) = 0$，$\Delta F^{**} = 0$，质点表面即是晶核的晶面，新相可在其表面外延生长。可见，只要新相与质点有所润湿，即 $0° \leqslant \theta \leqslant 180°$，都有利于形核。

非均质形核和表面润湿理论有其不完善性。由于在非均质形核中只有几十个原子参与形核过程，用实验方法测定润湿角 θ 十分困难。大量的实验观察和理论分析表明，两相之间的润湿现象与界面上两相相应晶面的结构和原子间的结合力存在密切关系，即晶格对应（或共格）对非均质形核有重要影响。

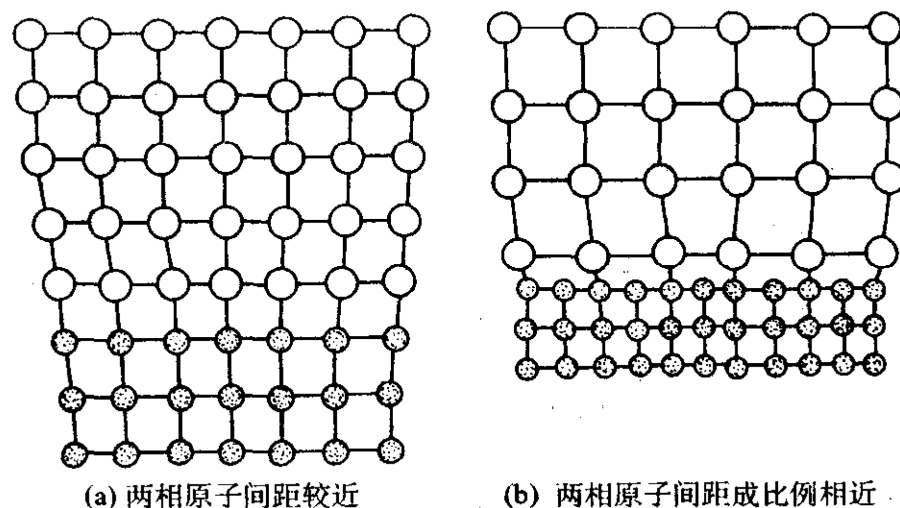
晶格对应理论说的是，新相原子在固相质点表面生核，一定要符合共格对应原则，它可用两相界面上原子的间距（或点阵常数）差来说明：

$$\delta = \frac{\Delta a}{a} \times 100\% \tag{2.3-11}$$

式中，a 为新相表面的原子间距；Δa 为新旧两相界面上原子间距之差。

有观点认为，当 $\delta \leqslant 5\%$时，两相的相界面能共格，即易于生核；当 $\delta = 5\% \sim 25\%$，两相界面共格不好，生核稍难。

新相与固体质点产生界面共格对应，两相的晶格类型可以相同，也可以不同，但固体质点表面上原子的排列方式与新相中某一晶面上原子的排列方式相似，而其原子间距相近或成比例（如图 2.3-11）。



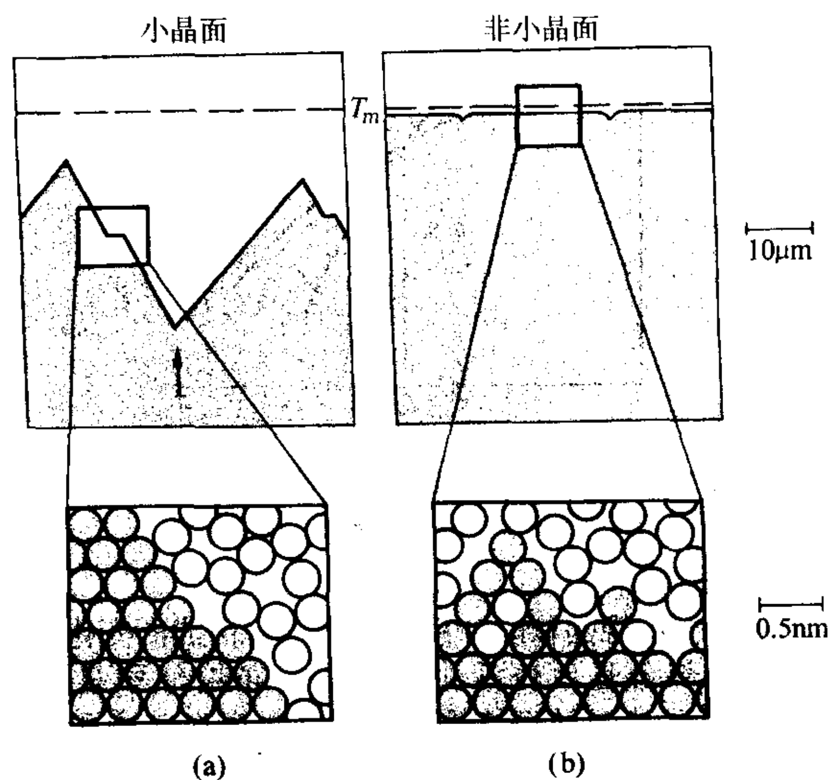
(a) 两相原子间距较近　(b) 两相原子间距成比例相近

图 2.3-11 结晶相在固体质点上外延生长及原子对应情况

而晶格对应理论也有一些与事实不符现象，如金属基底比非金属基底更能为形核提供有效结晶基底。金属来源原子间距不超过 9%时，也可为结晶生核提供有效基地。

2.2 固-液界面结构

晶体的生长是单个原子向生长表面的堆砌，所以界面结构对原子堆砌方式和堆砌速度有较大影响，并影响晶体的生长方式、生长速度和最后的形态。从原子的尺度看，固－液界面的结构可以分为粗糙界面（非小晶面）和光滑界面（小晶面）两类（如图 2.3-12）。



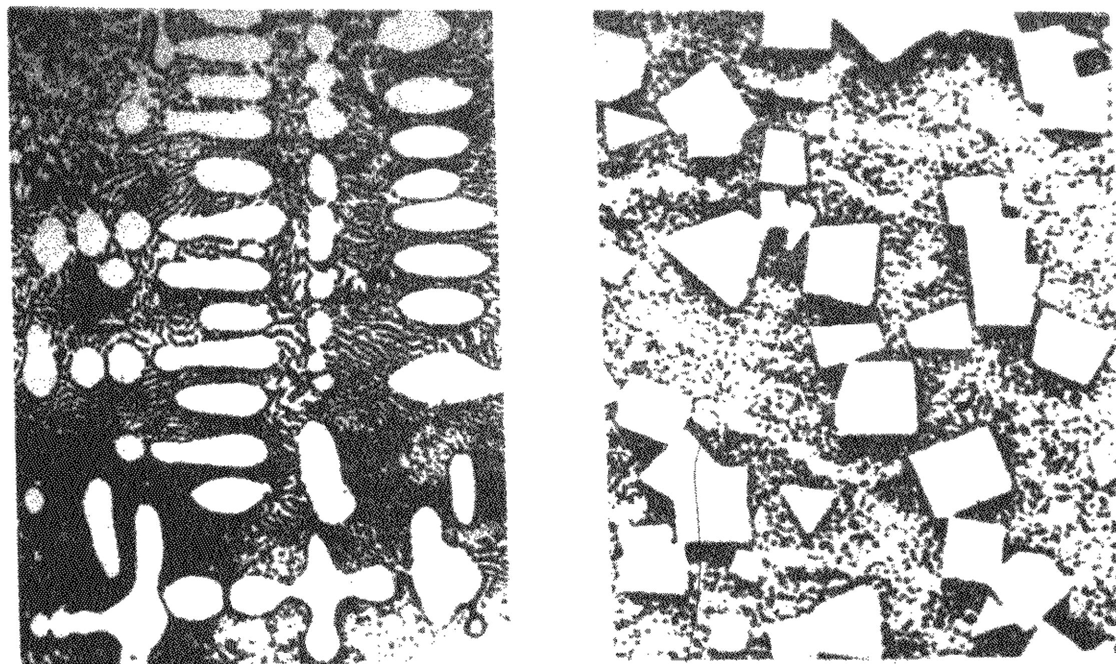
(a)　(b)

图 2.3-12 晶体生长的两种固液界面结构

1）光滑界面（也称小晶面 faceted)：固相表面原子层基本上是充满的（95%），因而可以说固液界面上的原子排列是光滑的。但从宏观尺寸度看却是不光滑的（如图 2.3-12a），具有锯齿的固－液界面，并显示出明显的结晶面特征。这种晶体的不同晶面长大的速度不一样，高指数的晶面、长大时向前（垂直于晶面方向）推进的速度很快，最后晶体被低指数晶面包围，从而形成有棱角外形。类金属及金属间化合物、矿物、一些有机物晶体均属此类。

2）粗糙界面（也称非小晶面 monfaceted)；固相表面最

外几个原子层均有 50% 左右的位置未被充满，界面上的原子排列是粗糙的（如图 2.3-12b）。但从宏观尺度看固液界面形貌是平滑的，且显示不出任何结晶面的特征，原子在向固-液界面上附着时是各向同性的。原子供给取决于热流和溶质原子的扩散状态，哪个方向传热、传质快，哪个方向就长大得快。与此同时，由于界面能的各向异性，这类晶体在长大方向上有择优取向的倾向，表现在树枝晶的主干有一定的结晶取向。大部分金属结晶时，固相界面属于粗糙界面。图 2.3-13 即为粗糙界面和光滑界面结晶形貌的实例。



(a) Cu- Ag共晶基底上Ag的非小晶面枝晶　(b) 富Sn基底上 β －SnSb 化合物小晶面立方晶体

图 2.3-13　金属合金晶体生长时的固－液界面组织

究竟哪类物质属于粗糙界面结晶，哪类物质属于光滑界面结晶，主要处决于它们的熔化熵值。

假定在一个平整界面上无规律堆砌一些原子，用统计物理方法可计算出在固液生长界面上相对吉布斯自由能变化（ΔF_S）与界面沉积原子几率的关系

$$\frac{\Delta F_S}{NkT_m} = \alpha x(1-x) + x\ln x + (1-x)\ln(1-x) \quad (2.3\text{-}12)$$

式中，N 为界面上可供原子占据的全部位置；$x = N_A/N$ 为全部 N 位置中原子所占据位置 N_A 的分数；k 为玻耳兹曼常数；T_m 为熔点或平衡凝固温度；α 为系数。将式（2.3-12）按 $\frac{\Delta F_S}{N_k T_m} = f(x)$ 作图（见图 2.3-14）可知，按 α 取值的不同，可得出两种性质不同的结晶界面类型。

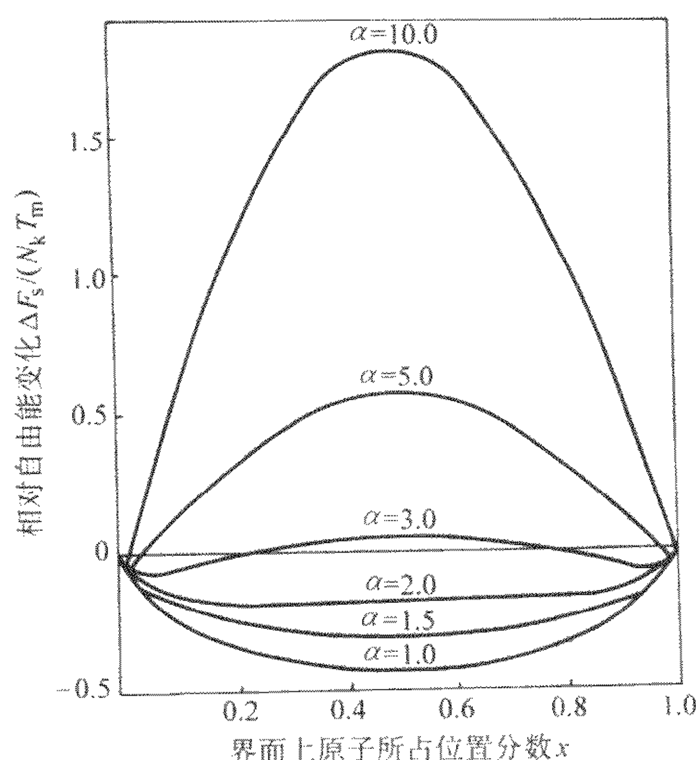


图 2.3-14　不同 α 值时相对吉布斯自由能变化与界面上原子所占位置分数的关系

当 $\alpha \leqslant 2$ 时，曲线可出现一个极小值，并在 x 约等于 0.5 的位置出现这个最低值，其物理意义是此时界面上可能被原子占据位置为 50% 左右时，界面的相对自由能最小，状态最稳定。从原子尺度上看这类界面是"粗糙"的，即在这类界面上还有约 50% 的位置可以随意添加原子，因此这类物质结晶时，晶体长大速度较快。

当 $\alpha \geqslant 3$ 时，曲线仅在 $x < 5\%$ 和 $> 95\%$ 时，ΔF_S 才会有最小值，其物理意义是添加原子在界面只占据 < 5% 或占满 95% 以上时，界面的相对自由能最小，状态也最稳定。故而界面是光滑平整的。α 值愈大，界面愈光滑。由于这类晶体在长大过程中要保持"光滑"状态，所以长大速度相对较小，而且自由生长时，容易形成规则的几何外形。

这表明 α 是一个区分晶体界面结构的重要参数：

$\alpha > 2$，界面平整；

$\alpha < 2$，界面粗糙。

α 系数与熔化熵 ΔS_m 有关，可简化为下式表示

$$\alpha \approx \Delta S_m / R \approx \Delta S_m / 8.31$$

其中，$R = 8.31$ J/（mol·K）。

大多数金属的 $\alpha < 2$，而许多非金属的 $\alpha > 2$，某些金属的熔化熵值见表 2.3-11。

实际研究结果证实，判断物质是按粗糙界面长大还是按光滑界面长大，单靠熔化熵值的大小是不够的，它还和结晶动力学，即物质在溶液中的浓度及凝固过冷度有关。如 Al－Sn合金中，随 Al 浓度减少，初晶 Al_α 的形貌可由粗糙界面转变为光滑界面。又如白磷在低的长大速度时为光滑界面，当长大速度增加时却转变为粗糙界面。因此 D.E.Temkin 等人提出了固－液界面的多原子层模型（如图 2.3-15 所示）。在这个多原子层界面中，既存在着原子排列较为规则的原子簇，又存在排列非常紊乱的原子，在排列规则的原子簇中的晶体位置被部分填满，并与一定的晶面相对应，随着向固相一边靠近，原子簇中原子排列的有序化程度增加。因此在过冷度较小时（即熵值比较低的金属），界面原子层数较少，生长可按原子团中每层台阶的侧面扩展方式进行，其生长固－液界面为光滑界面。对过冷度较大的情况（熵值也较大），固－液界面原子层变厚，界面上排列混乱的原子数增多，粗糙度增加，因此，即使原来属于光滑界面生长的物质，也可以转变为粗糙界面生长。

表 2.3-11 某些金属的熔化熵 ΔS_m 值

晶格类型	晶体	T_0/K	L/J·mol⁻¹	ΔS_m/J·mol⁻¹·K⁻¹	晶格类型	晶体	T_0/K	L/J·mol⁻¹	ΔS_m/J·mol⁻¹·K⁻¹
面心立方	Al	933	10 470	11.2	体心立方	Ca	1 123	8 790	7.8
	Cu	1 356	12 980	9.6		Cr	2 163	19 260	8.9
	Au	1 436	12 770	8.9		Fe	1 912	15 490	8.1
	Pb	600	4 815	8.0		Li	459	2 930	6.9
	Ni	1 728	17 670	10.2		K	336	2 390	7.1
	Pt	2 046	21 770	10.6		Na	311	2 640	8.5
	Ag	1 233	10 890	8.8		W	3 683	35 170	9.5
密排六方	Cd	594	6 410	10.8	正方	Sn	505	6 990	13.8
	Mg	923	8 790	9.5	三方	Bi	544	10 890	19.6
	Zn	692	7 290	10.5	斜方	Ga	303	5 590	18.5

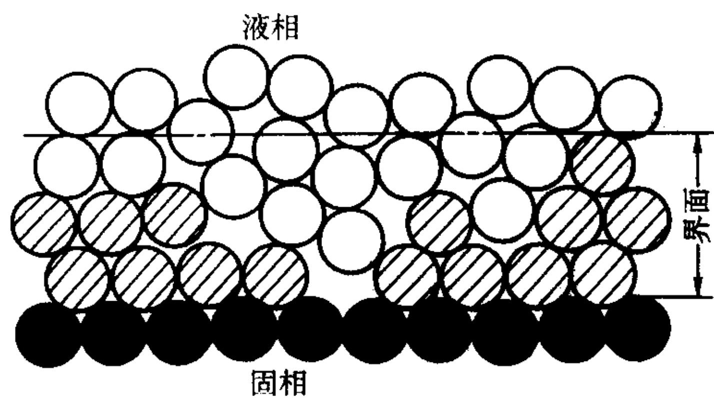


图 2.3-15 固-液界面的多原子层模型

2.3 凝固晶粒的长大方式

由于固液界面结构不同，晶体的长大方式也不一样，可分为以下三种。

1) 连续长大（continuous growth） 固-液界面在原子尺度上是“粗糙的”，大多数金属在固-液界面上所有位置，对添加原子而言是等效的。原子可以连续，无序地向界面添加，从而使界面迅速向液相推移，大多数金属都是通过这种方式长大，生长所需的动力学过冷度较小（ΔT_k 约 0.01～0.05 K）生长速度也较高。晶体长大时生长速度 R 与结晶过冷 ΔT_k 的关系最大

$$R = \mu_1 \Delta T_k \tag{2.3-13}$$

式中，μ_1 为连续生长系数。

2) 二维晶核台阶长大 这种生长方式属于光滑界面的侧面长大（lateral growth）。原子主要借助于界面上形成的台阶侧面堆砌，并通过台阶的侧面生长增厚，界面向前推进，因此界面向前推进的方向与台阶扩展方向相垂直（如图 2.3-16）。在这种光滑的生长界面上，具有独特的密排晶面的晶体学特征，晶面上原子间的结合力较强，原子不易脱落，界面保持完整。但在这种界面上，原子堆砌难，也很不稳定。台阶的来源可以是界面上的二维形核，也可以是界面上的晶体缺陷。

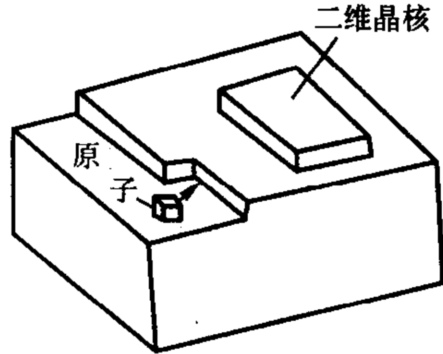


图 2.3-16 二维晶核台阶示意图

3) 沿晶体缺陷长大 实际晶体生长时总要形成许多生长缺陷，这些缺陷所造成的界面台阶容易使原子向上堆砌，其中对晶体生长过程影响较大的是螺旋位错和孪晶。

螺位错生长台阶如图 2.3-17 所示。螺位错使界面出现现成的生长台阶，原子在台阶上堆砌时，台阶便绕位错旋转，台阶每转一周，界面便生长出一个原子层。由于一般螺位错的台阶不会消失，可保证界面沿螺位错线连续生长下去。这种生长所需能量小于二维晶核生长模式，生长速度也较之要快。

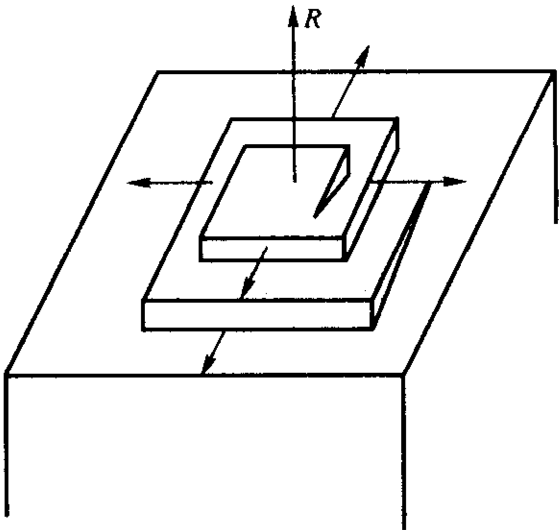


图 2.3-17 螺位错生长台阶及晶体的生长方式

螺位错生长方式在铸铁石墨基面（0001）的生长中作用很大，在片墨或球墨中的生长表面常可看到螺位错生长的六角形突台。

孪晶生长在片状石墨生长过程中也有很重要作用（如图 2.3-18 所示），石墨晶体具有以六角网络为基面的层状结构，晶面之间结合力弱，在生长过程中由于原子排列的层错，使上下层之间产生一定角度旋转，构成旋转孪晶。在孪晶的旋转边界上存在许多台阶可供碳原子堆砌，使石墨晶体侧面 <10$\bar{1}$0> 向生长加快而成为片状。

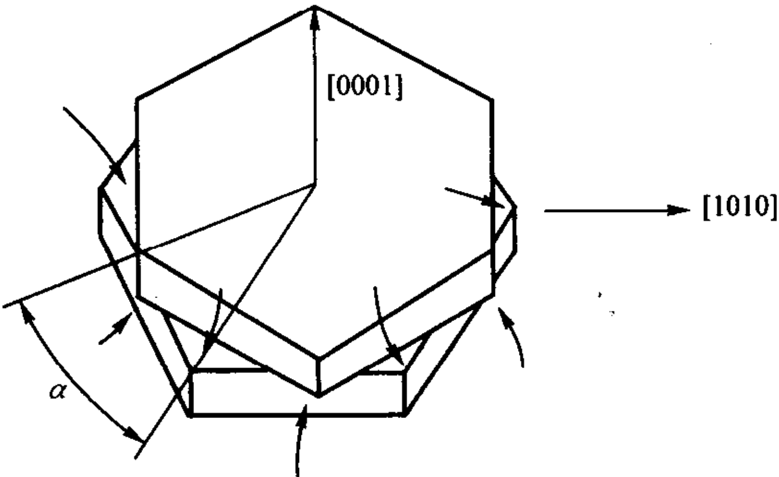


图 2.3-18 石墨的旋转孪晶及其边界所形成的生长台阶

而面心立方 Si、Ge 等晶体又是以反射孪晶模式，依靠孪晶（111）面组成的凹坑（如图 2.3-19）生核长大。

三种生长方式的生长速度与过冷度之间关系比较如图 2.3-20，生长速度最快者为连续式生长方式（粗糙界面），

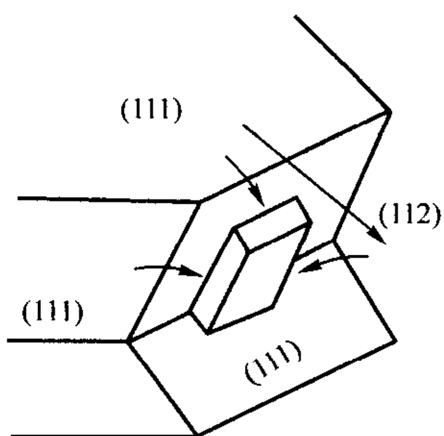


图 2.3-19 面心立方晶体反射孪晶及其凹坑边界

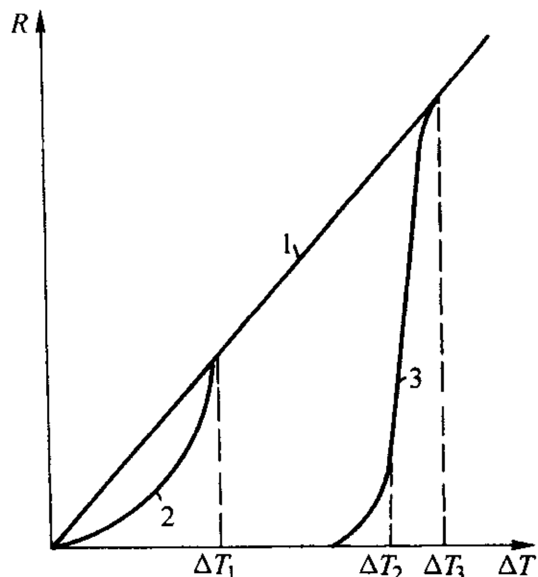


图 2.3-20 三种生长方式的晶体生长速度与过冷度关系
1—粗糙界面的连续生长；2—通过螺型位错机理生长；
3—通过二维生核生长

二维晶核台阶生长方式较慢。

3 液态金属凝固过程中的传热

3.1 液态金属凝固过程中的传热特点

热量传递是液态金属凝固过程中的重要现象。高温的液体金属浇注铸型时，金属所含的热量（包含凝固潜热）将通过液态金属、已凝固的固态金属、金属－铸型的界面和铸型而传出。液态金属在这个传热过程中存在两个主要界面，即金属固－液相界面和固相金属－铸型间的界面，并且这两个界面随凝固进程而发生动态迁移（如图 2.3-21）。值得注意的是，金属和铸型间界面的接触通常是不紧密的，存在接触热阻或界面热阻，是影响热量传递的重要因素。由图 2.3-21 可见，凝固开始后，型壁处的液态金属逐步变为固相，并在凝固前沿释放出潜热，在凝固过程中，由于铸型的膨胀和金属的凝固收缩，金属和铸型界面的热阻发生变化，可能形成间隙（或气隙），并使传热方式发生变化（以对流换热为主要传热方式），其热量传递也大受影响。

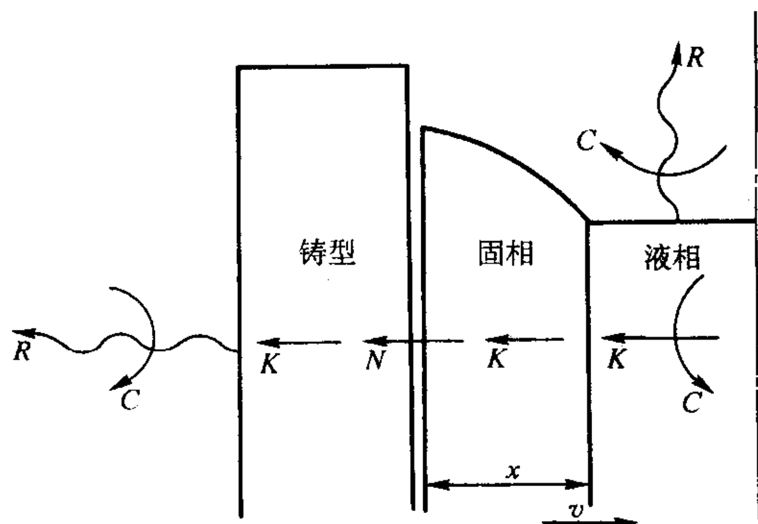


图 2.3-21 纯金属在铸型中凝固时的传热模型
K—导热；C—对流；R—辐射；N—牛顿界面换热

因此，金属在铸型中的凝固过程伴随有三种传热方式：传导、对流和辐射。在凝固过程中，热传导是最主要的传热方式，其次是对流，辐射的影响最小。在液态金属的成形过程中，铸型或锭模传热直接影响晶体的形核和生长、影响液态金属在凝固过程中的流动、影响补缩的大小和分布，从而对晶体形貌和溶质的偏析，以及最终性能产生影响。我们可以把金属假定为连续介质，并认为金属是均匀的和各向同性的。对于稳定的一维热传导过程，我们可以用傅里叶导热定律来描述

$$q = -\lambda G \tag{2.3-14}$$

式中，λ 为热导率；G 为温度梯度。

对于非稳态的传热，我们可以用如下三维热传导方程来描述，它反映了热传导过程中的能量守恒

$$\lambda\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}\right)+q^* = c_p\rho\frac{\partial T}{\partial t} \tag{2.3-15}$$

或

$$\frac{\partial T}{\partial t} = \alpha\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}\right)+c \tag{2.3-16}$$

式中，T 为热力学温度；q^* 为单位体积物体单位时间释放的热量；c_p 为金属比热容；ρ 为金属密度；t 传热时间；α 为导温系数，$\alpha=\frac{\lambda}{\rho c_p}$；$C$ 为系数，$C=\frac{q^*}{\rho C_p}$。

在对流换热情况，如已凝固金属与铸型的界面换热过程，可引入如下界面换热系数 h（$J/m^2\cdot s\cdot K$）来计算传热量：

$$q = h\ (T_s - T_m) \tag{2.3-17}$$

式中，T_s 为界面上已凝固金属温度；T_m 为界面上铸型温度。

由于凝固过程中的界面现象使传热过程在微观上变得复杂。当固液界面凹凸不平或生长为枝晶时，在凝固前沿，热总是沿垂直于这些界面的不同方位从液相传入固相，发生三维传热现象。在金属和铸型界面上，除导热过程外，也可能同时存在微观的对流和幅射传热，但在实际计算过程中，由于对流和幅射值较小，特别是幅射热可略去不计，以使问题简化。

在铸件凝固过程中，如果不计液态金属的热阻，金属的凝固速度主要受如下三种热阻的控制

$$R_s = s/\lambda_s \tag{2.3-18}$$

$$R_m = I_m/\lambda_m \tag{2.3-19}$$

$$R_i = 1/h_i \tag{2.3-20}$$

式中，R_s 为已凝固的固体金属层热阻；R_m 为铸型热阻；R_i 为固相金属和铸型界面热阻；s 为已凝固金属层厚度；I_m 为铸型厚度；h_i 为固相金属和铸型界面换热系数。

在砂型铸造中，铸型的热阻 R_m 远大于 R_s 和 R_i，凝固速度主要取决于铸型热阻 R_m，在金属型铸造、压铸或连铸中，R_i 远大于 R_s 和 R_m，是凝固速度的主要制约因素。

3.2 非金属型铸造时（如砂型）的凝固传热

非金属型（如砂型）的特点是，与浇注其中的金属相比热导非常小，因此，金属的凝固速度处决于铸型的传热性能。由于铸型导热能力差，在金属凝固时，铸型外壳的温度变化不大，可假定铸型为半无限大，或铸型外表温度等于环境温度 T_0。假定浇注的为纯金属、浇注温度为熔点 T_m，即金属无过热，在浇注结束瞬间（浇注时间 $t=0$），铸型内表面温度可立刻升至 T_m。若不考虑铸型和固相金属间的热阻（假定很小忽略），则温度场分布如图 2.3-22 所示，这时非金属型中的温度场 T 可用如下一维非稳定传热方程求解

$$\begin{cases}\frac{\partial T}{\partial t} = \alpha_m\frac{\partial^2 T}{\partial x^2}\\ x=0,\ T=T_m\\ x=-\infty,\ T=T_0\end{cases} \tag{2.3-21}$$

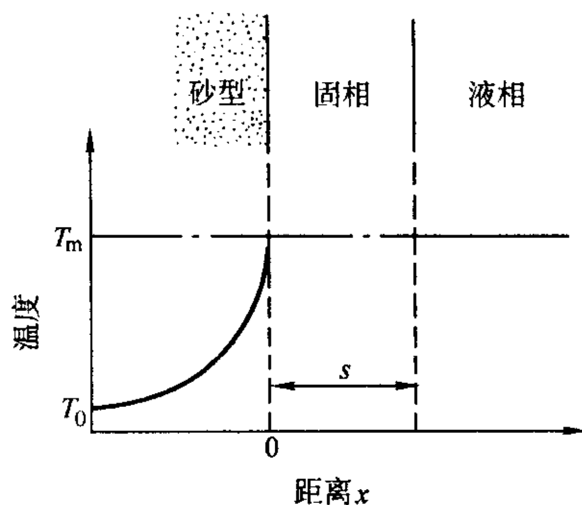


图 2.3-22　无过热液态纯金属在砂型中凝固时的近似温度分布

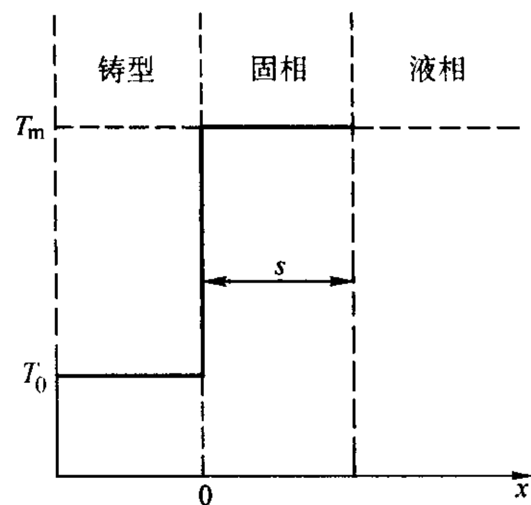


图 2.3-23　无过热液态纯金属在金属型中凝固时的温度分布

式中，α_m 为铸型热扩散率，$\alpha_m=\frac{\lambda_m}{c_m\rho_m}$；$t$ 为凝固时间。

因而可求出已凝固金属层厚度 s 与凝固时间 t 的关系

$$s=\frac{2}{\sqrt{\pi}}\left(\frac{T_m-T_0}{\rho_s H}\right)\sqrt{\lambda_m c_m\rho_m}\sqrt{t} \tag{2.3-22}$$

或

$$s=c\sqrt{t} \tag{2.3-23}$$

式中，ρ_s 为已凝固固相金属密度；H 为金属熔化潜热；c 为凝固常数。

上式为著名的平方根定律，由著名学者 Chvorinov 在 1930 年通过实验提出。其物理意义是金属凝固层厚度与凝固时间的平方根成正比，即金属的凝固速度开始较快，尔后随铸型的温度升高而逐渐变慢。平方根定律是铸造成型中，浇注系统设计的重要理论依据。

由式（2.3-22）可知，金属凝固速度由两部分因素决定，其一部分 $\sqrt{\lambda_m c_m\rho_m}$ 反映了铸型的热物性参数对凝固过程的影响，即反映了铸型在凝固过程中的吸热能力；$\frac{T_m-T_0}{\rho_s H}$ 与金属本身的性能有关，如熔点高而熔化热和密度小的金属则有利于较快凝固。

对于一个体积为 V，表面积为 A 的实际铸件，在非金属铸型中凝固时，也可得到一个求凝固时间的公式

$$V/A=\frac{2}{\sqrt{\pi}}\left(\frac{T_m-T_0}{\rho_s H}\right)\sqrt{\lambda_m c_m\rho_m}\sqrt{t_f} \tag{2.3-24}$$

或

$$M=c\sqrt{t_f} \tag{2.3-25}$$

式中，M 可称为等效凝固厚度，$M=\frac{V}{A}$。

3.3　以界面热阻为主（如金属型铸造）的凝固传热

以界面热阻为主的铸件凝固现象在铸造成形中十分普通，如金属型铸造、压铸等。由于金属型有很高的导热性，所以在铸件凝固过程中，热流的限制环节通常不在铸型，而在铸件与铸型之间的界面，铸件凝固收缩和铸型受热膨胀在铸件铸型间形成气隙时，界面热阻的作用将更加突出。

我们仍讨论一维问题，假定金属液无过热，浇注温度为熔点 T_m。金属型半无限大，远离界面处等于环境温度 T_0，且铸型和固相金属内的温差忽略，其温度场 T 分布如图 2.3-23 所示。

在铸型和固相金属界面，由于固相析出的潜热量等于铸型通过热交换所排出的热量，所以可得如下平衡式

$$h(T_m-T_0)=\rho_s H\frac{\mathrm{d}s}{\mathrm{d}t} \tag{2.3-26}$$

式中，s 为凝固金属的厚度；ρ_s 为固相金属密度；H 为金属熔化潜热；h_i 为铸型和固相金属界面换热系数。

引入初始条件 $t=0$，$s=0$ 并解方程得

$$s=h\frac{T_m-T_0}{\rho_s H}t \tag{2.3-27}$$

或

$$s=ct \tag{2.3-28}$$

可见，当热阻主要在界面时，或金属型与固相金属导热优良时，已凝固金属层厚度与所用时间 t 为线性关系。

与热阻在铸型时的情况类似，金属的凝固速度与界面换热能力和金属本身的性质有关。

4　液态金属凝固过程中的溶质分配

按照合金在凝固过程中晶体形成特征，我们把合金分为两类：①单相合金——在凝固过程中只析出一个固相，如固溶体、金属间化合物等。在结构材料中常用的单相合金有高碳钢、钛合金，以及部分低碳钢、高温合金，铜合金、镁合金、铝合金等。亚共晶铸铁在高温时先析出的初生奥氏体的过程也可认为是单相合金凝固；②多相合金——在凝固过程中同时析出两个以上相的合金，如共晶成分的 Fe－C 合金、Al－Si 合金，还有包晶或偏晶转变的合金，也可称为多相合金。

在铸造生产中，单相合金主要指互相无限固溶或合金元素含量少，在凝固时只形成单一固溶体的合金（如图 2.3-24 所示），溶质原子主要以间隙或置换方式存在于基体原子的点阵中。

在合金凝固过程中，随温度变化，固相和液相浓度随时间变化而不同。我们可将单相合金凝固的相图简化为图 2.3-25，即假定液相线和固相线均为直线，则依此关系有一平衡分配系数 k，其值可大于或小于 1。

$$k=\frac{\text{某温度 }T\text{ 时固相的溶质成分}}{\text{该温度时液相中溶质的浓度}}=\frac{C_s}{C_L} \tag{2.3-29}$$

4.1　单相合金的平衡凝固

在平衡结晶条件下，即在凝固过程中，固相和液相都能充分扩散，因此在凝固的任一时刻，固相和液相成分都是均匀的，则在凝固中途某时刻，固相浓度 C_s 和液相浓度 C_L 有如下分布（见图 2.3-26）。其最终固相成分是均匀的。

其结晶过程中浓度表达式可如下式所示，该式也称为杠杆定律。

$$C_s f_s+C_L f_L=C_0 \tag{2.3-30}$$

式中，f_s、f_L 为凝固某时刻固相和液相所对应的体积分量，$f_S+f_L=1$。

4.2　溶质在固相中不扩散，在液相中充分扩散的凝固

在普通铸造过程中，冷速较快，因此都是不平衡凝固。由于固相中的扩散速度比液相低几个数量级。所以我们可假定固相中的溶质在凝固过程中不扩散，而液相中的溶质可充分扩散。则其在凝固过程的某一时刻，其浓度分布如图 2.3-27 所示。凝固过程中开始析出固相的瞬间浓度为 kC_0 并随界面推进，固相浓度 C_s 不断提高，而液相浓度始终是均匀的。由于溶质在液相中的富集，因此在凝固过程中，有可能

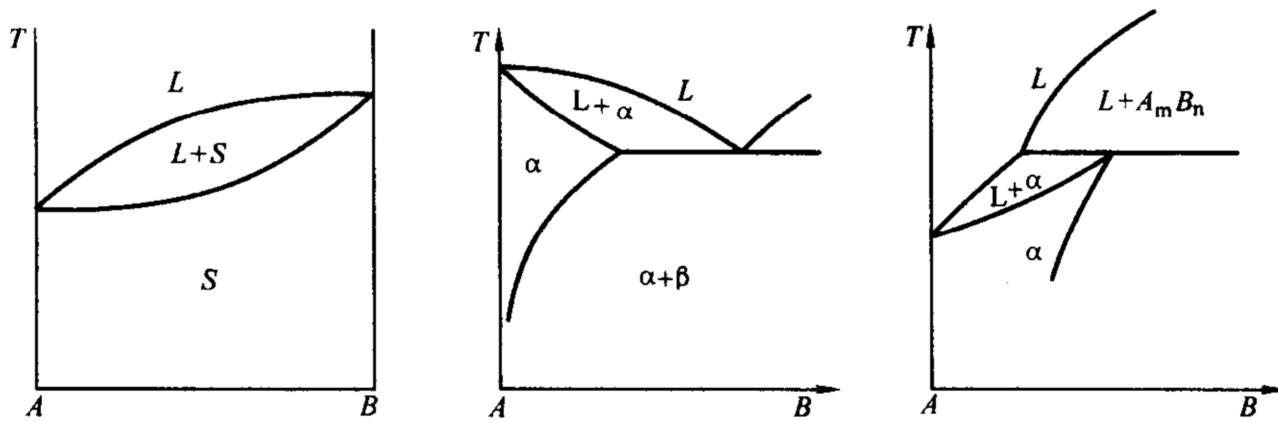
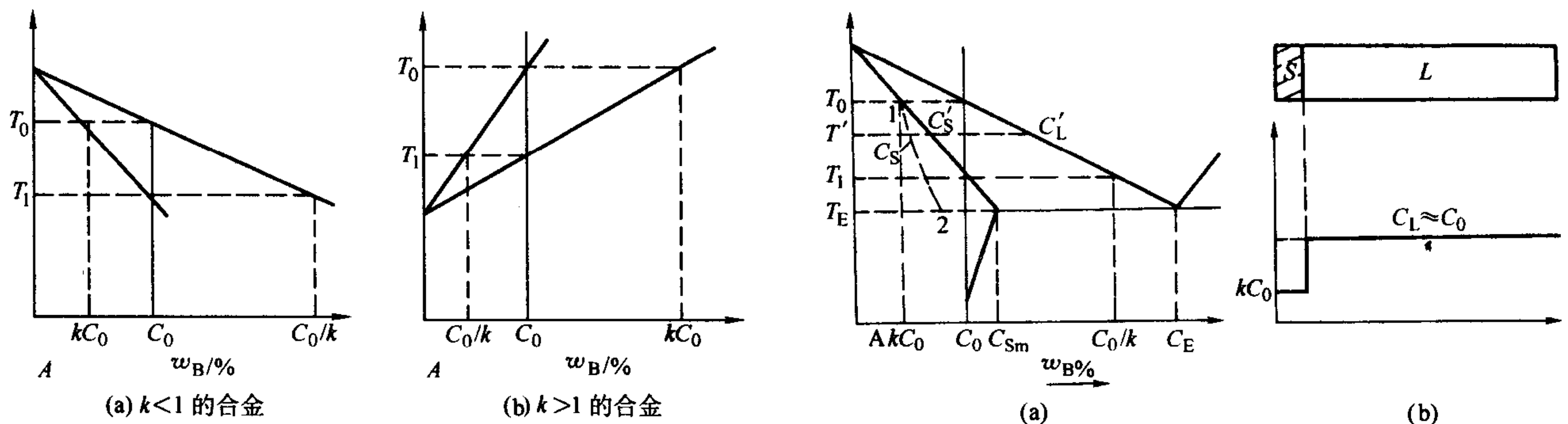


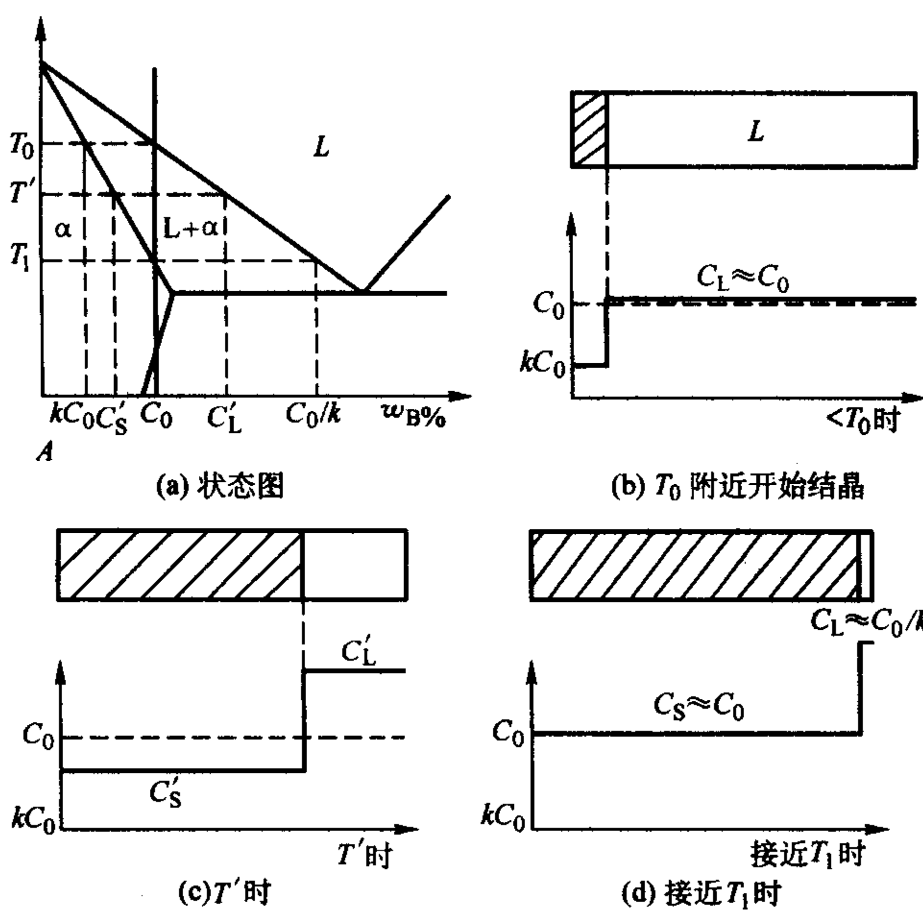
图 2.3-24　几种合金系中凝固时只析出单相固溶体的合金



(a) $k<1$ 的合金　(b) $k>1$ 的合金

图 2.3-25　单相合金单相区部分

(a)　(b)



(a) 状态图　(b) T_0 附近开始结晶

(c) T'时　(d) 接近T_1时

图 2.3-26　单相合金平衡凝固固液界面浓度分布

C_s 浓度在某时刻会大于 C_0，且 C_L 浓度会在凝固末期达到共晶成分而发生少量共晶反应。愈接近凝固末期，析出固相的溶质含量愈高。这种成分不均匀产生于晶粒之内，称之为晶内偏析。

这种凝固过程某时刻固液相界面两相浓度分别为

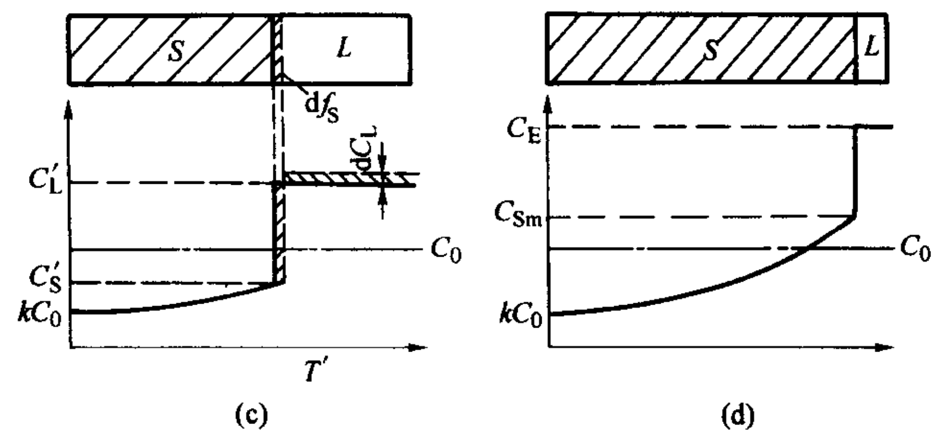
$$C_s = kC_0\ (1-f_s)^{k-1} \tag{2.3-31}$$

$$C_L = C_0 f_L^{k-1} \tag{2.3-32}$$

该式也可称为非平衡杠杆定律，由夏尔（Scheil）在 1942 年提出。

4.3　溶质在固相不扩散，在液相中有限扩散而无对流的凝固

如图 2.3-28 所示，在液相中没有对流或搅动，在开始



(c)　(d)

图 2.3-27　单相合金固相不扩散，液相完全扩散时的固液界面浓度分布

过度区内，固相浓度可从 kC_0 逐步增加到 C_0，大量溶质原子富集在界面液相内，并逐步由界面向液体金属内扩散，由于熔池长，液相中的浓度曲线 C_L 成指数衰减函数分布，在 $x=\infty$ 处，液相浓度 C_L 趋近于原始浓度 C_0 时，界面液相浓度从 C_0 逐渐增加至 C_0/k。当界面固相达到该合金成份 C_0 时，固液界面开始稳定生长，在这个阶段，固相成分和界面处液相成份分别为 C_0 和 C_0/k 始终不变。在这个稳定生长过程中，界面析出固相排出的溶质数与液相扩散送走的溶质量是相等的。在凝固末期，当液相内溶质富集层厚度等于剩余液相区长度时，溶质扩散受到试样末端边界的阻碍，使固-液界面处的 C_L^* 与 C_s^* 同时升高。由于质量守恒，最初过渡区溶质贫乏总量等于最后过渡区溶质的过剩总量，即图 2.3-28 中 $A_1 = A_2$。

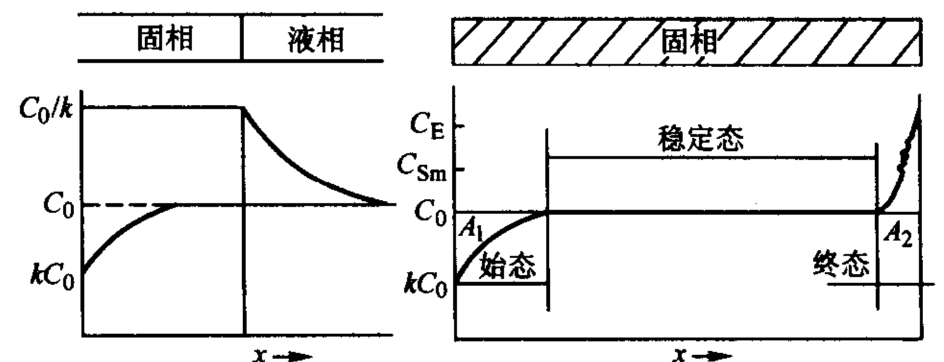


图 2.3-28　单相合金固相不扩散、液相扩散受阻但无对流时的固液界面浓度分布

存在界面液相溶质富集是这种结晶方式的最大特点，稳定生长阶段界面附近液相浓度 C_L 可由下式导出

$$\begin{cases} R\dfrac{\partial C}{\partial x}+\dfrac{\partial^2 C}{\partial x^2}=0 \\ x=0,\ C_L=C_0/k \\ x=\infty,\ C_L=C_0 \end{cases} \tag{2.3-33}$$

得解

$$C_L=C_0\left[1+\frac{1-k}{k}\exp\left(-\frac{R}{D_L}x\right)\right] \tag{2.3-34}$$

式（2.3-34）中的$\dfrac{D_L}{R}$是一个决定液相浓度分布曲线的函数，称为特征距离（如图2.3-29所示），其数值为浓度函数 C_L 在 $x=0$ 处的导数（或斜率）与 x 轴的交点至原点的距离。从物理意义上，可以认为$\dfrac{D_L}{R}$表示溶质堆积层的厚度和浓度曲线 C_L 的陡峭程度。当特征距离$\dfrac{D_L}{R}$减小时，表示界面附近液相溶质扩散阻力大，溶质富集严重，C_L 浓度曲线陡峭；反之，则扩散阻力小，溶质富集减弱，浓度曲线平缓。

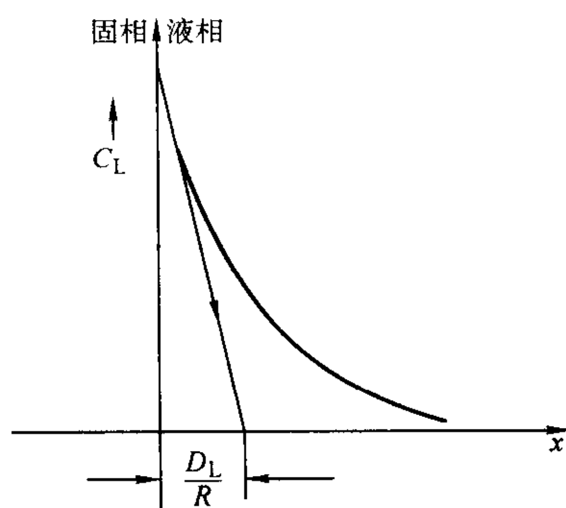


图2.3-29 特征距离示意图

4.4 固相中无扩散，液相中有限扩散但有对流的凝固

在这种条件下，虽然对流（或强烈地搅拌）打乱了液相中溶质按指数函数扩散分布的规律。但由于液体的黏性，在界面附近总会保持一薄层不受对流影响的液体。也就是说，界面前沿保留了一个厚度为 δ 的扩散层（溶质堆积层），而其余液体的浓度则视熔池的长短有如下两种情况，即如果熔池很长，则当 $x=\delta$ 时，有 $C_L=C_0$，即扩散层 δ 外的液相浓度仍为原始成分 C_0；而当熔池较短时，$x=\delta$，有 $C_L=C_b$，且 $C_b>C_0$，并随界面推进浓度逐渐增加（如图2.3-30所示）。

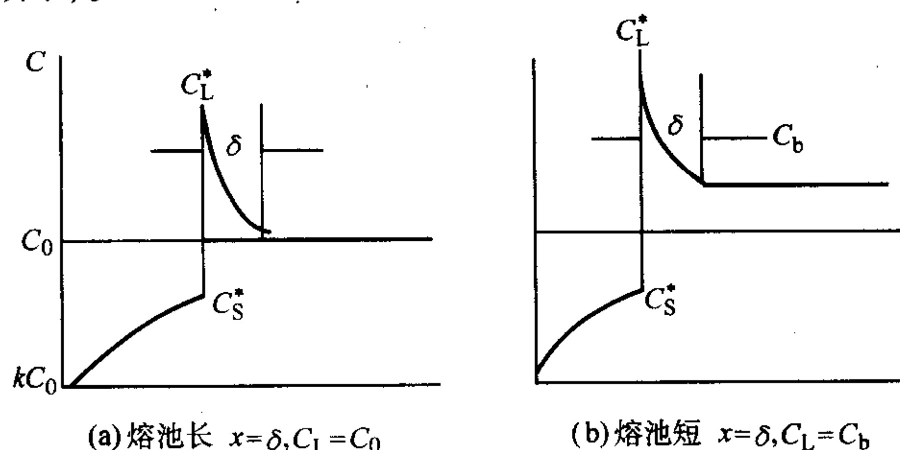


图2.3-30 单相合金固相不扩散、液相有限扩散有对流时固液界面浓度分布

对于熔池较长时的情况，扩散层内的界面固液相浓度和仍可用下式求出

$$\begin{cases} R\dfrac{\partial C}{\partial x}+D_L\dfrac{\partial^2 C}{\partial x^2}=0 \\ x=\delta,\ C_L=C_0 \\ x=0,\ C_L=C_L^* \end{cases}$$

其解为：

$$\frac{C_0-C_S^*}{C_L^*-C_S^*}=\exp\left(-\frac{R\delta}{D_L}\right) \tag{2.3-35}$$

式中，C_S^* 和 C_L^* 均为界面瞬时浓度，他们随界面移动而变化，但始终保持 $C_S^*=kC_L^*$ 的平衡关系。

若熔池有限，液相主体浓度不能不受到固相析出溶质数量的影响而发生变化，但此时液相中浓度方程仍可用下式求解

$$\begin{cases} R\dfrac{\partial C}{\partial x}+D_L\dfrac{\partial^2 C}{\partial x^2}=0 \\ x=\delta,\ C_L=C_b \\ x=0,\ C_L=C_L^* \end{cases}$$

其解为

$$\frac{C_b-C_S^*}{C_L^*-C_S^*}=\exp\left(-\frac{R\delta}{D_L}\right) \tag{2.3-36}$$

注意到 $C_S^*=kC_L^*$，上式可写为

$$C_S^*=\frac{kC_b}{k+(1-k)\exp\left(-\dfrac{R}{D}\delta\right)} \tag{2.3-37}$$

或者为

$$C_S^*=k'C_L^* \tag{2.3-38}$$

k'称为“有效分配系数”而不同于正常分配系数 k。它的意义是在实际应用中可以将固相前沿的瞬时浓度和液相主体浓度联系起来，以便于近似地计算实际铸件的浓度分布。

k'实际上是对流强度的参数，由式（2.3-37）和式（2.3-38）可知：当扩散层 δ 较小时，$\dfrac{R}{D_L}\delta\to 0$，即液相内扩散速度大，或生长速度较小，可得 $k'=k$，即表示充分对流与充分扩散等效，或对流作用使液相接近完全扩散状况。正如前述此时可按固相不扩散、液相完全扩散方式处理生长过程中的界面固液两相浓度（如图2.3-27）。即我们可近似地将液相主体浓度 C_b 作为全部液相的浓度，建立以下溶质平衡方程式

$$(C_L^*-C_S^*)\mathrm{d}f_s=(1-f_s)\mathrm{d}C_b \tag{2.3-39}$$

并可求出界面处固相浓度和液相浓度。

$$C_b=C_0f_L^{(k'-1)} \tag{2.3-40}$$

$$C_S^*=k'C_0\ (1-f_s)^{(k'-1)} \tag{2.3-41}$$

当扩散层 δ 较大时，$\dfrac{R}{D_L}\delta\to\infty$，即液相内扩散速度过小，生长速度较大，可得 $k'=1$，此时 $C_S^*=C_b$，即对流影响非常小，生长界面固液两相浓度可按液相有限扩散无对流方式处理，生长过程中有成分稳定阶段（如图2.3-28）。

5 金属凝固过程中的“成分过冷”现象及胞晶形态

5.1 “成分过冷”现象及界面生长稳定性判据

实际生产的单相合金在凝固成形过程中，界面液相溶质原子的扩散会受到限制，因此界面液相中溶质原子总会产生不同程度的富集，并对界面生长的稳定性，凝固后的晶体形貌以及成份偏析状况产生重要影响。

当液体中的溶质有限扩散时，正如前分析，单相合金凝固前沿（固液界面液相一侧）将出现一个厚度为 δ 的溶质富集层。在这个层中，溶质的浓度高于主体浓度。以分配系

数为 k 的某单相合金为例。假定合金的液相线和固相线都为直线，其斜率分别为 m_L 和 m_S，固液两相的平衡分配系数 $k=C_S/C_L$ 为常数，以 C_L 代表界面液相富集层中相应的浓度，以 T_L 代表平衡状态图上随浓度变化着的液相线温度，根据简单几何关系（见图 2.3-31）有

$$T_L = T_0 + m_L C_L \tag{2.3-42}$$

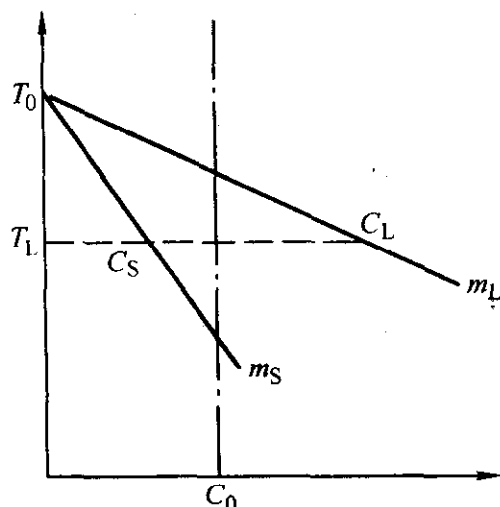


图 2.3-31　单相合金相区简图

将固液界面液相扩散层内浓度表达式（2.3-24）代入式（2.3-31）得

$$T_L = T_0 + m_L C_0\left[1+\frac{1-k}{k}\exp\left(-\frac{R}{D_L}x\right)\right] \tag{2.3-43}$$

T_L 即为生长界面液相侧的理论液相线温度（或称空间液相线温度）。T_L 曲线所代表的并不是界面处结晶液体中的实际温度分布，而仅仅是依据平衡状态图规定的与液相中 C_L 浓度相对应的开始结晶温度（液相线温度），如图 2.3-32 所示。或者说，是在凝固过程中由于界面前沿液相中溶质富集，所引起的液相平衡凝固温度的降低。

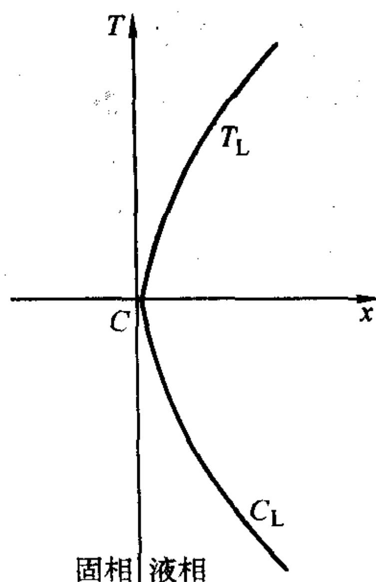


图 2.3-32　单相合金界面液相溶质富集与理论液相线温度的关系

求生长过程中固液界面液相侧的温度梯度，得

$$\left.\frac{dT_L}{dx}\right|_{x=0} = G_L = \frac{-m_L C_0 R\ (1-k)}{kD_L} \tag{2.3-44}$$

式（2.3-44）即为单相合金生长界面稳定性判据，G_L 为空间液相线 T_L 在界面处（$x=0$）的斜率。其物理意义是，当实际界面温度梯度 G_1 大于或等于 G_L，即有 $G_1 \geqslant G_L = \frac{-m_L C_0\ (1-k)}{kD_L}$ 时，界面前沿的生长是稳定的，以平界面推进，当实际温度梯度 $G_2 < G_L = \frac{-m_L C_0\ (1-k)}{kD_L}$ 时，则远离界面的部分液相区域处于结晶过冷状态（见图 2.3-33）中阴影区），界面生长处于不稳定状态，并随过冷度增大，平界面生长逐渐转变胞晶生长或枝晶生长，并使结晶生长有可能在液体内部发生。

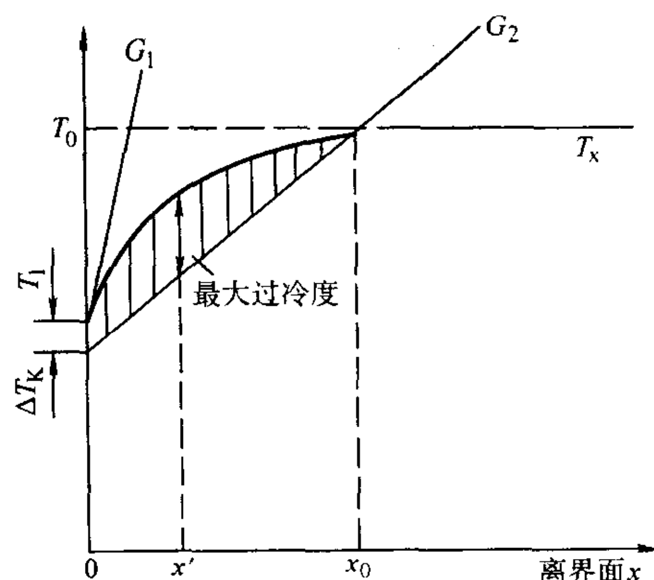


图 2.3-33　生长界面成份过冷的形成

ΔT_K 为界面结晶时的结晶过冷度。

图 2.3-34 是在成分过冷情况下，采用倾倒法显示的 Sn-Pb合金（0.95% Pb）固液界面胞晶形成和发展的组织变化。

“成分过冷”是液态金属成形中的普遍现象，对结晶过程、组织形貌以及最终性能有显著影响。

如果将某成分合金的结晶温度区间 $\Delta T = T_L - T_S$ 引入成份过冷判据，则成分过冷判据也可以写成另一种表达方式

$$G_L \geqslant \frac{\Delta T}{D_L/R} \tag{2.3-45}$$

其物理意义是该合金结晶生长过程中，界面液相侧浓度曲线 C_L 的特征距离 $\frac{D_L}{R}$ 上的实际温升 ΔT^* 大于或等于该合金结晶温度区间 ΔT 时，生长界面稳定，呈平面晶生长，反之则呈胞晶生长方式。或者说界面液相侧实际温度梯度 G_L^* 大于或等于该合金温度区间 ΔT 与特征距离的比值时，界面生长稳定，见图 2.3-35。

由式（2.3-44）可知影响成分过冷的因素有两种，一种为合金本身性质的影响，如平衡分配系数 k，液相线斜率 m_L，该合金的原始浓度 C_0 及液相扩散能力 D_L 等；另一种为工艺因素如界面温度梯度 G_L、界面推进速度 R 等。

5.2　“成分过冷”对组织形成的影响

如前述，固液生长界面的“成分过冷”大小决定着界面的生长过程、生长方式和最后晶体的形状。按照“成分过冷”的大小不同，界面的基本生长方式可分为以下三种：

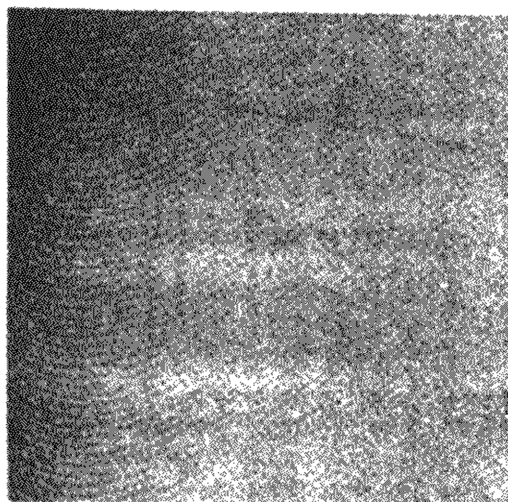
1）平面生长（planer interface growth）；
2）胞晶生长（cellular interfere growth）；
3）枝晶生长（dendrictic interface growth）。

5.2.1　平面生长

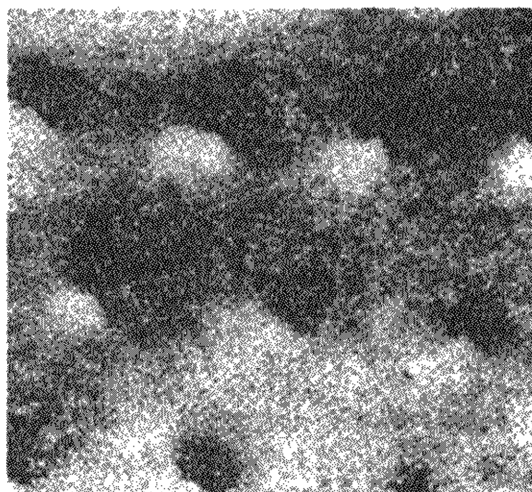
界面前沿液相的实际温度梯度 G_1^* 大于或等于界面的平衡结晶温度梯度 G_1，即满足“成分过冷”判据 $\frac{G_L}{R} \geqslant \frac{mC_0\ (1-k)}{D_L k}$，界面没有大于 ΔT_k 的“成分过冷”，相界面始终保持平面生长方式，如固相偶然出现突出生长部分，即会隐于过热的液体中被重熔，并使界面推进速度减缓（见图 2.3-36）。如果整个凝固过程都保持上述条件，则每个晶体将平行向内伸展成一个个条状的晶体。如果开始时是一个晶体，凝固时此晶体可逐渐向液体内生长成为一个大的单晶体。由于界面平整，液固转变时的体积收缩可由液体补充，没有晶间疏松。

5.2.2　胞晶生长

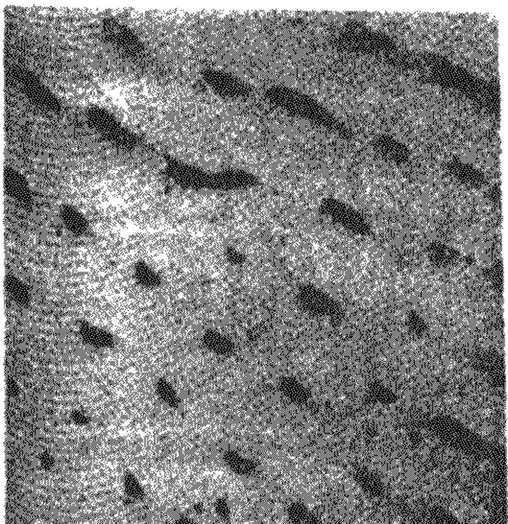
界面液相中的温度梯度为 G_2 时，出现成分过冷（见图



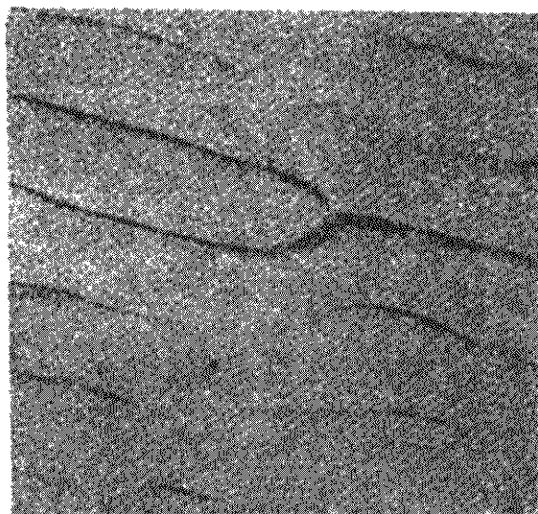
(a)无成分过冷，平面界



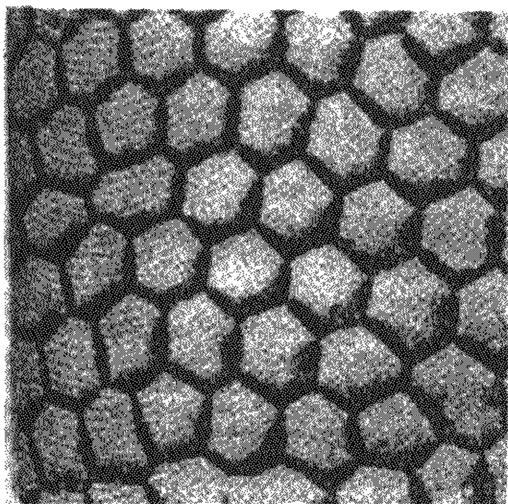
(b)“成分过冷”很小，界面上出现凹坑



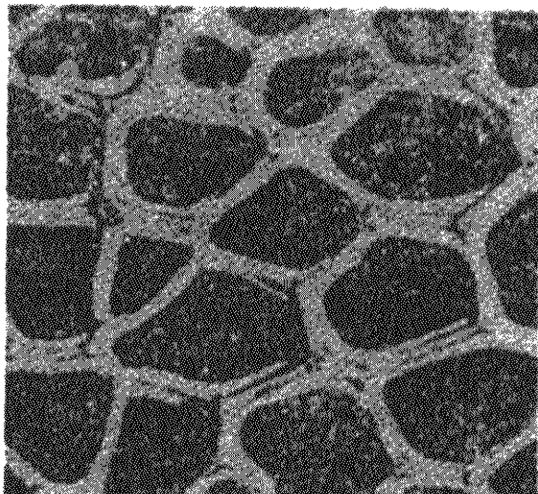
(c)“成分过冷”稍大，凹坑增加并趋于连接



(d)“成分过冷”继续增大，凹坑连接成为沟槽



(e)“成分过冷”增大，形成规则的六角形胞晶



(f)“成分过冷”再增大，胞晶变得不规则

图 2.3-34 “成分过冷”增加对胞晶形成的影响

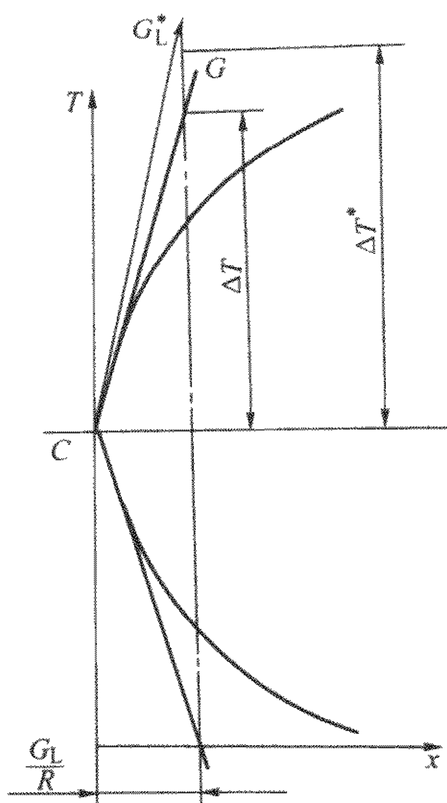


图 2.3-35 单相合金凝固前沿界温度梯度与结晶区间 ΔT 的关系

2.3-37)，平的界面开始不稳定，产生许多凸起伸入成分过冷区内，由于成分过冷区窄，不会产生侧向分枝，晶体也以条状向液体内生长，由于生长突出部分排出的溶质向周围扩散，侧面溶质富集较高，熔点降低，形成富集低熔点物的凹坑沟槽，如将界面液体迅速倾出，可看到胞状结构（如图 2.3-34c、d）。

当成分过冷较小时，在固－液界面上出现的凸起和溶质富集的凹坑则通过沟槽联结在一起，形成规则或不规则的胞状晶（见图 2.3-34 e、f）。

由本条件成分过冷判据 $\frac{G}{R} \geqslant \frac{mC_0\ (1-k)}{D_k}$ 可以得知，影响胞状生长的因素为溶质的存在及 G/R 值的减小。在同样 G/R 值下增加溶质含量，可使生长界面由平面转变为胞状。同样，在一定的 C_0 时，随 G/R 值的减小，也发生从平面向胞状生长的情况。

当成分过冷增大时，胞晶可突出伸入液体至远，其侧向的固液界面也会出现短小分枝，胞状晶转变为胞状枝晶（见图 2.3-38）。它们垂直于固－液界面生长，与热流方向相反而与晶向无关，如生长速度增大，晶体学的影响逐渐发生作

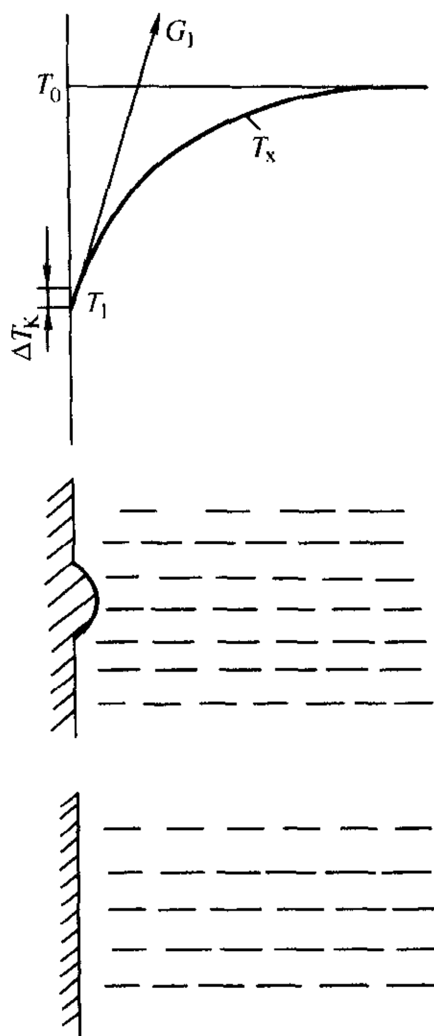


图 2.3-36 平面生长示意图

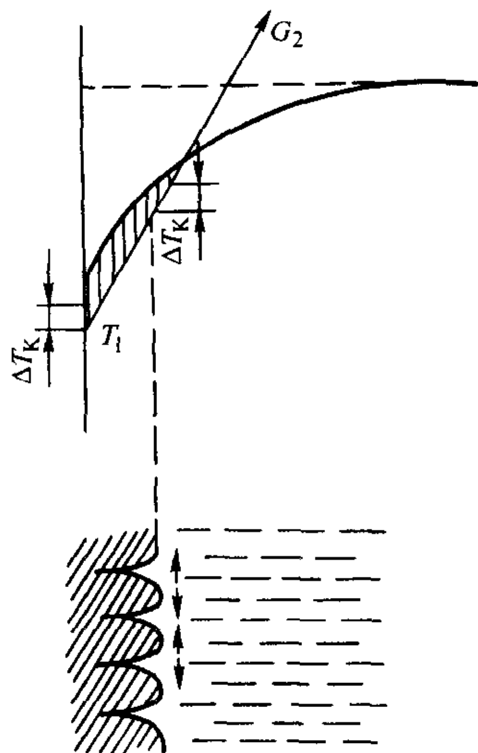
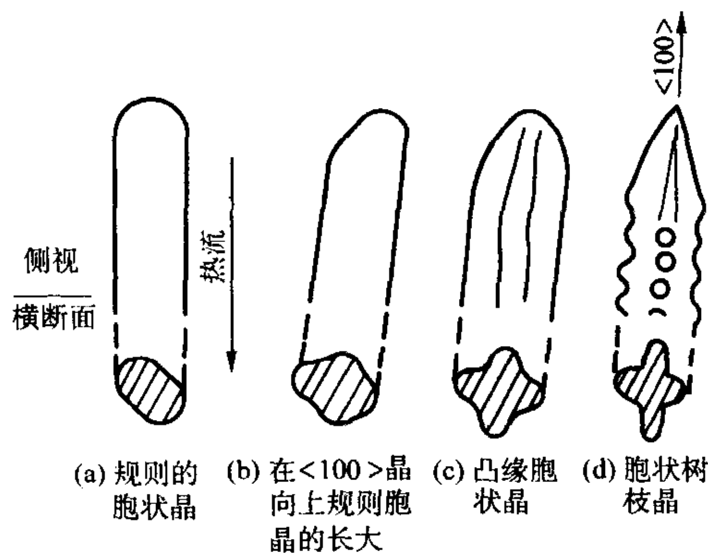


图 2.3-37 胞状生长

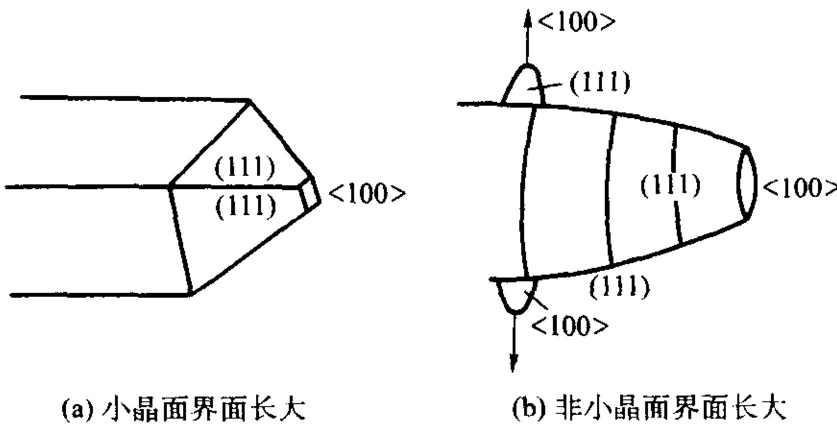


(a) 规则的胞状晶 (b) 在<100>晶向上规则胞晶的长大 (c) 凸缘胞状晶 (d) 胞状树枝晶

图 2.3-38 随生长速度增加胞状晶转变为胞状树枝晶

用，胞晶的生长方向转向优先的结晶生长方向（如立方晶体为 < 100 > 向)，胞晶的横断面出现凸缘结构，或出现锯齿状二次枝晶臂。

可以把胞状树枝晶的主干称为一次臂，由于它是由胞晶发展而来，因此，许多一次晶臂来自一个晶粒。对于溶质少或凝固温度区间窄的合金，其胞状树枝晶形貌为在主干上长出短而密的二次枝晶臂，而大多数合金在较大冷却速度时，有高度的分枝形态，即在二次枝晶臂上还会长出三次枝晶臂。通常将这种一次臂与热流方向平行且高度分枝的晶体称为柱状树枝晶，它的生长方向是密排晶形成的维体的主轴方向（如图 2.3-39)。



(a) 小晶面界面长大 (b) 非小晶面界面长大

图 2.3-39 胞状树枝晶的长大方向

5.2.3 枝晶生长

当成分过冷区域较大，生长速度较快时，胞晶生长转为枝晶生长。枝晶生长有两种类型：①柱状树枝晶；②等轴树枝晶。

柱状树枝晶与胞状树枝晶在本质上是相同的。可以说，柱状树枝晶由胞状树枝晶发展而来。柱状树枝晶生长时，界面上的突出部能进入液体较深处形成一次分枝。一次分枝的侧面也有成分过冷，可形成二次分枝，如成分过冷区足够大，在二次分枝上可出现三次分枝。柱状树枝晶的生长方向是依赖晶体学特征择优生长，表 2.3-12 为不同晶系的合金择优生长晶向。

表 2.3-12 枝晶择优生长晶向

晶系	面心立方	体心立方	体心正方	密排立方	
生长晶向	< 100 >	< 100 >	< 110 >	< $10\bar{1}0$ >	< 0001 >
举例	Al	δ - Fe	Sn	H_2O（雪）	$CO_{17}Sm_2$

还有一种情况，在成分过冷区，当成分过冷极大值 ΔT_{max} 大于某些外来质点非均质形核所需要的过冷度 ΔT^* 时，可以不依赖于型壁而在液态成分过冷区内形核生长（如图 2.3-40)。

晶体生长时，结晶潜热可向四周散发，受型壁散热方向的影响较小，故枝晶在各个方向生长均匀，没有明显的单向延伸的分枝，所形成晶粒成颗粒状，内部可显示出各方等轴的枝晶组织，称为等轴晶。

柱状树枝晶和等轴树枝晶晶体学最优长大方向是一致的，但是，它们的组织形貌却有所不同，主要受到溶质扩散和温度扩散的影响。图 2.3-41 为不同溶质含量的合金在不同界面液相温度梯度情况下的生长形貌。

5.3 细胞晶间的溶质分配

无论是柱状树枝晶还是等轴树枝晶，都会在凝固的固相中留下严重的溶质偏析。这种偏析对铸件性能危害较大，可造成材料的冲击韧性，塑性及耐蚀性的降低。分析细胞晶凝固中的溶质再分配问题，有助于了解和控制溶质偏析倾向。将柱状树枝晶视作胞晶（见图 2.3-42)，可以定性地预测胞晶（树枝晶）的微观偏析情况。

为使问题简化，仍作出以下假设：①假设固相无扩散；②液态金属 x 向浓度梯度为常数，即$\dfrac{\partial C_L}{\partial x}=G/m_L$，且$\dfrac{\partial C_L}{\partial y}=0$

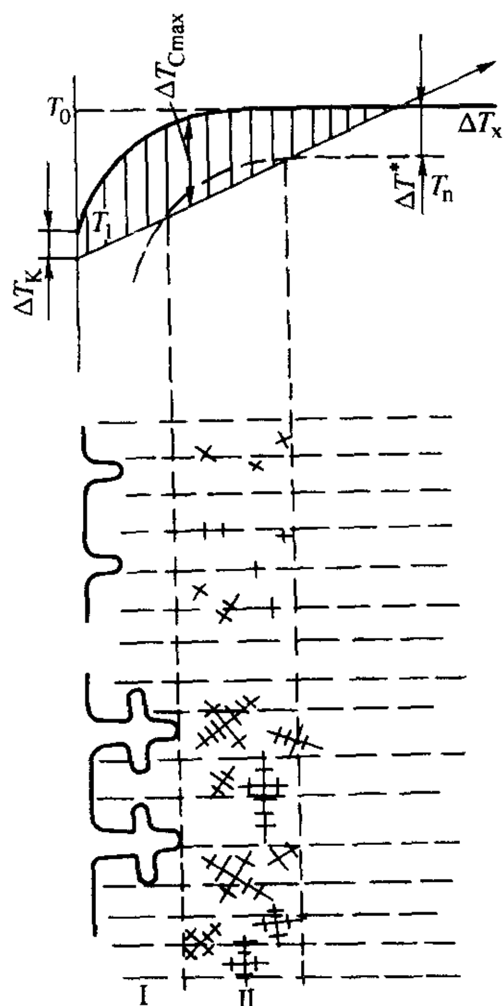


图 2.3-40　内生生长－界面前方等轴晶形成
Ⅰ—枝晶生长区；Ⅱ—等轴晶形成区

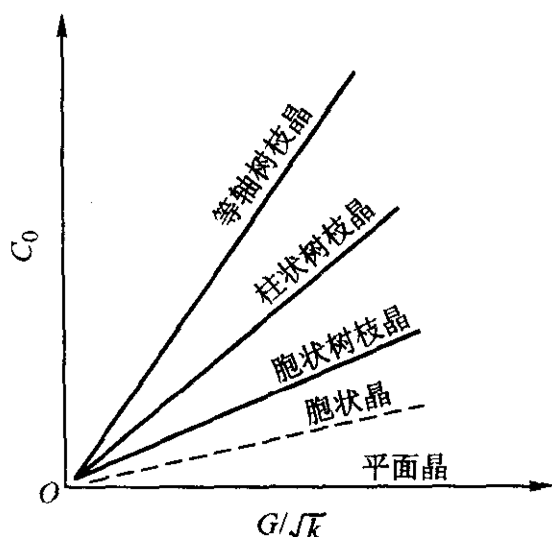


图 2.3-41　$\frac{G}{\sqrt{k}}$和 C_0 对固溶体晶体形貌的影响

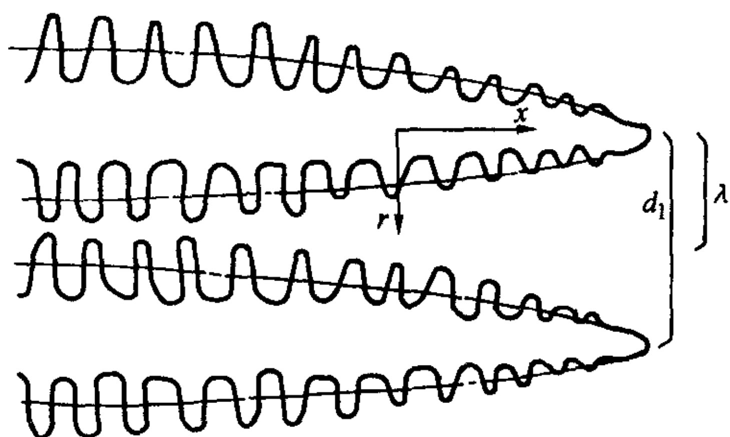


图 2.3-42　将柱状树枝晶视作胞晶示意图

(胞晶横截面较小，Y 向没有浓度梯度)，③等温面是平坦的，且垂直于生长界面。在胞晶中设有微元体 dxl（见图 2.3-43)，且在微元体生长过程中存在固液两相体积分量 f_S、f_L，$f_S+f_L=1$。

当胞晶生长时，即微元体中 f_S 增加时，微元体之溶质浓度变化可达成以下平衡，即单位时间内微元体平均浓度的增量$\left(\frac{dC_{av}}{dt}\right)$等于析出固相时所排出的溶质浓度$\left[\frac{df_L}{dt}(C_L-C_S)=\frac{df}{dt}C_L(1-k)\right]$和单位时间内液体 f_L 体积中溶质量的增加 $\left(f_L\frac{\partial C_L}{\partial t}\right)$ 的代数和

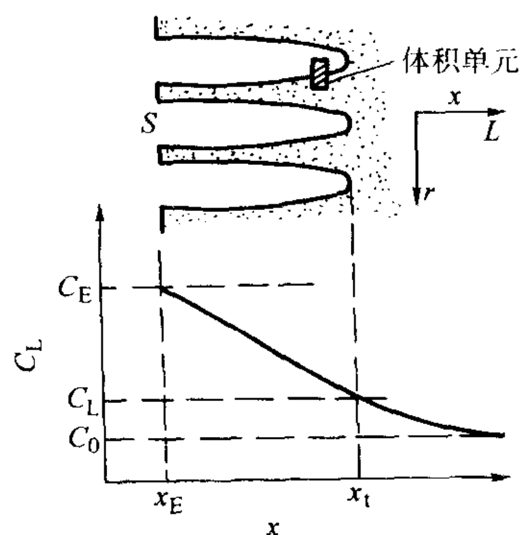


图 2.3-43　胞晶间溶质分布及微元体示意图

$$\frac{dC_{av}}{dt}=\frac{df_L}{dt}C_L(1-k)+f_L\frac{dC_L}{dt}\qquad(2.3\text{-}46)$$

引入边界条件 $f_L=1$ 时，$C_L=C_t=(1-a)C_0$

式中，C_t 为胞晶尖端处的浓度；$a=\frac{D_L}{m_L}\times\frac{G}{Rc_0}$。

解上述方程得：

$$C_L=C_0\left[\frac{a}{k-1}+\left(1-\frac{ak}{k-1}\right)f_L^{(k-1)}\right]\qquad(2.3\text{-}47)$$

这就是细胞晶间（或枝状晶内）无固相扩散时，固液两相的浓度分布和液态体积分量 f_L 消长的关系，f_L 可认为是从胞晶间液体中线到胞晶边界的相对距离，f_L 增加，C_L 逐步下降。利用该式，可描绘胞晶的大致形貌和胞晶间的溶度富集程度。

5.4　枝晶臂间距

生产试验证实，枝晶臂间距愈小，其结晶组织愈细，力学性能也愈高，元素偏析和夹杂物分布的情况也大为改善，因此了解合金生长过程中的枝晶臂间距大小，以及影响间距大小的因素有重要意义。

在有关枝晶臂的影响中，以一次枝晶臂间距和二次枝晶臂间距对合金组织和性能影响最为显著。

研究结果指出，一次枝晶臂间距 d_1 和二次枝晶臂间距 d_2 可分别用下式表达：

在平界面生长条件一次枝晶间距

$$d_1=\frac{2D}{R}=2\left[\frac{mD_LC_o(k-1)}{R}\right]^{1/2}\qquad(2.3\text{-}48)$$

在非平面生长条件

$$d_1=2\varepsilon\left[\frac{mD_LC_o(k-1)}{R}\right]^{1/2}\qquad(2.3\text{-}49)$$

上两式也可简化为

$$d_1=a\left(\frac{1}{G_LR}\right)^{1/2}\qquad(2.3\text{-}50)$$

两二次枝晶臂间距也可简化为下式

$$d_2=b\left(\frac{1}{G_LR}\right)^{1/3}\qquad(2.3\text{-}51)$$

式中，C_o 为合金原始浓度；m 为液相线斜率；D_L 为液相合金扩散系数；R 为界面生长速度；G 为界面液相温度梯度；a、b 为与合金种类有关的常数；ε 为晶体生长形态的修正系数。

由式（2.3-50）和式（2.3-51）可看出，在合金成分一定时，一定的生长条件（生长速度 R 和界面液相温度梯度 G 不变），一次枝晶臂及二次臂间距是稳定的，凝固速度 R 及

温度梯度G的增加，均可使臂间距减小而细化。一次臂间距在晶体生长过程中，由于生长条件的改变会自动调整，其调整的机制主要以分枝、淹没方式进行。对胞晶来说分枝是在胞晶端部产生与原胞晶轴向相近的分枝，对枝晶来说是在二次枝晶上出现三次臂，如图2.3-44所示。所谓淹没方式是在凝固过程中，那些生长较快的胞晶排出的溶质，一部分排向生长慢的胞晶顶端，使之生长受阻，最后被"淹没"。

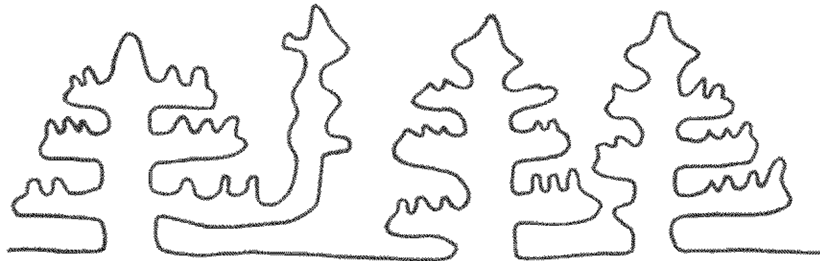


图2.3-44 树枝晶一次臂间依靠分枝调整示意图

实验研究发现，在凝固过程中早期凝固出现的二次枝晶臂间距在凝固后期有粗化倾向。Kattamis等人提出了三种粗化模式。如图2.3-45所示。

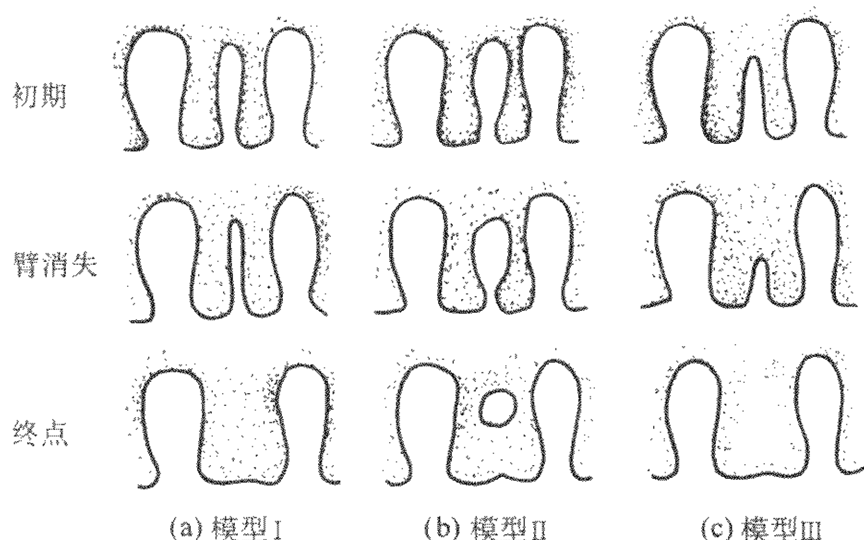


图2.3-45 二次臂粗化模型

模型Ⅰ的机理是细的二次臂在生长过程中溶解，导致二次臂间距增大；模型Ⅱ的机理是，对于根部比较细的二次臂，根部逐渐溶解，残留下的固相游离到液相中，使二次臂间距增大；模型Ⅲ的机理是，细的二次臂的尖端首先形成溶解，导致二次臂从端部向根部逐渐溶解，最后该二次臂完全消失，使二次臂间距增加。

粗化现象产生的原因主要与二次枝晶臂端部曲率半径不同（或晶臂粗细不同），使端部熔点发生变化有关，即枝晶端部曲率半径小者熔点低而易于熔解。

近年来的研究发现，除冷却速度外，溶质浓度、平衡分配系数等对二次枝晶臂间距也有不容忽视的影响；二次枝晶臂间距对力学性能的影响相关性甚密，甚至比晶粒度的影响还要明显，这是因为晶内偏析、缩松及夹杂物的分布随二次枝晶臂间距的减小而趋于均匀，从而减轻了对性能的不利影响。但需指出的是，二次枝晶细化的良好作用只有在铸件致密的情况下才能获得。

6 共晶凝固及包晶凝固

6.1 共晶组织分类

在结晶过程中，同时由液态合金中析出两个以上固相的过程称为共晶反应。工业上应用广泛的铸铁和铝硅合金材料的组织即主要由共晶相组成。由于化学成分及凝固条件的变化，相种类相同的共晶合金可以形成多种不同组织形态，使性能出现较大差异。按照熔融熵值的大小，或固液生长界面粗糙程度，共晶合金可分为规则共晶（非小平面共晶 non-facet-nonfacet）和非规则共晶（非小平面－小平面共晶 non-facet-facet）。

规则共晶组织多由金属－金属相或金属－金属间化合物相组成，如Al－$CuAl_2$、Sn－Zn、Pb－Sn、Ag－Cu等。它们的组织形态可分为层片状和棒状（或纤维状）（见图2.3-46）。这类共晶的固液生长界面在原子尺度上是粗糙的，即固液界面不是特定的晶面，决定它们长大的影响因素主要是热流方向和两组元在液相中的扩散能力。这类共晶的两相在长大过程中虽然会有一相先析出或领先生长，但由于一相的长大受制于另一相排出溶质的影响，所以在共晶凝固中，两相近似于并排推进，并垂直于固液界面长大，界面基本保持平面，其等温面也是平直的（见图2.3-47）。

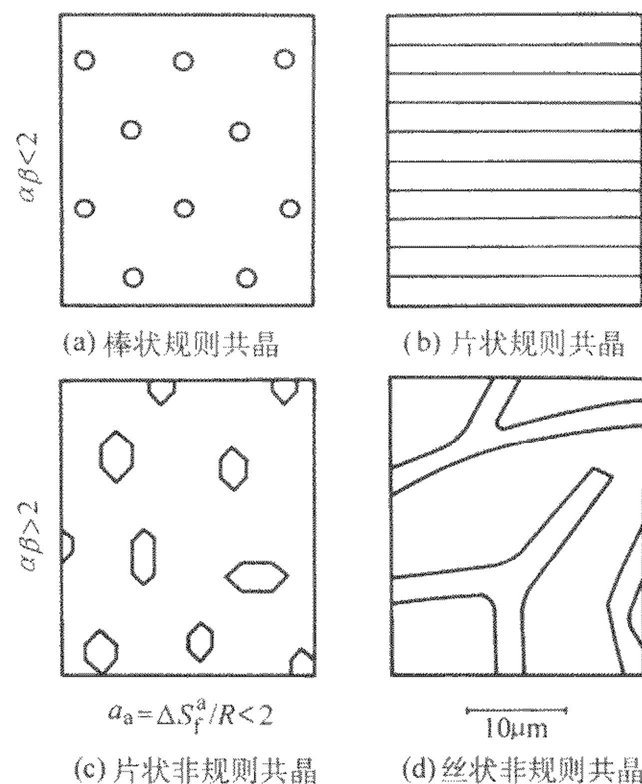


图2.3-46 共晶组织形态的基本模型

非规则共晶多由金属－非金属相组成，如Al－Si合金，Fe－C合金等，它们的组织形态可分为片状与丝状两类（见图2.3-46），由于其小平面相的晶体长大时呈各向异性（受界面能、热传导及最优生长方向等因素影响），其固液界面一般有特定晶面。虽然这类合金在共晶凝固过程中，两相生长仍受界面液相中的溶质原子横向扩散的影响而"协同"长大，但其固液界面的形态是非平面的，而且极不规则（如图2.3-47），等温面也不是平直的。

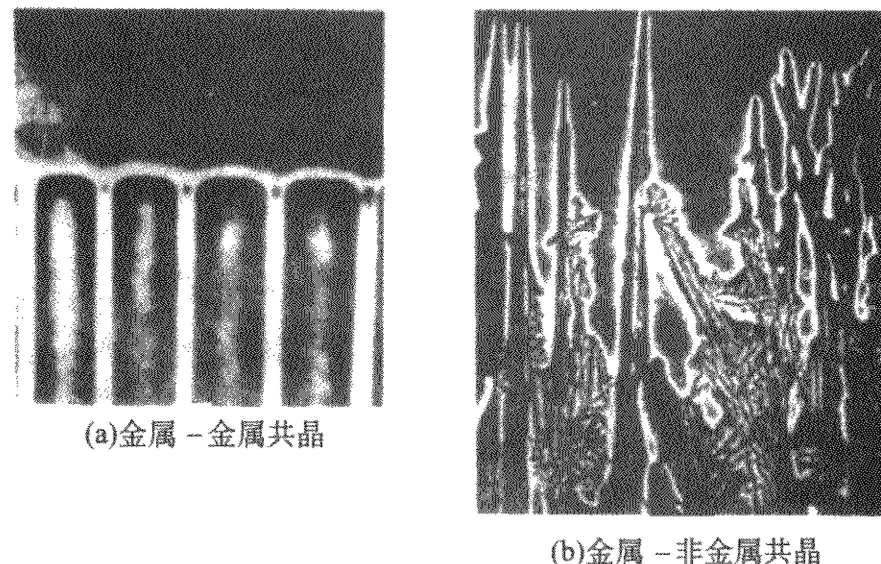


图2.3-47 单向凝固条件下共晶的固－液界面形貌

由于小平面相主要依靠晶体缺陷侧向扩展长大，方向性强，其在共晶体中的体积分数φ_f对共晶形貌有很大的影响。Croker绘制出在长大速度为5 $\mu m \cdot s^{-1}$情况下，φ_f与熔解熵值对共晶显微结构的影响（见图2.3-48）。在溶解熵值$>23J/(mol \cdot K)$的非规则共晶中，$\varphi_f<10\%$时，组织为破断

的层状结构（见图2.3-49）；当共晶体中的小平面相体积分数 φ_f 在10%～20%时（见图2.3-48*D*区），共晶组织为不规则的片状结构，Al－Si及Fe－C合金是这类组织的典型（见图2.3-50）；在熔解熵值较小的*G*区，层片状结构可向丝状结构转变。当共晶中的小平面相体积分数 φ_f 在20%～35%之间时（见图2.3-48*E*区），共晶组织为复杂的规则结构，这种组织由许多胞晶组成，在胞晶内部，其显微结构为在脊椎的周围规则地排列着的一些板条组织所构成（如图2.3-51Al－Ge共晶合金所示）。

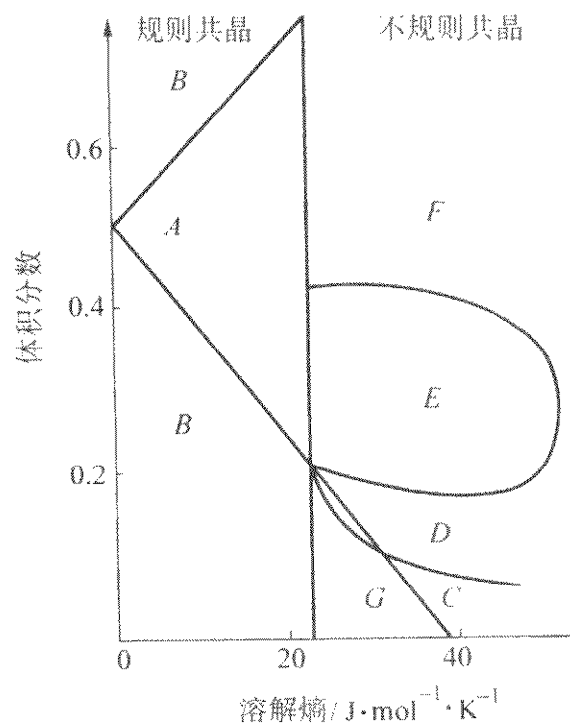


图2.3-48　共晶中小平面相体积百分数与溶解熵对组织形貌的影响

A—规则层片；*B*—规则棒状；*C*—不规则的分枝层片；*D*—紊乱的薄片；*E*—复杂的规则结构；*F*—准规则结构；*G*—不规则的丝状结构

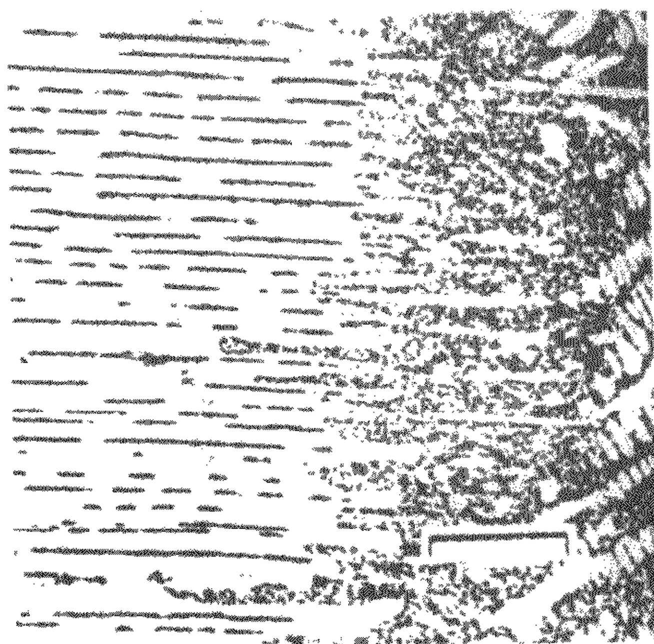


图2.3-49　Pb-Ag合金破断层片结构的淬火界面

当 $\varphi_f>40\%$（图2.3-48*F*区），共晶组织为准规则结构，如图2.3-52所示Fe－Fe_3C共晶合金（莱氏体），它们由非小平面相的板片状或少量棒状组成，而基面则为小平面相，只是小平面相Fe_3C不能按小平面方式而是按非小平面方式长大。在板片与棒状混合存在的组织中，一般板片的扩展方向与热流方向一致，而棒状结构则与板片相垂直。当液相中的温度梯度减少或者长大速度增加时，棒状结构的比例增加，第三组元的存在也会使侧向的棒状结构增加，并促使胞状组织的形成。

6.2　非平衡状态的共晶凝固

在平衡条件下，共晶反应只发生在一个固定成分，而偏离这一成分的合金凝固后都不能获得100%的共晶组织。从热力学观点看，在非平衡凝固条件下，具有共晶型的合金，

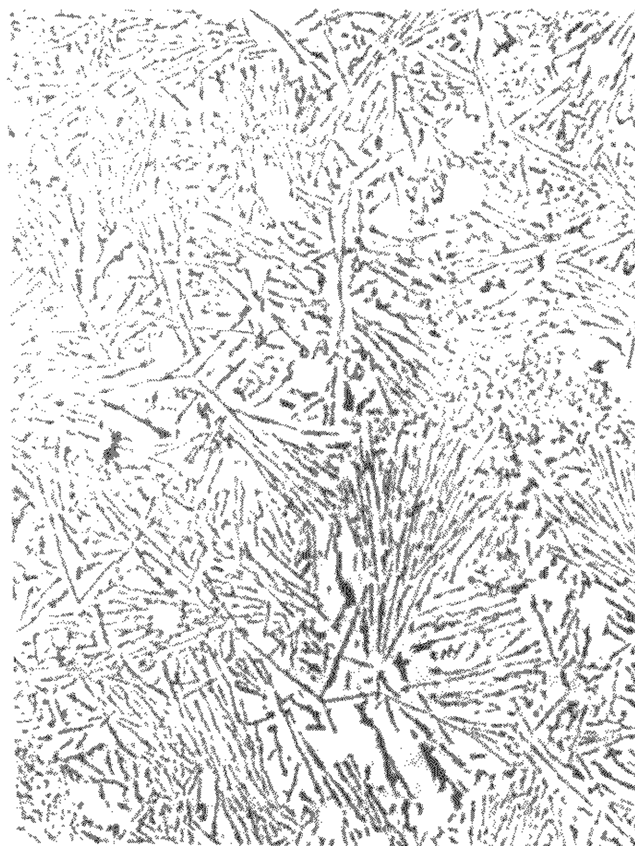


图2.3-50　Al-Si合金不规则的片状结构

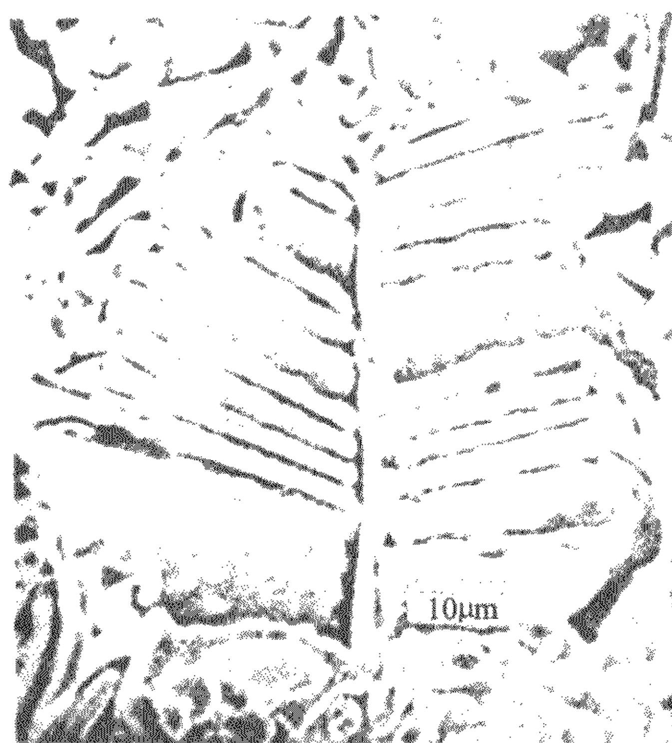


图2.3-51　Al-Ge合金复杂结构的胞晶内微观排列

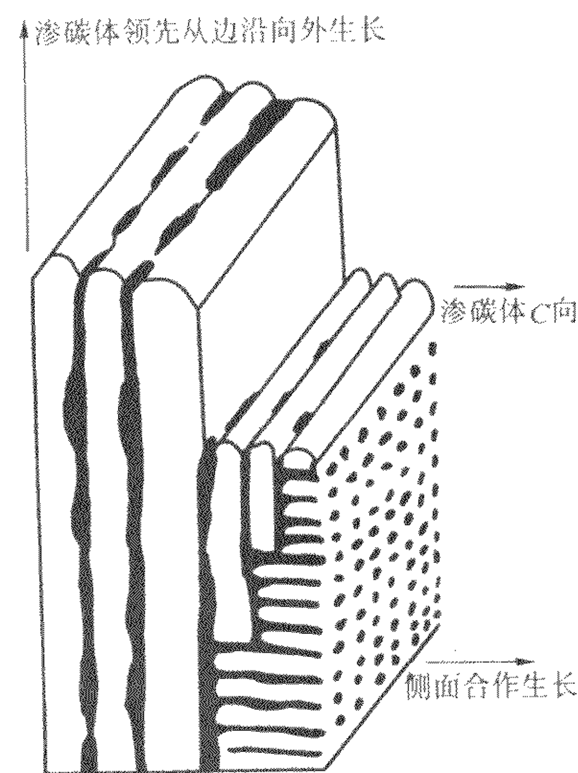


图2.3-52　Fe-Fe_3C的准规则结构（白色为Fe_3C，黑色为奥氏体）

当快冷到两条液相线的延长线所包括的区域内时，即使是非共晶成分的合金，也可以获得 100% 的伪共晶组织（见图 2.3-53），图中的影线部分即是共晶共生区。共生区确定了共晶稳定生长的温度和成分范围，超过这个范围，组织将变为亚共晶或过共晶结构。但是以共晶成分为中心的对称型共生区，只发生在共晶中两相的熔点相近的金属 - 金属共晶系中，其共晶共生区形状也不如图 2.3-53 那样简单。图 2.3-54 中阴影部分即为界面温度梯度 $G>0$ 时的铁砧式对称型金属 - 金属共晶共生区。当晶体生长速度较小时（阴影区上部），为单相凝固可获得平界面共晶组织，凡处于共晶相图上 $C_{\alpha m}-C_{\beta m}$ 之间的成分（$C_{\alpha m}$ 和 $C_{\beta m}$ 分别为平衡相图中共晶反应时 α 和 β 两相的浓度），均可得到共晶组织。随着长大速度增加，共晶组织变为胞状、树枝状，在图中虚线与其延长线所夹的范围为 $G=0$ 时的情况，共晶只能成为等轴晶。

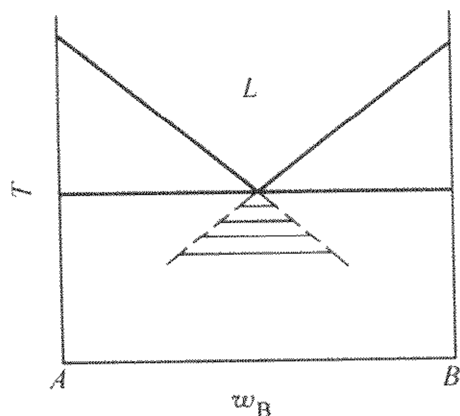


图 2.3-53 从热力学观点考虑的共晶共生区

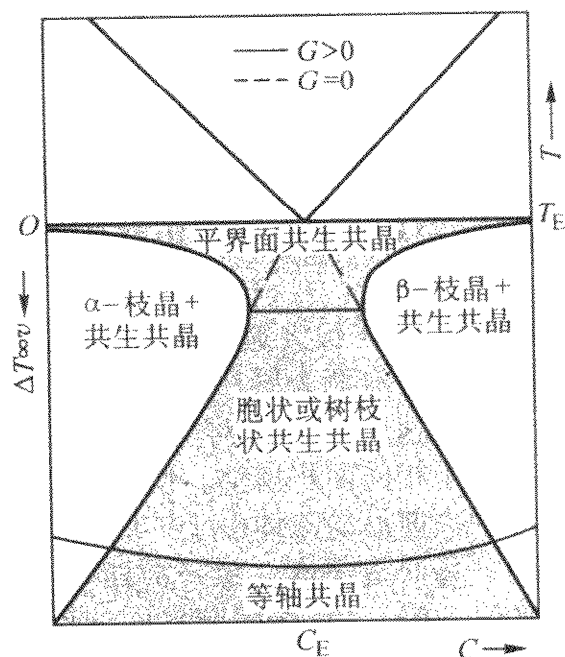


图 2.3-54 金属 - 金属共晶共生区

对于金属 - 非金属共晶，其共生区是非对称型的，相图上的共晶点靠近金属组元一方，共晶共生区偏向非金属一方（如图 2.3-55），由图可知，这类共晶成分合金，在快冷条件下往往得不到共晶组织，其原因是在大的过冷度下，共晶中的小晶面相在特定晶向上的长大及分枝难以进行，金属 - 非金属共晶生长被抑制，且单相的金属相在富非金属组元的合金中形成。

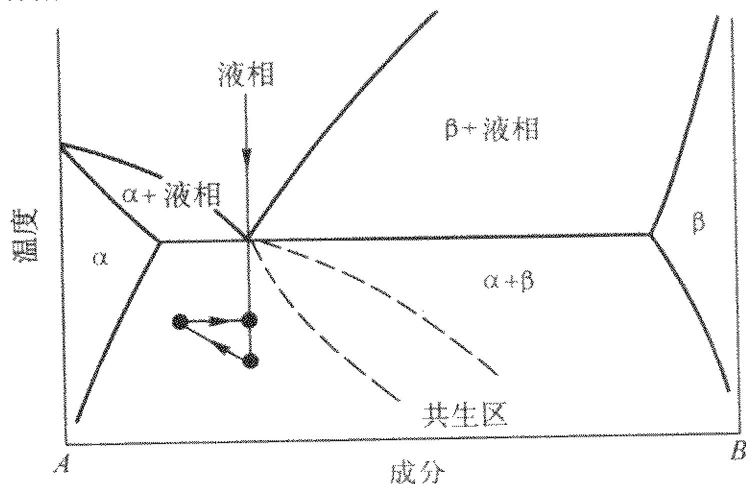


图 2.3-55 金属 - 非金属共晶共生区

6.3 金属 - 金属共晶凝固

6.3.1 层状共晶的生长

（1）层状共晶生长特征

共晶开始时，总有生长领先相或先析出相出现（一般为高熔点物质），然后第二相以它为基底相继析出，共晶结晶过程中，两相交替成长，主要靠搭桥方式使同类相的层片增殖长大（如图 2.3-56），并不需要每一个层片都要单独形核。由于共晶长大速度和四周方向是均一的，因此同一晶核产生的共晶生长界面具有球形生长前沿，并最终长成为球形共晶体。这类共晶体称为共晶团（或共晶晶粒）。在球形共晶团内部两相之间却是层片结构，在生长过程中，在层片间距大小范围固液生长界面大体上是平面的。图 2.3-57 可看出 Pb - Cd 合金几个不同的共晶晶粒凝固终了接触时的组织状况。

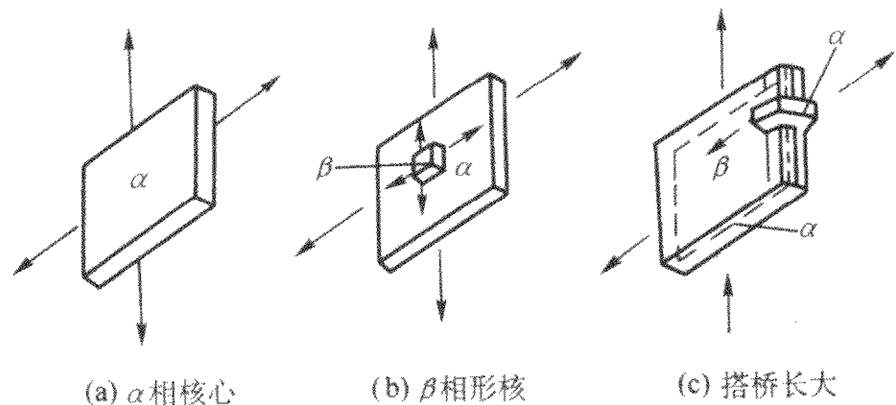


图 2.3-56 层片状共晶搭桥长大过程

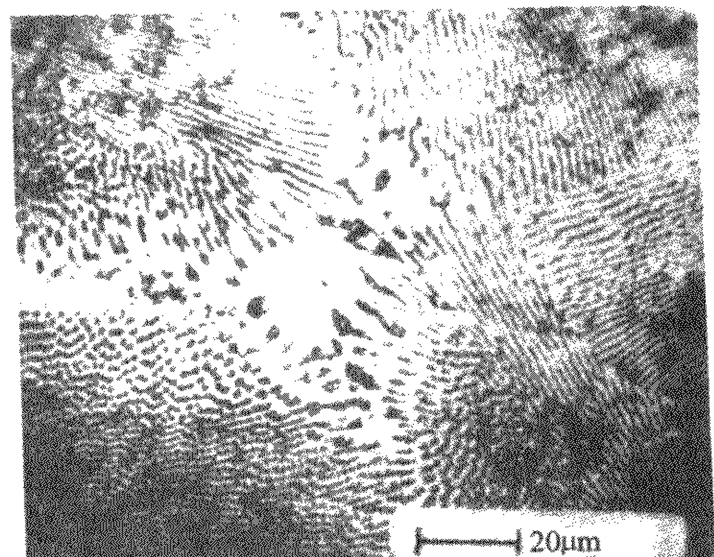


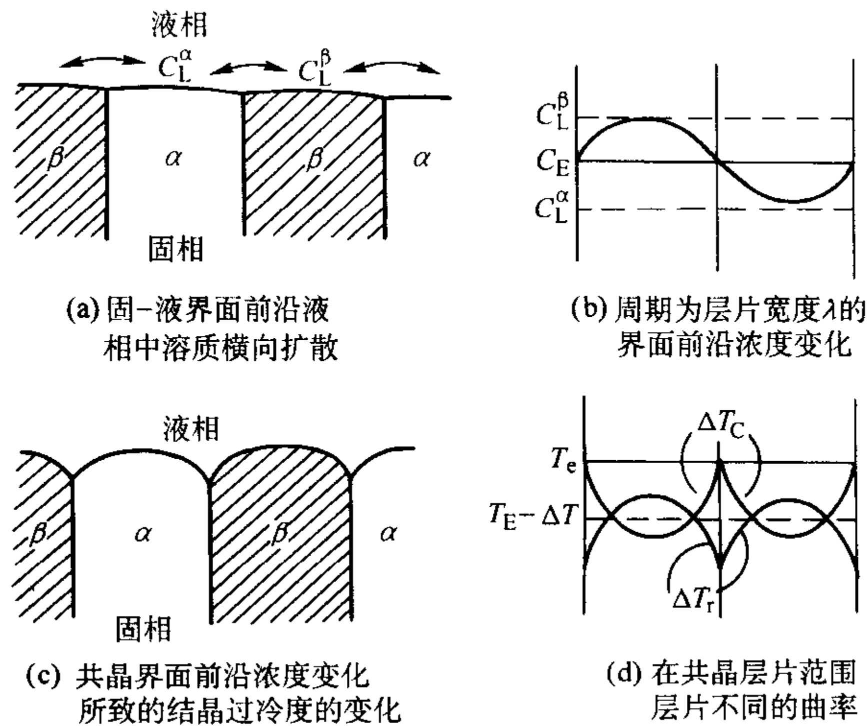
图 2.3-57 Pb-Cd 共晶晶粒

（2）共晶层片生长时固 - 液界面浓度分布

假定由 A 组元和 B 组元构成层片共晶，在共晶凝固期间，φ 相生长界面会排出 B 原子而在 β 相生长前沿则排出 A 原子，因此必在 α、β 生长前沿有溶质富集现象和"成分过冷"抑制界面推进。为使共晶生长得以继续，A 原子需从 β 相前沿横向扩散到 α 相前沿。而 B 原子则从相关前沿横向扩散至 β 相前沿（见图 2.3-58a）。α 相中央的前沿液相由于距离 β 相较远，所以排出 β 原子不如 α 与 β 交界处的前沿那样很快可以扩散掉，因此形成 B 原子的富集程度最大，而在 α 与 β 交界处几乎没有富集，此处界面前沿液相成分等于 CE；同理在 β 相中央前沿液相中，势必也富集了很多 A 原子，$\alpha\beta$ 相交界处 A 原子富集也很少。这种溶质富集现象的存在和横向扩散能力的大小，导致了两相界面浓度的变化和结晶过冷度 ΔT_c 的变化，并影响共晶生长及形态变化。

若假定共晶平衡相图完全对称，即 α、β 两相体积分数相等，并具有相同的平衡分配系数，则界面前沿液相浓度分布可如图 2.3-58b 所示。

在一般凝固条件下，纯共晶合金的固 - 液界面在层片间距大小范围内是平面的和等温的，其结晶过冷度 ΔT 由三个



(a) 固-液界面前沿液相中溶质横向扩散　(b) 周期为层片宽度λ的界面前沿浓度变化

(c) 共晶界面前沿浓度变化所致的结晶过冷度的变化　(d) 在共晶层片范围层片不同的曲率

图 2.3-58　共晶前沿 α、β 两相溶质扩散和浓度分布

单独的过冷度所组成（即由三个因素影响最终结晶过冷度值）。其一是上述溶质在界面前沿液相中局部富集所引起的成分过冷度 ΔT_e（见图 2.3-58c）；其二是由于与液相接触的固相表面曲率变化所造成的过冷度 ΔT_r（见图 2.3-58c）；其三是凝固动力学过程所需要的过冷度 ΔT_k，而 ΔT_k 与前两相比相对较小而忽略，所以固液界面的总结晶过冷度 ΔT 可写为

$$\Delta T=\Delta T_e+\Delta T_r$$

正如前述，在稳定状态下长大时，固液界面是等温的，也即结晶过冷度 ΔT 力求稳定。为满足这一条件，也即为了弥补由于生长前沿浓度的变化所带来的平衡结晶的温度的降低，则要求 α 相与 β 相层片表面具有不同的曲率半径（曲率半径愈大则平衡凝固温度愈高）。因此可以发现，在 α、β 相交界处，成分过冷最小，而层片表面的曲率半径也最小，弯曲最大；而在层片中心接触的液相溶质富集最多，成分过冷最大，层片中心部分的弯曲也最小（见图 2.3-58d）。

共晶凝固生长界面前沿的浓度不均匀，仅在固 - 液界面层片间距范围内发生，而在层片间距范围外，液相成分仍可保持着共晶成分 CE。

因此在层片厚度范围内，距离固 - 液界面愈远成分波动的幅度愈小。假定 x 为距固 - 液界面的距离，C_L 为界面前沿 x、y 处的液相成分，y 为以 α 相中心为原点，横切 α、β 相长大方向的距离；D_L 为溶质在液相中的扩散系数；R 为共晶固 - 液界面向液相中推进的速度，则可定量求出界面液相浓度表达式（如图 2.3-59）

$$C_L=\sum_{n=1}^{\infty}\frac{\lambda R}{(n\pi)^2D_L}(C_{\beta m}-C_{\alpha m})\sin\frac{2n\pi}{\lambda}S_\alpha\cos\frac{2n\pi}{\lambda}y\mathrm{e}^{\frac{-2n\pi}{\lambda}x}+C_E \qquad (2.3\text{-}52)$$

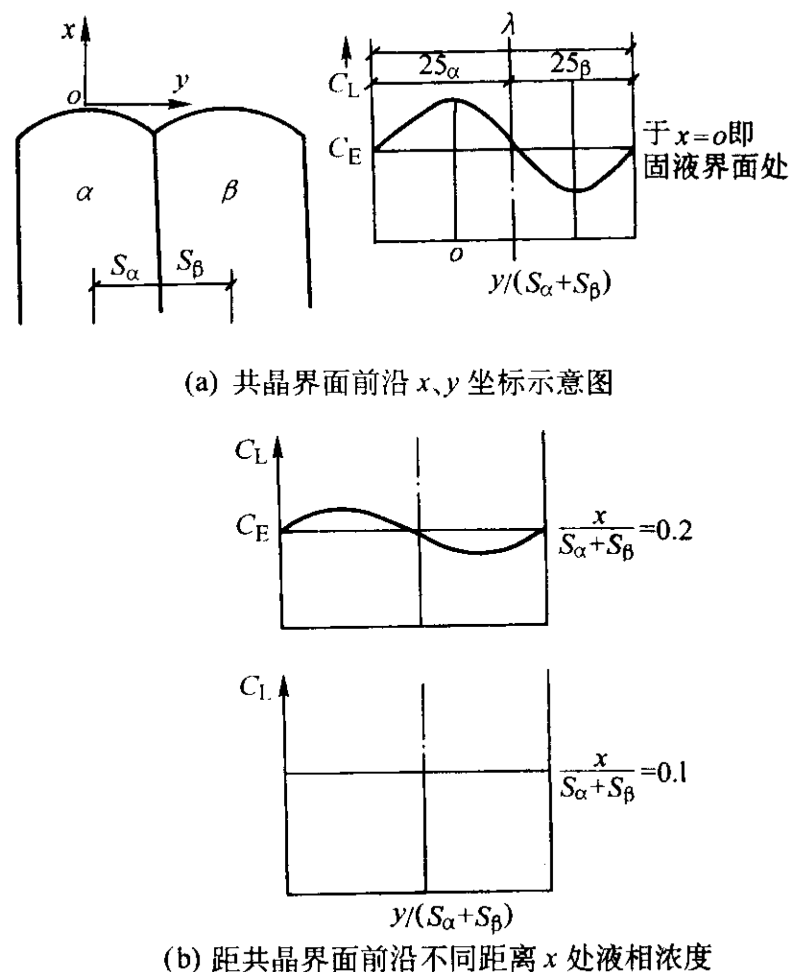
设 $S_\alpha=\lambda/4$，上式可变为

$$C_L=\sum_{n=1}^{\infty}\frac{\lambda R}{(n\pi)^2D_L}(C_{\beta m}-C_{\alpha m})\mathrm{e}^{-\frac{2n\pi}{\lambda}x}\cos\frac{2n\pi}{\lambda}y+C_E \qquad (2.3\text{-}53)$$

由式可看出，在固 - 液界面前沿、溶质富集的程度与 $(C_{\beta m}-C_{\alpha m})$ 成正比，长大速度 R 愈大，或 R 不变，层片间距愈大时，其界面前沿溶质富集愈严重。

(3) 共晶层片间距

所谓层片间距指的是共晶组成相的单元层片宽度的平均值 λ，对于给定的共晶合金成分而言，层片间距 λ 的大小（即弥散度）对合金力学性能有很大影响。共晶结晶过冷度



(a) 共晶界面前沿 x、y 坐标示意图

(b) 距共晶界面前沿不同距离 x 处液相浓度

图 2.3-59　共晶结晶时固 - 液界面前沿在层片厚度距离范围内的成分分布

ΔT，界面生长速度 R 与片间距 λ 之间有如下关系：

$$\Delta T=B_1k\frac{1}{\lambda}+m_L\frac{\lambda R}{\pi^2D_L}(C_{\alpha m}-C_{\beta m})=\frac{B}{\lambda}+AR\lambda \qquad (2.3\text{-}54)$$

式中，$B_1=\frac{\sigma_LT_s}{\Delta H_m\rho_S}$ 为常数；k 为平衡分配系数；m_L 为液相线斜率；$C_{\alpha m}$ 和 $C_{\beta m}$ 分别为平衡相图中共晶反应时 α 和 β 两相的浓度。

从图 2.3-60 可知：在长大速度 R 一定的情况下，除 m 点外，同样的过冷度会有两个层片间距，这在实际上是不可能的，因为一个长大速度 R 与之对应的只有一个层片间距，片间距过小时，由于相间面积增加，使界面能增大；片间距过大时，在片层中央前沿的液体由于扩散距离较远，富集了大量的溶质原子，从而迫使固 - 液界面曲率半径出现负值，形成凹袋（见图 2.3-61），并逐渐向界面的反向延伸，直到产生另一相为止，事实上也就自动调整了层片间距。总之，一个长大速度，只有一个最小过冷度与之对应，图 2.3-61 中的 m 点即为某一长大速度所需要的最小过冷度以及与之对应的一定大小的层片间距。

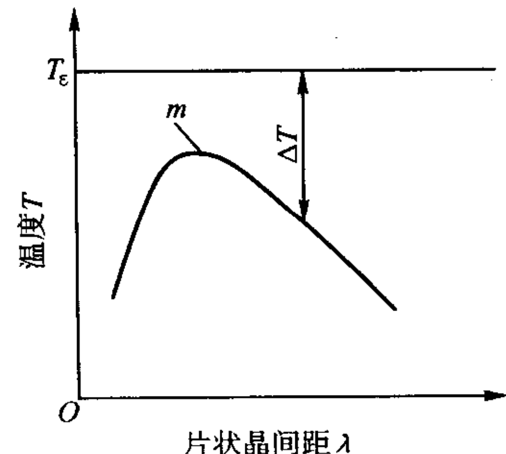


图 2.3-60　长大速度一定的情况下，界面结晶过冷度变化与片状晶间距的关系

在生长速度 R 不变的情况下，对图 2.3-55 中的曲线求极值可得

$$\frac{\mathrm{d}\Delta T}{\mathrm{d}\lambda}=AR-B\frac{1}{\lambda^2}=0$$

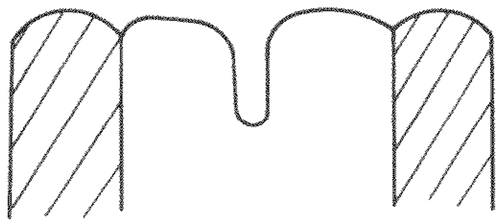


图 2.3-61 片间距过大引起的凹坑

故有 $$\lambda^2 = \frac{B}{A}\frac{1}{R} \quad 或 \quad \lambda = \alpha R^{-1/2} \qquad (2.3\text{-}55)$$

上式表明共晶层片间距 λ 与长大速度 R 之间的关系，即层片间距与长大速度的平方根成反比。在实验中，通过测量共晶的层片间距，可以确定长大速度 R 的大小。

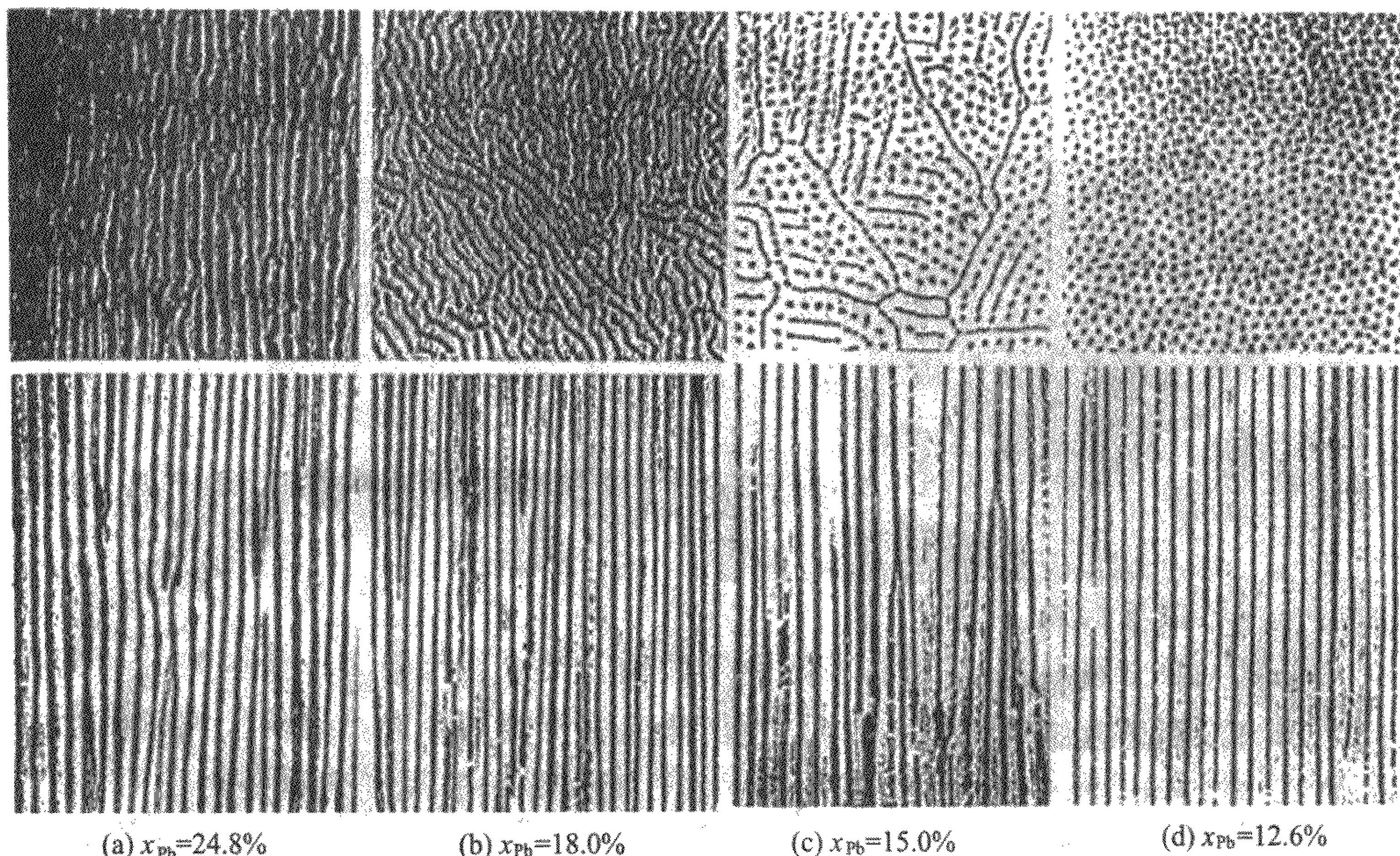
6.3.2 棒状共晶生长

在金属－金属共晶组织中，除层片结构外，还有棒状结构，究竟形成何种结构，取决于共晶中 α 和 β 相间的体积比及第三组元的存在。

（1）共晶中两相体积分数的影响

在 α、β 两固相界面张力相同的情况下，如果共晶中两相的体积比差别不大接近 $\frac{1}{\pi} \sim \frac{1}{2}$，倾向于形成片状共晶，如果共晶中一相的体积含量相对于另一相较低时，即如果一相体积分数小于 $1/\pi$，该相多半以棒状结构出现。如图 2.3-62 为 M. C. Flemings 等人对 Pb－Sn 合金进行的单向凝固试验发现，随着 Pb 量减少。共晶将由层片逐渐转为棒状（共晶成分为 $x_{Pb}=26.1\%$），造成这种情况的原因是晶体表面积大小的影响（或者说表面能的大小）所致。

（2）第三组元对共晶结构的影响



(a) x_{Pb}=24.8%　(b) x_{Pb}=18.0%　(c) x_{Pb}=15.0%　(d) x_{Pb}=12.6%

图 2.3-62 Pb－Sn 合金原始成份对共晶单相凝固组织影响

当第三组元共晶两相中的分配系数相差较大时，其在某一相的固－液界面前沿的富集，将阻碍该相的继续长大；而另一相的固－液界面前沿由于第三组元富集较少，其长大速度较快。这样，由于搭桥作用，落后的一相将被长大较快的相分隔为筛网状组织（如图 2.3-63），继续发展，即成为棒状组织。通常可以看到共晶晶粒内部为层状组织，而在晶粒交界处成为棒状，其原因就在于：在晶粒之间有第三组元富集，从而造成其在共晶两相中分配系数的差别，导致在某一固相前沿出现“成分过冷”，生长滞后，被另一相领先包围生长而成为棒状结构。

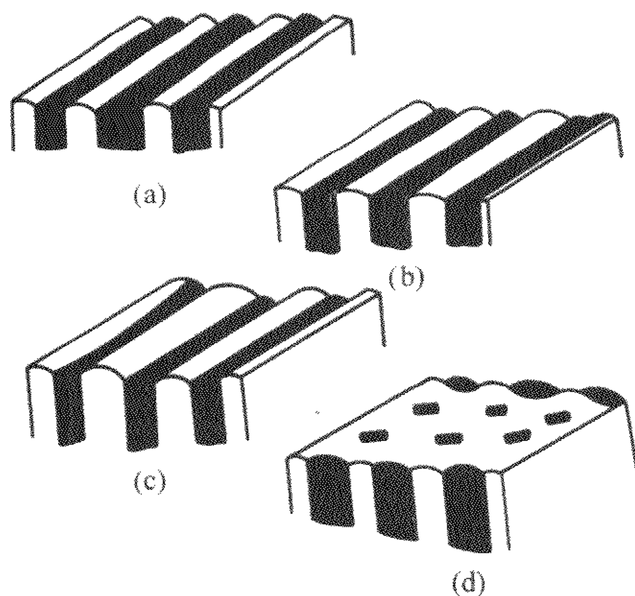


图 2.3-63 层状共晶转变为棒状共晶示意图

6.4 金属－非金属共晶凝固

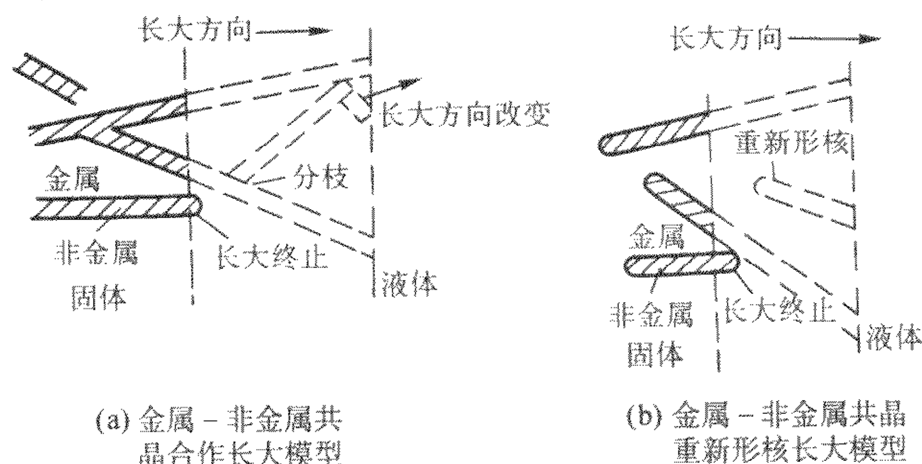
6.4.1 形核与生长

金属－非金属共晶结晶时，其热力学和动力学原理与金属－金属共晶相似，其差别在于非金属相与金属相长大机制不同。共晶凝固时，非金属的固液界面在原子尺度是光滑的，其固－液界面为一特定晶面，因此长大时有明显的方向性，反映出结晶形貌上有较大差异。

金属－非金属共晶的形核与金属－金属共晶相似，凝固开始时，领先相独立在液相中长大，第二相依附于领先相形核并按“协同”方式生长。在通常铸造条件下，一个共晶晶粒的外形是一个球体，我们也可称之为共晶团。但在共晶团内部，两相不是片状或棒状结构，而是互相连接在一起的排列异常紊乱的分枝片，由于金属－非金属共晶两相熔点一般来说差距较大，所以其共生区偏向于高熔点一方。如果高熔点相为领先相，在其形成之后，第二相像光环一样将它包围起来，一直到进入共生区后，两相才开始“合作长大”，因此，在这类共晶中，常形成光环（或称晕圈）组织。

金属－非金属共晶凝固长大时，由于非金属相只能在某些方向长大，所以非金属晶体会出现互相隔离或相互靠近长大的现象。如两个邻近的非金属晶体相对靠近长大时，界面

处将出现非金属原子贫乏的现象，从而使一个或两个晶体停止长大、相反，当非金属晶体相互隔离生长时，在它们之间的金属相前沿有非金属原子的富集，有可能重新形成非金属晶核，然后继续长大，那么在这个非金属原子富集区是重新形成非金属晶核继续长大，还是在原有非金属晶体处发生分枝而长向非金属原子富集区，存在两种生长模式（如图2.3-64）。



(a) 金属-非金属共晶合作长大模型　(b) 金属-非金属共晶重新形核长大模型

图2.3-64　金属-非金属共晶长大的两种模型

第一种长大模型称为合作长大。当一个非金属晶体由于缺乏非金属原子供应而停止长大时，它可以通过孪生或形成亚晶界（小角度晶界）将长大方向改变到非金属原子富集区，这样就产生了非金属晶体的分枝，按照这种模式生长，非金属相内部是相连的。

第二种长大模型称为重新形核长大。当两个非金属晶体相对长大靠近时，将导致一个或两个晶体长大的停止，而新的晶核将在非金属原子富集区重新形成，在这种情况下，非金属晶体将是不相连的。

将Al-Si和Fe-C（石墨）共晶试样深腐蚀后观察共晶硅晶体和片状石墨结构，可见它们相互连在一起，见图2.3-65，前者称为纤维状组织，后者如网状。实验证实金属-非金属共晶是按合作长大模型进行。且共晶中的非金属相在长大过程中不断分枝以改变其生长方向。

6.4.2　非金属晶体在共晶凝固过程中的分枝机制

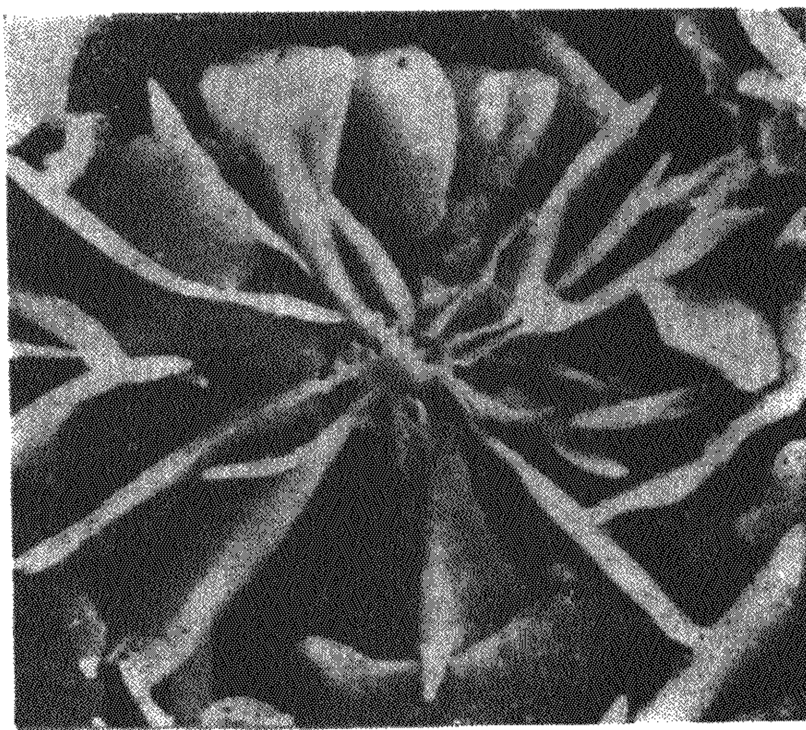
如上所述，金属晶体属于粗糙界面，非金属晶体属于光滑界面，其生长靠光滑界面侧面扩展长大，所以一般认为非金属的固-液界面生长落后于金属晶体，然而实验结果正好相反，在共晶凝固时，非金属相总是领先于金属相长大。其机制在于非金属相在固-液界面上有改变其长大方向的能力。下面以Al-Si和Fe-C（石墨）合金为例来说明这个过程。



(a) 未变质共晶Al-Si合金，层片共晶硅



(b) 共晶Al-Si合金Na变质纤维状共晶硅



(c) 灰铸铁中的片状石墨

图2.3-65　金属-非金属共晶合作长大典型共晶结构

X射线分析表明，硅晶体只能在{111}晶面的<211>或<110>晶向上长大，因此长大后的晶体为片状。解剖单向凝固的Al-Si合金的横断面，发现存在有不少{111}面孪晶沟槽，（如图2.3-66），硅晶体的长大就是通过硅原子优先吸附在这些{111}沟槽上进行的，同时这些孪晶{111}的存在，也为硅晶体在长大过程中改变方向提供了条件。在

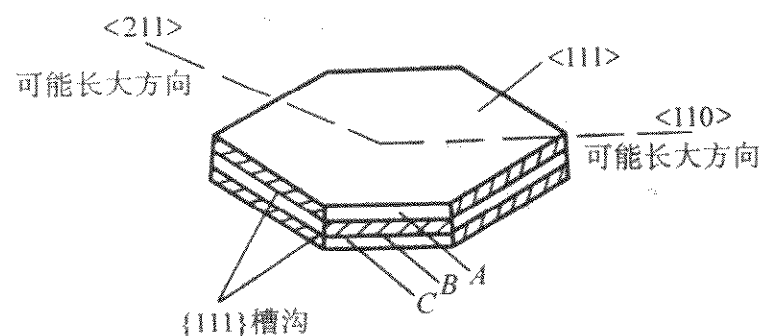


图2.3-66　非金属硅晶在共晶凝固中的长大方式

A、B、C互相孪生，长大是通过原子优先吸附在晶体边缘的{111}沟槽上进行的

共晶凝固中由于 Al－Si 两相在凝固过程中的收缩不同和原子错排，这就为 Si 晶的领先生长创造了条件。生长速度愈大（过冷度 ΔT 增加），硅晶的分枝愈厉害（减少片间距），其共晶硅可由板片状转变为多分枝的纤维状（见图 2.3-65）。

在 Fe－C（石墨）共晶中，石墨也不是单晶体，而是由许多组织单元聚合而成，每一个亚组织单元是一个单晶体。它们之间通过孪晶或亚晶界相互连接。这些孪晶界和亚晶界也是在凝固过程中由于石墨与奥氏体在收缩上的不同或原子错排造成的。X 射线研究发现，石墨的基面含有旋转孪晶（如图 2.3-67）有利于石墨片垂直于棱面长大，同时也为石墨晶体在长大过程中改变方向创造了条件。共晶石墨的分枝就是靠这些孪晶形成的。当冷却速度增加时，石墨频繁地弯曲和分枝，以致形成过冷型石墨组织（如灰铸铁中的 D、E 型石墨）。

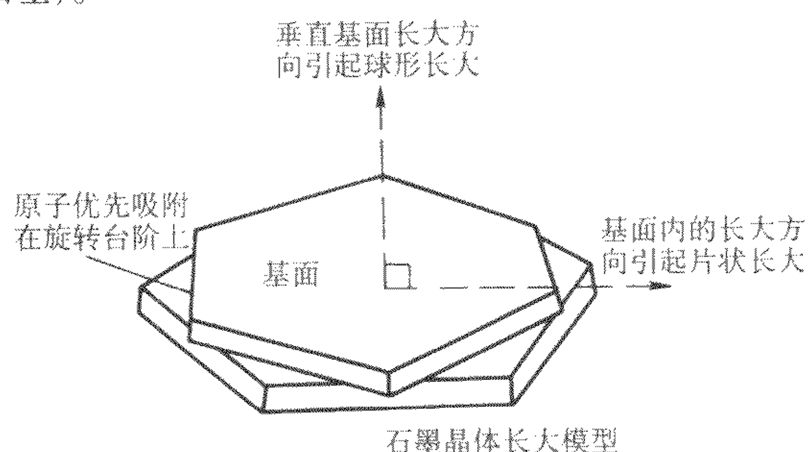


图 2.3-67　石墨晶体长大模型

6.4.3　第三组元的作用

第三组元对金属－非金属共晶凝固生长及组织结构的形成也有重要影响。对常用铸造合金（如 Al－Si、Fe－C－Si 合金等）的变质，包括孕育、细化处理的研究，是人们认识第三组元作用的最好途径。

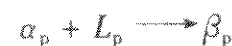
研究表明，Al－Si 共晶合金加入 Na 可细化共晶硅，大幅度提高力学性能。对 Na 的变质机制，较多的看法是：Al－Si合金加入 Na 后，可吸附在硅晶 {111} 孪晶间沟槽中，抑制硅晶长大、使 α_{Al} 的生长速度提高，并强迫硅晶多分枝以维持共晶生长，成为纤维状结构。

在 Fe－C 合金中第三组元影响的最好实例是 O、S、Mg、Ce 等元素对球铁石墨球化的影响。在工业用 Fe－C 合金中，由于含有大量 S、O 元素，他们大量吸附于石墨六方晶系中的棱面，提高了棱面的原子结合能，使石墨晶体在凝固过程中 a 轴方向生长速度提高，成为片状结构。另外，S 的存在还使得石墨的（0001）基面容易成为奥化体生长的基底，阻碍了碳原子向基面的堆砌。增强了石墨的片状生长倾向。加入 Mg、Ce 球化元素后，与 S、O 结合，一方面去除了硫对石墨 a 轴生长的促进作用，又使（0001）面难以成为奥氏体结晶的基底，从而提高了石墨在 C 轴方向的生长速度，使石墨成为球状结构。

由于熔融合金的凝固观察有相当困难，元素间的相互作用也不易评估，因此，第三组元对共晶凝固的影响的研究还有待深入，但合金元素或微量元素在共晶凝固中的作用，已成为现代物理冶金学的重要内容，也是合金凝固及控制的重要手段。

6.5　包晶凝固

典型的包晶平衡相图如图 2.3-68 所示。其特点是：①液相完全互溶，固相部分互溶或完全不互溶；②有一对固、液相线的分配系数小于 1、另一对固、液相线的分配系数大于 1。平衡条件在温度 T_p 可发生包晶反应



$$\alpha_p + L_p \longrightarrow \beta_p$$

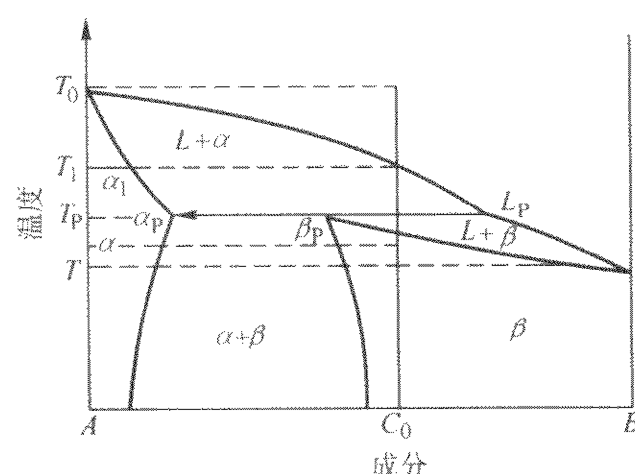


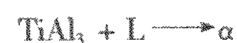
图 2.3-68　包晶平衡相图

β 相的形核可以 α 相为基底，也可在液相中直接形成。平衡凝固要求溶质组元在两个固相和一个液相中进行充分的扩散，但实际上由于冷却速度很快，非平衡凝固是经常发生的。

在非平衡凝固时，由于溶质在固相中的扩散不能充分进行，包晶反应之前凝固出来的 α 相内部的成分是不均匀的，即树枝晶心部溶质浓度低，而枝晶边缘溶质浓度高，当温度达到 T_P 时，在 α 相的表面发生包晶反应。从形核功的角度看，β 相在 α 相表面非均质地形核更为有利。因此，在包晶反应中，α 相很快被 β 相包围，此时，液相与 α 相脱离接触，包晶反应只能依靠溶质组元从液相一侧穿过 β 相向 α 相一侧进行扩散才能继续下去，因此受到很大限制。当温度低于 T_P 后，β 相继续从液相中凝固，并可能出现残余共晶组织，图 2.3-69 为非平衡凝固条件包晶反应示意图。

多数具有包晶反应的合金，其溶质组元在固相中的扩散系数很小，因此，在非平衡凝固条件，包晶反应进行得不完全。由于溶质组元在固相中扩散的不充分，本来是单相组织却变成了多相组织，如图 2.3-70 为 Pb－Bi（$w_{Bi}=20\%$）合金在非平衡凝固条件下溶质分布及组织示意图。当然，一些固相扩散系数大的溶质组元，如钢中的 C，包晶反应可以充分进行，具有包晶反应的碳钢，初生 α 相可以在冷却到奥氏体区后完全消失。

利用包晶反应获得晶粒细化是非常有效的，如向 Al 合金液中加入少量 Ti，可以形成 $TiAl_3$，当 Ti 的质量分数超过 0.15%时将发生包晶反应



$$TiAl_3 + L \longrightarrow \alpha$$

包晶反应产物 α 为 Al 合金的主体相，它作为一个包层包围着非均质核心 $TiAl_3$，由于包层对于溶质组元扩散的屏障，使得包晶反应不易继续进行，也就是包晶反应产物 α 相不易继续长大，因而获得细小的 α 晶粒组织。

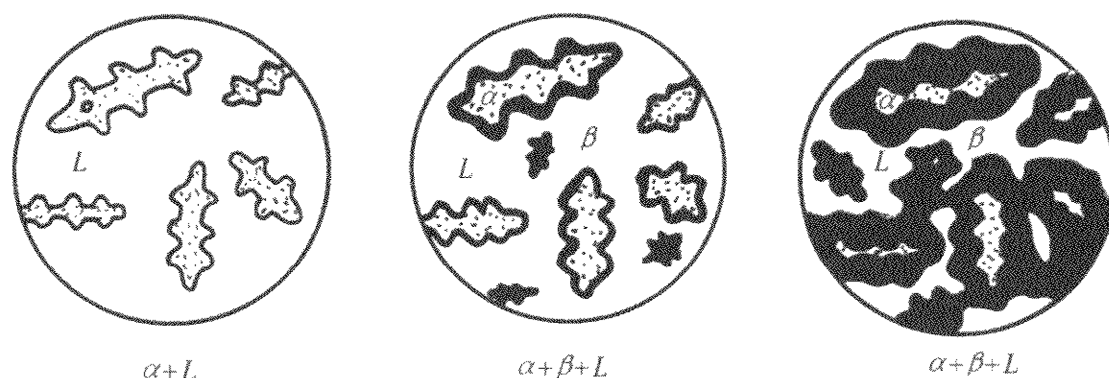
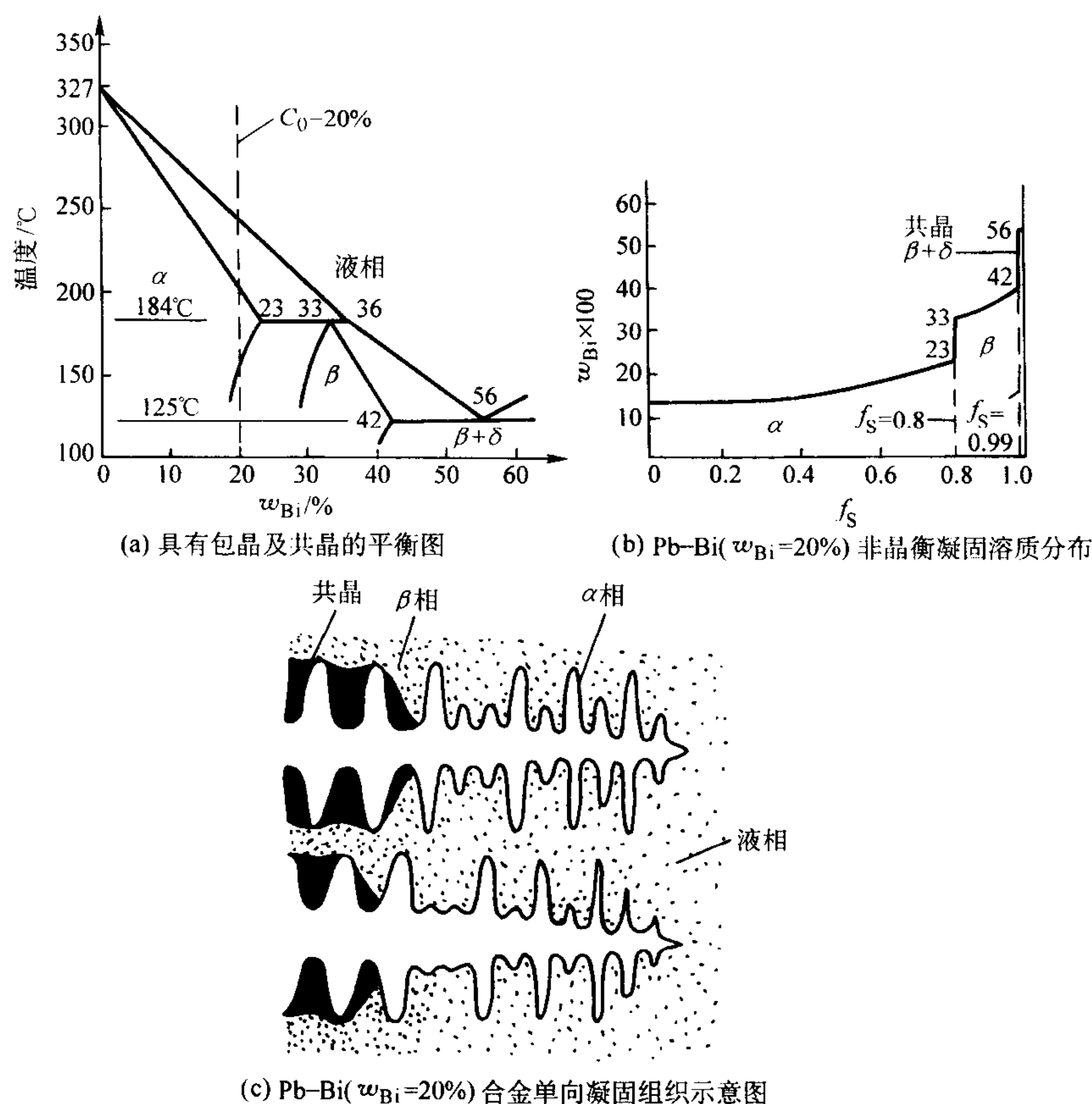


图 2.3-69　非平衡凝固条件下包晶反应示意图



(a) 具有包晶及共晶的平衡图　　(b) Pb-Bi(w_{Bi}=20%) 非晶衡凝固溶质分布

(c) Pb-Bi(w_{Bi}=20%) 合金单向凝固组织示意图

图 2.3-70　Pb－20%Bi 合金在非平衡凝固条件溶质分布及组织图

7　定向凝固和快速凝固

7.1　定向凝固

定向凝固又称为单向凝固或单向结晶，是制备高纯度和高性能单晶、柱状晶和自生复合材料的重要手段。20 世纪 20－30 年代 Northcoff 等人就证实用定向凝固技术可得到拉长的合金晶粒，20 世纪 60 年代定向凝固技术首次在燃气涡轮发动机高温叶片的生产获得成功，大幅度提高了叶片的使用寿命与工作温度，促进了涡轮发动机的发展。90 年代后，随着定向凝固理论研究的发展和深入，定向凝固铸件生产、单晶喷气发动机叶片制造、连续单晶及定向结晶钢棒等技术陆续开发出来，使定向凝固技术成为获得高性能（高温等）合金、电磁功能材料、共晶复合材料的重要手段，也成为现代高新技术发展的重要内容。

7.1.1　定向凝固原理

为了保证定向凝固组织，应保证做到：①使液态合金在铸型中产生定向散热以达到定向凝固；②保证有足够大的 G_L，避免晶体前沿出现“成分过冷”，以防止枝晶及等轴晶的形成；③防止型壁及熔体表面生长出晶体而成为等轴晶。

以坩埚下降法为例，讨论定向凝固固液生长界面温度分布（如图 2.3-71）。若将生长过程中的热量传递简化为一维传热，则在热平衡条件下，依傅里叶一维传热方程，在固－液界面稳定推进时，有如下热量平衡

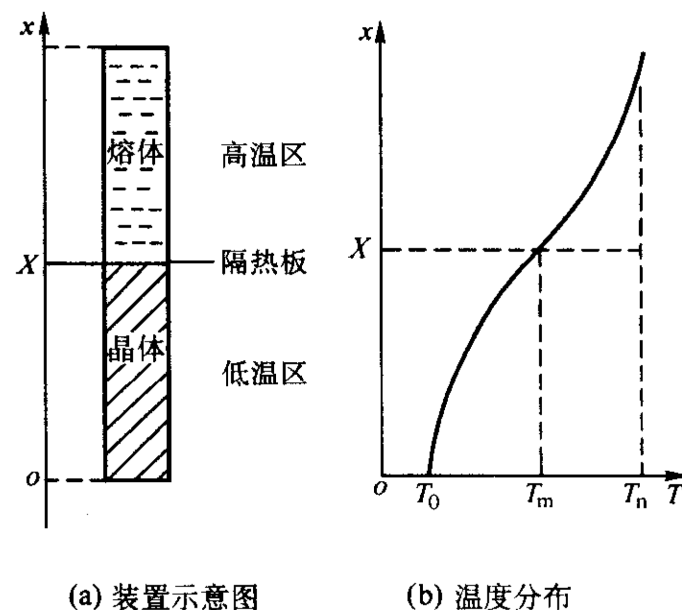
$$k_s G_s - k_L G_L = H\rho_m R \qquad (2.3\text{-}56)$$

或

$$G_L = \frac{k_s G_s}{k_L} - \frac{H\rho_m R}{k_L} \qquad (2.3\text{-}57)$$

式中，k_s、k_L 分别为固液两相的热导率；H 为合金的凝固潜热；ρ_m 为合金密度；R 为界面推进速度；G_L、G_s 为界面液固两相温度梯度。

在实际生产和实验条件，只有充分大的 $\frac{G_L}{R}$ 才能保证单晶质量；只有较大的 R 才能保证一定生产率。由上两式可见，当凝固速度 R 一定时，G_L 与 G_s 成正比，增大 G_s（如改进设备散热方向和强度），可增强凝固固相的散热强度，可获得较大的 G_L，从而保证单晶稳定推进，保证铸件（锭）质量。固相散热强度的提高，不仅提高 G_L，也提高了生长速度 R，但易造成“成分过冷”，不利于晶界稳定，而且固相温度梯度 G_s 过大，使生长的晶体产生大的内应力，甚至开裂，因此在生产和实验中，常用提高固－液界面前沿熔体的温度来增加 G_L，即在单向凝固设备中在液固界面附近加上幅射板即是此目的。但熔体温度过高，会导致液相剧烈挥发、分解和受到污染，反而影响晶体质量。



(a) 装置示意图　　(b) 温度分布

图 2.3-71　坩埚下降法定向凝固装置和温度分布

假定定向凝固的晶体为圆柱体（断面半径为 a），晶体在稳定生长过程中除与结晶器有一维导热以外，还通过柱周

围与环境换热，Flemings提出了凝固界面固相侧晶体中的温度分布函数 T 及固相温度梯度 G_s 表达式

$$T=(T_m-T_0)\exp\left\{\left[-\frac{R}{2\alpha_s}+\sqrt{\left(\frac{R}{2\alpha_s}\right)^2+\frac{2h}{ak_s}}\right]x\right\}+T_0 \tag{2.3-58}$$

$$G_s=\frac{\mathrm{d}T}{\mathrm{d}x}\bigg|_{x=0}=(T_m-T_0)\left[\frac{-R}{2\alpha_s}+\sqrt{\left(\frac{R}{2\alpha_s}\right)^2+\frac{2h}{ak_s}}\right] \tag{2.3-59}$$

且当 $\frac{R}{2\alpha_s}\ll\frac{2h}{\alpha k_s}$时

$$G_s\approx\left(\frac{2h}{\alpha k_s}\right)^{1/2}(T_m-T_0) \tag{2.3-60}$$

式中，α_s 为导温系数；h 为界面换热系数；k_s 为合金固相热导率；T_m 为合金熔点；T_0 为环境温度。

由式（2.3-59）可看出，界面温度梯度 G_s、晶体生长速度 R 及界面散热系数 h 三者之间有着相互制约的关系。必要的换热条件，是凝固界面稳定平推的必要条件。要想得到大的 G_s（如低熔点金属）应增大换热效率 h，或减小晶体半径 a。

7.1.2 单向凝固的主要工艺

（1）发热剂法

将型壳置于绝热耐火材料箱中，底部安放水冷结晶器。型壳中浇入金属液后，在型壳上部盖发热剂，使金属液处于高温。建立自下而上的凝固条件。由于无法调节凝固速率和温度梯度，该法只用于制备小型柱状晶铸件（如用于磁钢生产）。

（2）功率降低法（P.D法）

如图2.3-72所示，铸型加热感应圈分两段，铸件在凝固过程中不移动，当模壳预热到一定温度时，浇入过热合金液，切断电源，上部继续加热，G_L 随着凝固距离的加大而减小，因此柱状晶在锭上部粗化严重。该工艺 G_L、R 值都不能控制，只适于高度在120 mm以下的定向凝固铸件生产。

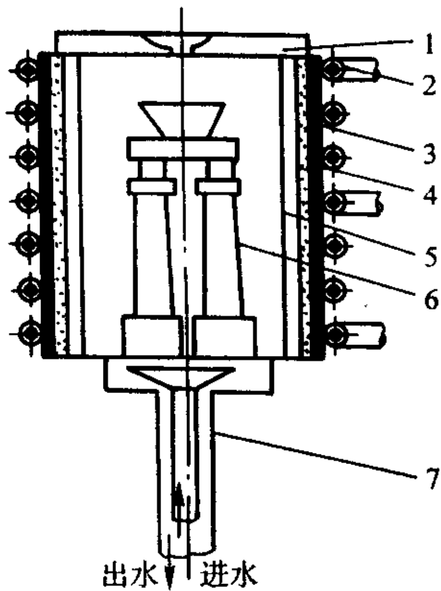


图2.3-72 P.D法示意图

1—保温盖；2—感应圈；3—玻璃布；4—保温层；5—石墨套；6—模壳；7—结晶器

（3）快速凝固法（H.R.S法）

如图2.3-73所示。与P.D法的主要区别是：铸型加热器始终加热，凝固时，铸件与加热器之间产生相对移动。另外，在热区底部使用辐射挡板和水冷套。在挡板附近有较大的温度梯度 G_L、G_s。与P.D法比可大大缩小凝固前沿两相区，局部冷却速度增大。有利组织细化，提高力学性能。

（4）液态金属冷却法（L.M.C法）

L.M.C法是1974年出现的一种新型单向凝固方法，如

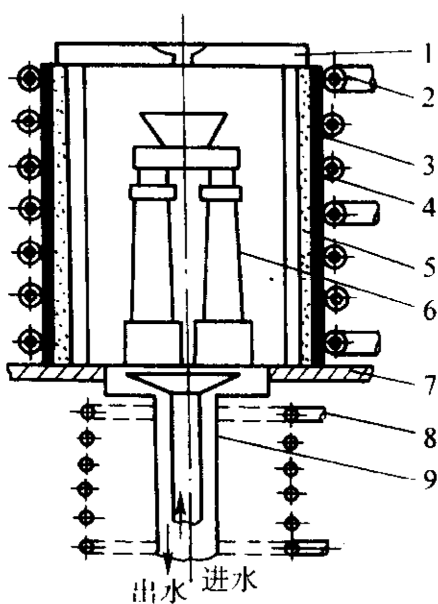


图2.3-73 H.R.S法示意图

1—保温盖；2—感应圈；3—玻璃布；4—保温层；5—石墨套；6—模壳；7—挡板；8—冷却圈；9—结晶器

图2.3-74，其工艺与H.R.S法基本相同。当合金浇入型壳后，按选择的速度将壳型推出炉体，浸入金属浴，金属浴的水平面保持在固-液界面处，并使其保持在一定温度范围，从而获得稳定的 G_L/R 最大值。由于采用低熔点的液态金属冷却，其冷速可为P.D法的500倍以上。常用的金属物质有：锡液、镓铟合金，镓铟锡合金等。

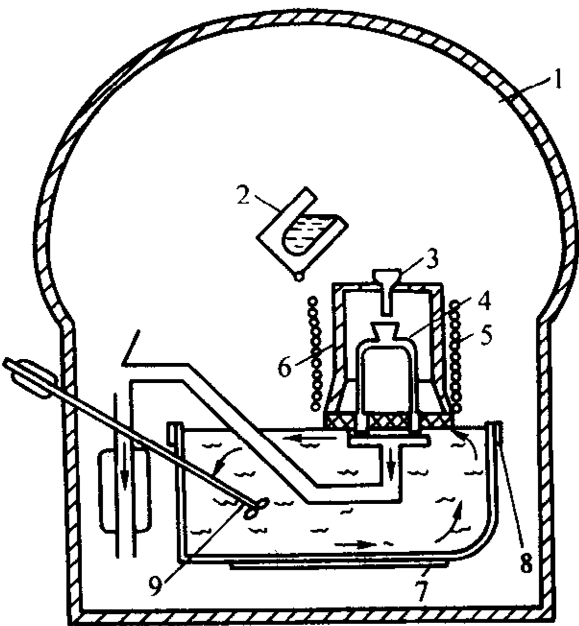


图2.3-74 L.M.C法示意图

1—真空室；2—熔炼坩埚；3—烧杯；4—炉子的热区；5—挡板；6—模壳；7—锡浴加热；8—冷热罩；9—锡浴搅拌器

7.1.3 单晶生长

单晶是人们认识固体的基础，单晶体的研究和生产可使人们发现金属新的性质（如各向异性、超导性、核磁共振等），开发高性能材料（如晶须、半导体件等）。定向凝固是制备单晶体的最有效办法。

（1）单晶生长特点

要在金属熔体中形成一个单晶核，可以引入籽晶或让其均质形核，然后单晶体在晶核和熔体界面上不断长大。在单晶体生长过程中，要绝对避免固-液界面不稳定而长出胞晶和柱状晶，因而在固-液界面前沿不允许有温度过冷和成分过冷。界面前沿的熔体应也处于过热状态，结晶潜热只能通过生长着的晶体导出。定向凝固可满足上述要求，生产中只要恰当地控制固-液界面前沿熔体的温度和晶体生长速度，可以得到高质量的单晶体。

按成分和晶体特征不同，单晶体可有以下几种：①晶体和熔体成分相同。纯元素和化合物属于这一类，由于是单元素，在生长过程中晶体和熔体的成分保持恒定、熔点不变，生长速度也允许较快；②晶体和熔体成分不同。在单晶中掺

入一定浓度的杂质以改善电学等性能。这类材料掺入元素或化合物后变为二元或多元系，在生长的固－液界面有溶质再分配，受熔体中溶质的扩散和对流的影响，单晶成分较难均匀；③有第二相或出现共晶的晶体。如 Ni 基高温合金单晶铸造组织不仅含有大量基体 γ 相和沉淀析出的 γ' 强化相，还有共晶析出于枝晶干间，这些枝干同属一个晶体，不存在晶界。

（2）单晶生长方法

根据熔区特点，单晶生长可分为正常凝固法和区域熔化法。

1）正常熔化法　正常熔化法制备单晶，最常用的有坩埚移动法、炉体移动法及晶体提拉法。

① 坩埚移动或炉体移动单向凝固法　如图 2.3-75 所示，坩埚可垂直或水平置于炉内，凝固过程由坩埚一端开始，最常用的是将尖底坩埚垂直沿炉体逐渐下降（见图 2.3-75a），单晶体从尖底部缓慢向上生长。也可将“籽晶”放在坩埚底部，当坩埚向下移动时，“籽晶”开始结晶，并随固－液界面移动，单晶不断长大。该法的主要缺点是晶体和坩埚壁接触容易产生应力或寄生成核，在生产高完整性单晶时，较少采用。

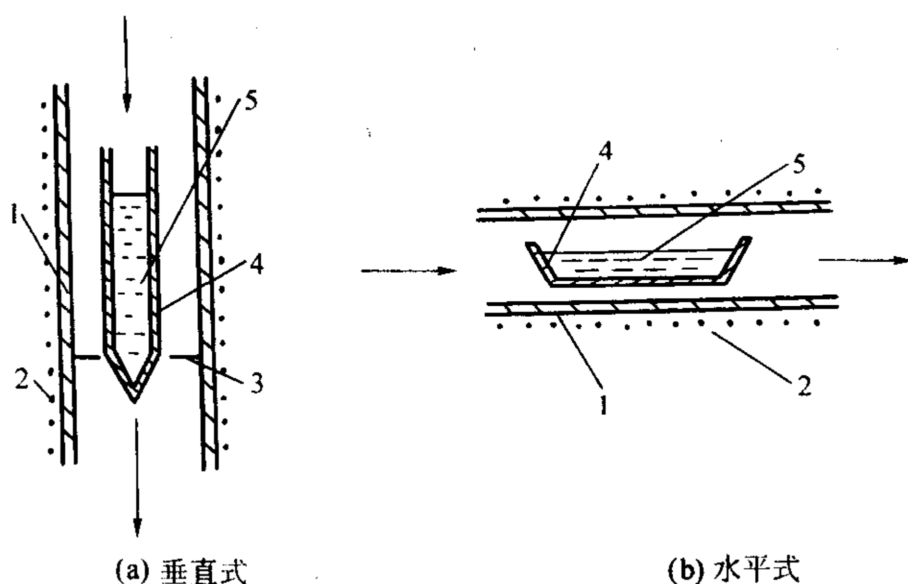


图 2.3-75　坩埚移动单向凝固示意图
1—炉管；2—炉丝；3—挡板；4—坩埚；5—熔体

② 晶体提拉法　或称为邱克拉斯基技术，其原理如图 2.3-76 所示，将欲生长的材料放在坩埚内熔化，然后将“籽晶”插入熔体，在适当的温度下，“籽晶”稳定长大。该法的优点是：可方便观察晶体生长；减少坩埚壁与晶体的接触生长应力，并防止寄生生核，在较快冷速形成低位错密度和高完整的单晶。

2）区域熔化法　区域熔化法（或区域提纯）是获得超纯材料的有效手段，区域熔化法有水平区熔法和悬浮区熔法两种。

以水平区熔法为例，说明材料提纯原理，如图 2.3-77，全试样只有一小熔区 l，当这个熔区从试样左端移向右端时，由于析出固相时溶质扩散困难，使析出固相的浓度总低于界面固相浓度和平均浓度 C_0（见图 2.3-78a），大量溶质

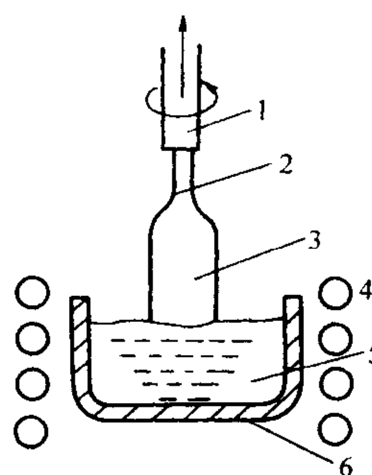


图 2.3-76　单向结晶晶体提拉法
1—提拉杆；2—籽晶；3—晶体；4—射频感应圈；5—熔体；6—坩埚

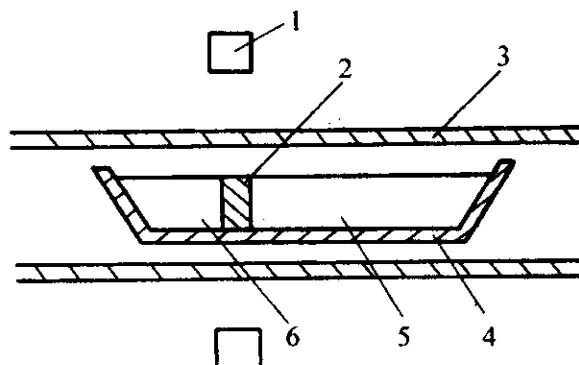


图 2.3-77　水平区熔示意图
1—加热器；2—熔区；3—炉管；4—坩埚；5—多晶；6—晶体

在界面液相中富集，并逐渐扩散进入液相内部。一次提纯后已凝固固相大部分浓度小于平均浓度 C_0，而在锭的右端尾部，浓度大于 C_0（见图 2.3-78b）。一次提纯操作，并不能达到目的，但多次重熔提纯后，材料的纯度大幅度提高（见图 2.3-79）。

假定区域提纯中的小熔池左边界面在提纯过程中固液界面的固相浓度为 C_s^*，合金原始浓度为 C_0，固液界面平衡分配系数为 k，固相浓度 C_s^* 与熔池所处锭中位置的关系为

$$C_s^* = C_0[1-(1-k)\exp(-\frac{k}{l}x)] \tag{2.3-61}$$

7.2　快速凝固

快速凝固指的是金属液在比常规工艺快得多的冷却速度下（$10^4 \sim 10^9$ K/s），或大的过冷度（几十至几百 K），以快的冷却速率（> 10 cm/s）从液态向固态转变的过程。由于凝固速率极快，合金凝固远离平衡状态，生核力大幅提高，溶质扩散困难，晶体不易长大，合金组织发生突变如许多有限互溶合金出现无限互溶现象，枝晶间距缩小，偏析程度大大降低，出现许多亚稳相、甚至微晶和非晶态结构、从而获得许多独特性能。

快速凝固技术由美国加州理工学院 P.Duwez 等在 20 世纪 60 年代首先提出，他采用冷淬气枪法在试验室获得 Cu－Ag

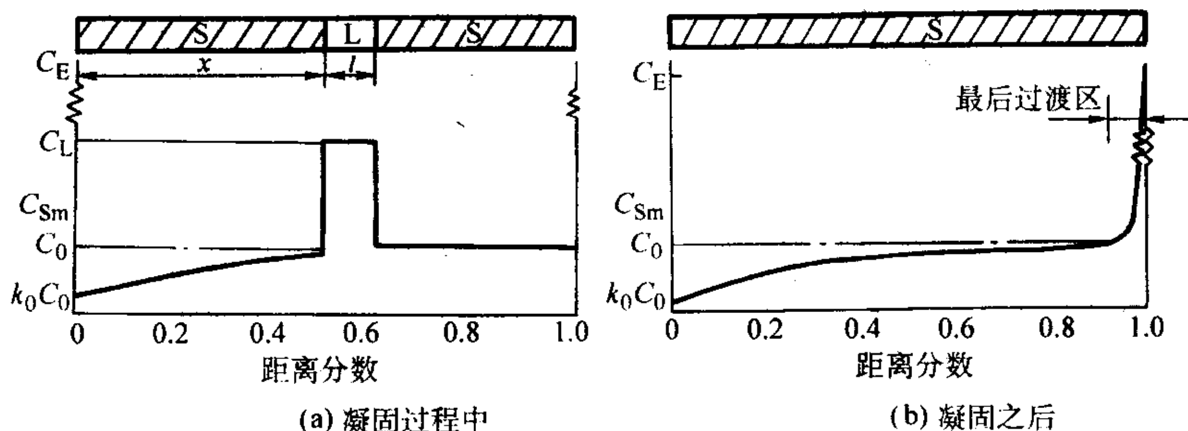


图 2.3-78　区熔的溶质分布

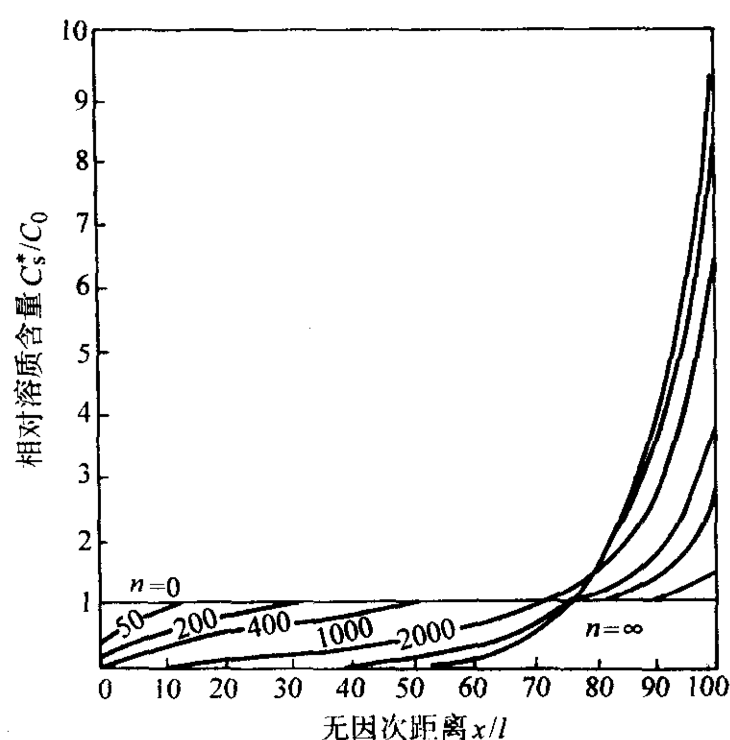


图 2.3-79　多次区熔后溶质分布
n—熔区提纯次数；$k=0.952\ 4$

合金单晶薄膜和 Au - Si 非晶态结构（金属玻璃），之后，快速凝固技术受到普遍关注，并得到快速发展，成为现代物理冶金和材料科学研究的新领域，也成为制取高性能新型材料的重要手段。

7.2.1　快速凝固主要工艺

要使合金熔体产生超常规的快速凝固过程，必须要有高效的快速凝固方法。为了得到高冷却速度，要满足两个条件，第一是合金流至少一个方向上很薄，以利导走热量；第二是要有一个高效的吸热衬底，以便从熔融金属流快速吸取热量。大体上，有三种金属流型和三种冷却类型可互相配合形成不同的生产方法。三种金属流是金属滴、柱状流和带状流。而三种冷却类型是气体、液体和固体冷却。图 2.3-80 是上述配合得到的生产方法。

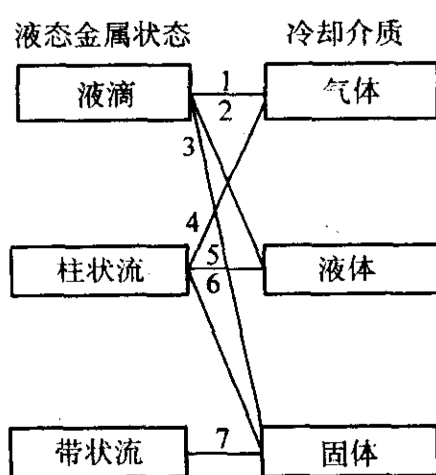


图 2.3-80　熔融金属流和冷却介质配合情况
1—雾化，快凝离心雾化，旋转电极处理；2—水雾化，快速旋转杯；3—冷淬枪，活塞砧，电子束 - 薄片冷淬，可控雾化沉积，雾化沉积；4—泰勒拉丝，挤压造丝；5—挤压造丝，旋转水造丝；6—熔体提拉，悬滴拉丝；7—熔体挤喷，片状流铸造，熔体拉拔，熔体溢流

快速凝固的方法虽然各不相同，但大致可分为急（激）冷、雾化和表面处理三类。急冷法是凝固前和凝固过程中都保持着金属的连续性；雾化法是金属在凝固前被分割为液滴；表面处理法则是使固体金属表面快速熔化一薄层并随后快速凝固。在这三类方法中，整个基体都是冷却源。图 2.3-81 是上述三种快速凝固方式和产品的分类。

下面简要介绍几种典型的快速凝固工艺。

（1）高凝固速率

通过高热导率的材料衬底，使喷射至衬底上的薄层熔体快速冷却，其代表方法有以下几种。

1）气枪法（cun technique）　该方法主要用于实验室制取快速凝固合金样品，其基本原理是将熔融的合金液滴（常

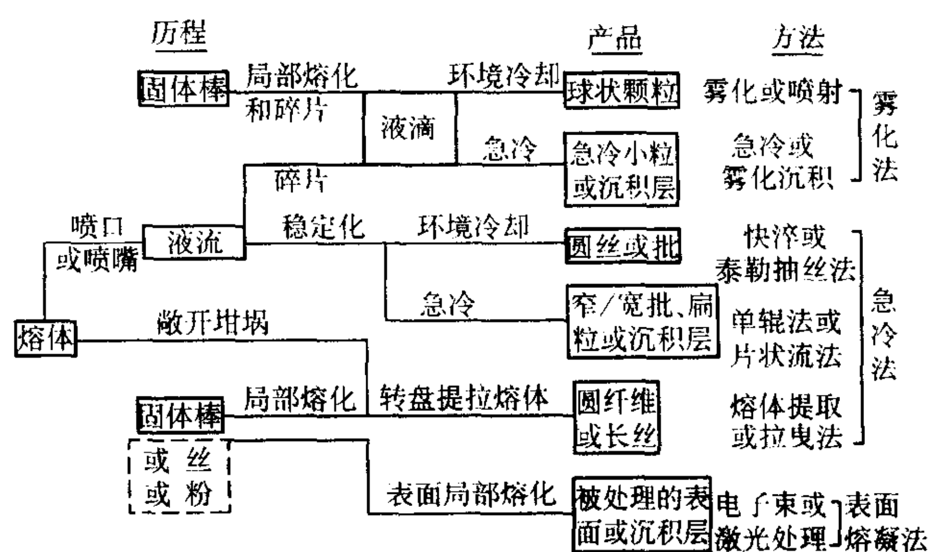


图 2.3-81　快速凝固方法和其相应产品的分类

为数克）在气枪中高压（> 5 MPa）惰性气体流（Ar 或 He）的带动下，射向急冷衬底上（常为纯铜）。由于极薄的液态合金与衬底紧密相贴，可获得较高的冷速（$>10^7$ K/s），该法所获得的薄膜外形不规则，厚度不均及多孔，最薄处的厚度在 200 nm（冷却速度 10^9 K/s）。这类薄膜仅可供作分析用，无法测定力学性能，不能用于工业生产。该法示于图 2.3-82a。

2）锤钻法（piston and auril technique）　该方法中，垂直落下的液滴被水平方向迅速合拢的两块高热导率衬底所挤压，形成的合金箔片均匀致密、厚度可达几十微米，冷却速度也可达 10^6 K/s。由于不能大量制备，主要用于实验室研究。

3）旋铸法（chill block melt spining）与平面流铸法（planar flow casting）　该方法产生 20 世纪 70 年代，是将熔融的合金液在惰性气体（如 Ar）的压力下射向高速旋转的，以高热导率材料制成的辊子外表面。辊面运动的线速度很高（> 30 ~ 50 m/s），液态合金与辊面相贴，连续地凝固成一条厚度均匀的致密合金薄条带。合金条带的厚薄和冷却速度受辊面运动的线速度影响，条带厚度一般有几十微米，最薄可达 15 ~ 20 μm（冷速为 10^6 ~ 10^7 K/s），旋铸法可大量提供致密合金带，因而可工业应用。为生产非晶软磁合金的需要，进一步发展了平面流铸法，其不同之处在于采用了窄缝喷射咀以增加条带宽度，可增加合金条带宽度几十到几百毫米。我国钢铁研究总院已建成一次浇注量为近百公斤的非晶软磁合金宽条带生产设备。此法见图 2.3-82b。

4）旋转叶片打击法（rotating blades melt quenching）　其工作原理为：在气压作用下（0.3 ~ 0.4 MPa），合金熔体细流射向一高速旋转（4 500 ~ 10 000 r/min）的内水冷铜制叶轮。合金液流在高热导率叶片的打击下，紧贴叶片分散和展开，并在径向脱离叶片之前迅速凝固。此法的特点在于旋转叶片给予合金液流一个巨大打击力（约 5 MPa），故界面传热及冷速较高。此法见图 2.3-82c。

（2）超声雾化法（utrasonic gas atomization）

在该法中，流动速度达到 2 马赫，脉冲频率为 60 000 ~ 80 000 Hz 的惰性气体冲击金属流，从而可达到雾化效果，冷却速度可达 10^4 ~ 10^5 K/s。雾化法制得的合金粉粒尺寸一般为 10 ~ 100 μm。雾化法所制得的合金粉粒，可通过热挤、热压后成为坯料或零件。此法见图 2.3-82d。

另有一种离心气体雾化法，是将合金融体射向一高速旋转的铜制激冷盘上表面，在离心力的作用下，合金液雾化成颗粒（见图 2.3-82e）

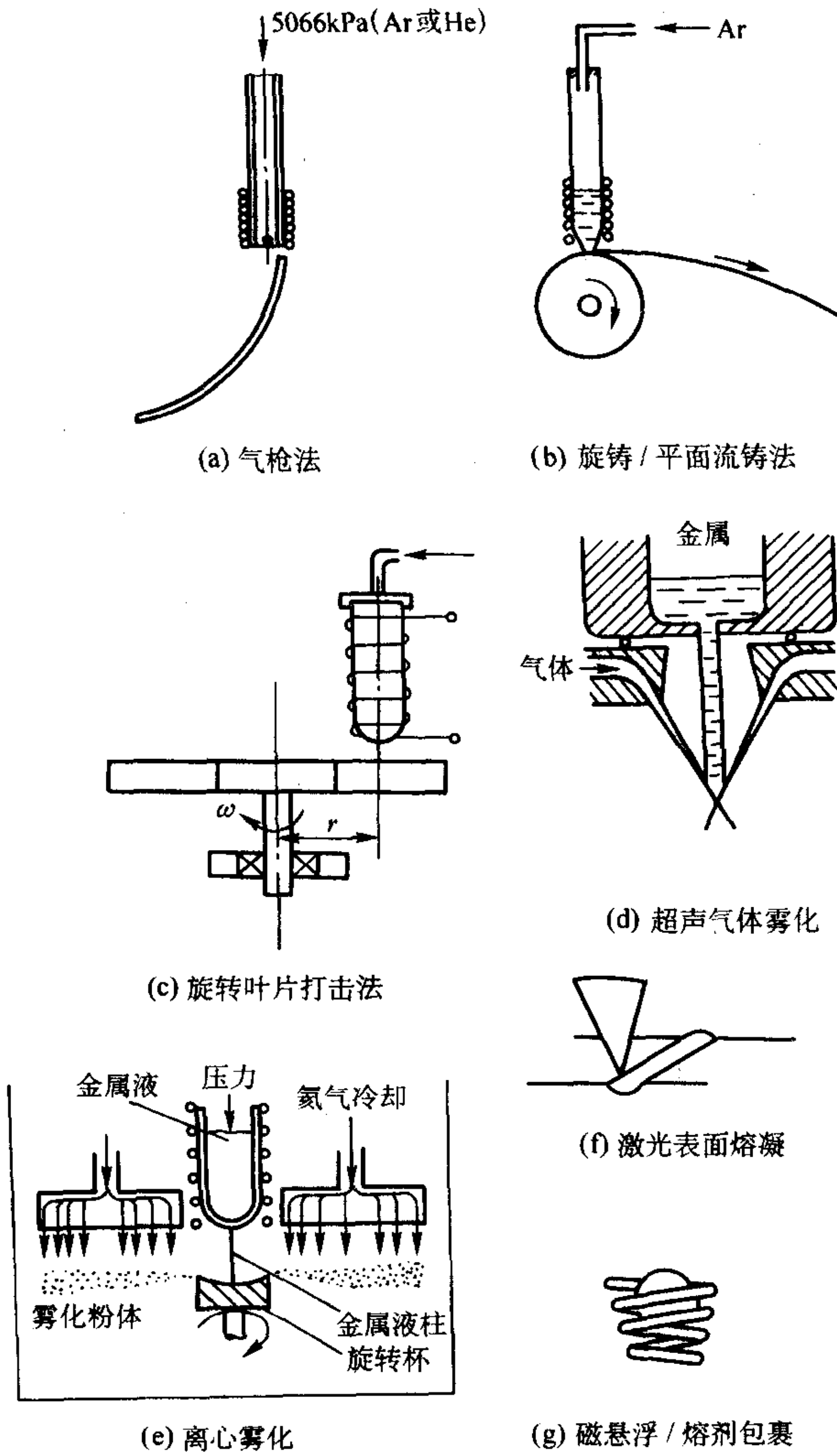
（3）水冷试样深过冷法

前面所述快速凝固技术所得制品至少在一个方向必须很小，如能开发大尺寸的块状试样快速凝固技术将有广阔的工业应用前景。实现该目标有两条途径：一是选择合金系熔体应能在保证不太高的冷却速度下达到大的起始过冷和快速凝

固；二是通过消除或部分消除熔体中的非均质形核作用，使其在缓慢的冷却过程中亦能达到大的起始形核过冷和发生快速凝固。目前已采用的技术有：磁悬浮无坩埚熔炼法及有净化作用的玻璃混合料包裹合金熔体法。目前仅用于试验室研究，见图 2.3-82g。

(4) 激光或电子束表面快速熔凝（laser/electron beam surface reatment）

该法用高能量的激光或电子束以很高的线速度扫描工件表面，在工件表面形成瞬间的薄层小熔池，热量由基底材料迅速吸收，在表面形成一个快速移动的温度场，从而实现快速凝固。表面的冷却速度可达 10^8 K/s 以上，快速凝固主要由快速移动的温度场所造成。该法可用于大尺寸工件表面的快速凝固，已有工业应用，其工艺见图 2.3-82f。



(a) 气枪法 (b) 旋铸/平面流铸法 (c) 旋转叶片打击法 (d) 超声气体雾化 (e) 离心雾化 (f) 激光表面熔凝 (g) 磁悬浮/熔剂包裹

图 2.3-82 几种快速凝固方法示意图

(5) 喷射成形（spray forming）

该法是将雾化的合金液连续喷射到移动的衬底材料上，形成连续沉积的、大尺寸的、几何形状接近实际工件的坯件。坯件密度可达理论密度的95%以上，经后续处理，可得到完全致密的大块材料。虽然该法冷却速度在 $10^2 \sim 10^3$ K/s，低于其他的快速凝固方法，但在大尺寸制品中能有如此高的冷却速度，其组织性能仍较优越，且有广阔的应用前景。

7.2.2 快速凝固的传热特征

(1) 薄层熔体在固态衬底上的传热

由于合金薄膜的顶面与边缘不与冷衬底接触，散热较少。所以此类传热可简化为单向传热。研究表明，影响温度场及冷却速度的最主要因素是金属衬底与合金的界面状况以及金属膜的厚度。

当金属薄膜与衬底界面传热系数 h 较大时，在试样及衬底中的温度梯度都较大，界面上无温差存在，属理想冷却方式；而当 h 较小时，即试样和衬底温度梯度很小，界面上有较大的温差，这种情况称为牛顿冷却方式，即热量传递由界面控制；而界面处于上述两者之间时，称为中间冷却方式。由此可用准则 hd/λ_s 值来判断何种冷却方式起主导作用（d 为试样膜厚度，λ_s 为试样导热率）即：

$hd/\lambda_s > 3$ 为理想冷却方式

$3 > dh/\lambda_s > 0.015$ 为中间冷却方式

$hd/\lambda_s < 0.015$ 为牛顿冷却方式

在牛顿冷却方式下，当温度处于凝固温度以上时，试样的温度 T 与时间 t 的关系为

$$T = T_b + (T_l - T_b)\exp\left(-\frac{ht}{\rho c_p d}\right) \tag{2.3-62}$$

式中，ρ 为金属密度；T_l 为液态金属起始温度；T_b 为衬底温度；c_p 为金属的比热容。

实验结果表明，目前大部分的快速凝固，试样厚度为几个微米到几十微米，界面传热系数 $h = 10 \sim 30$ W/(cm^2·K)，所以散热过程多属于牛顿冷却方式或中间冷却方式。表面熔化和自淬火方法，界面换热系数极高，可作为理想方式，或接近理想方式的中间冷却方式。

(2) 金属液滴在流体介质中的对流传热

在流体介质中以雾化法进行快速凝固时，热量通过“液滴-流体介质”界面传向介质，在此过程中，金属液滴的平均冷却速度 V_T 可估算为

$$V_T = \frac{h(T_l - T_g)A}{V\left(C_l + \frac{\Delta H_f}{\Delta T_S}\right)} \tag{2.3-63}$$

式中，T_g 为气体介质温度；A 为液滴表面积；V 为液滴体积；C_l 为液态金属体积比热；ΔH_f 为单位体积金属的熔化潜热；$\Delta T_S = (T_l - T_S)$；T_S 为凝固完毕的温度。

可见在熔体性质，过热温度和液滴尺寸一定的情况，传热强度取决于传热系数 h。

7.2.3 合金的快速凝固与组织结构的关系

(1) 结晶过冷度对组织形成的影响

冷却速度的变化对胞晶凝固中组织形成的影响已经讨论过了，它对快速凝固的影响则更为显著。以图 2.3-83 所示的简单共晶系为例，假定：①过冷状态下均质成核；②凝固潜热由体系吸收，不传入环境。那么，在回复温度（凝固潜热所导致的温度回升）T_R 及固相线温度 T_S 的影响下，就有三种可能的结晶过冷情况发生。

第一种情况，液态金属过冷到 T_1 后，由于潜热析出，不回复到 T_R，此时 $T_R > T_S$，处于 $\alpha + l$ 的相区内，在金属内可形成局部平衡和溶质再分配，导致化学成分偏析，这是浅过冷情况；第二种情况是由于过冷较深，到达 T_2 以致 T_R 恰好等于 T_S，称为临界过冷，用热力学计算则有

$$H = \int_{T_2}^{T_s} c_{pl}\,dT \tag{2.3-64}$$

式中，H 为凝固潜热；c_{pl} 为液相的比热容。在 α 相内发生均质凝固，没有化学偏析和固液界面上的局部平衡。第三种情况为深过冷情况，过冷温度很低为 T_3，且 $T_R < T_S$，溶质扩散困难，成分偏析无法产生，α 相的化学成分也是均匀的。

在第二、三种情况，α 相生核很快，但无法长大，最终得到细晶或微晶结构。图 2.3-84 即是过冷程度不同对 Fe-25%Ni 合金显微组织的影响。由图可看出，凝固速度增加所

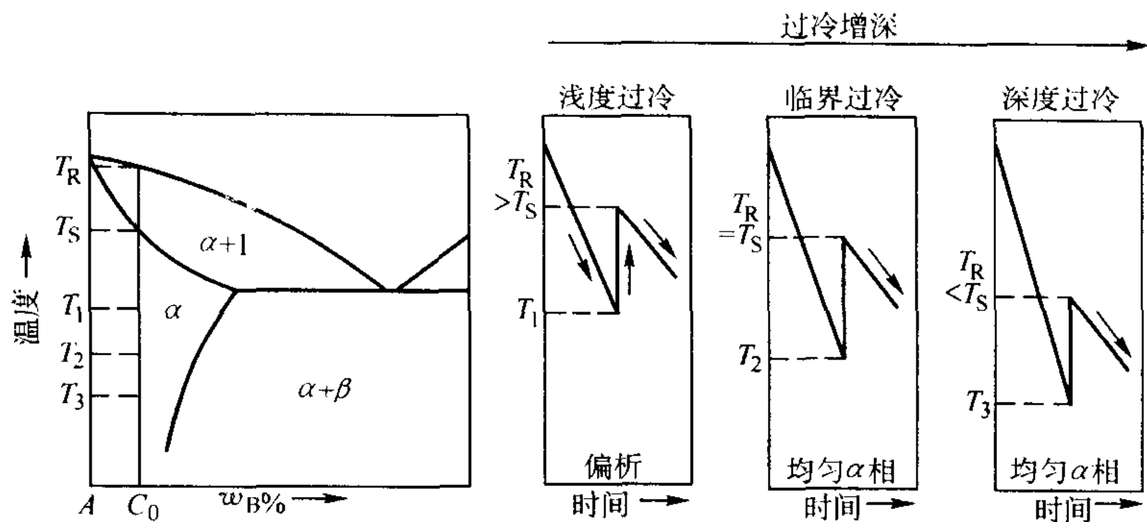


图 2.3-83　简单共晶体系中结晶过冷度对组织凝固的影响

带来的过冷度加大，使枝晶间距和晶粒度大幅度降低，表 2.3-13 为冷却速度和枝晶间距大小的关系。

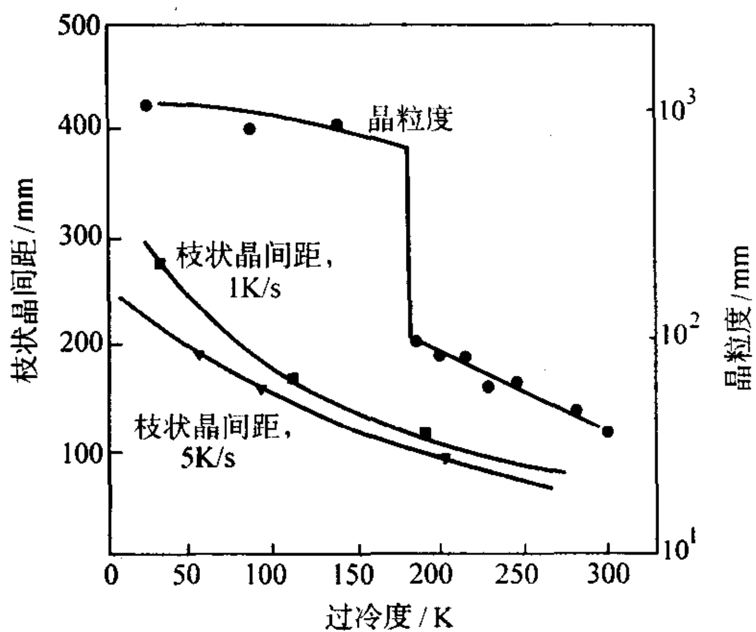


图 2.3-84　过冷程度对 Fe-25%Ni 合金组织的影响

表 2.3-13　枝状晶间距和冷却速度的关系及相应过程

冷却速度/K·s⁻¹	程度	应用过程	特征厚度	特征枝状晶间距
$10^{-6}\sim10^{-3}$	很低	大型砂铸和铸锭	>6 m	0.5～5 mm
$10^{-3}\sim10^{0}$	低	一般铸件和铸锭	0.2～6 m	50～500 μm
$10^{0}\sim10^{3}$	中高	薄带、压铸、粉末	6～20 mm	5～50 μm
$10^{3}\sim10^{6}$	高	特细粉末，熔体拉拔和挤压成型	0.2～6 mm	0.5～5 μm
$10^{6}\sim10^{9}$	超高	雾化沉积，熔体喷铸，电子/激光非晶化	6～200 μm	0.05～0.5 μm

一般凝固过程，冷却速度大概在 1 K/s 以内，其过冷程度只有几度，当冷却速度达到每秒百万度级时，过冷程度可达 100 K 以上。由于快速凝固过冷程度深，合金体系大都偏离平衡状态，有可能形成许多非平衡组织，而合金的溶解度也可增加上百倍，或形成非晶相，图 2.3-85 近似地描述了不同冷却速度引起的合金组织结构的变化。

(2) 快速凝固组织变化特征

1) 扩大了固溶极限　如在 Al－Cu、Al－Si、Al－Mg 等合金中，所获得的固溶量不仅大大超过最大的平衡固溶极限，并且超过了平衡共晶点的成分，形成单相固溶体组织，如表 2.3-14 所示。

2) 晶粒度超细化　快速凝固合金的晶粒尺寸比常规合金低几个数量级，一般为 0.1～1.0 μm，如在 Al－50%Cu 合金中，可观察到 3 nm 的晶粒。当快速凝固的合金中出现第

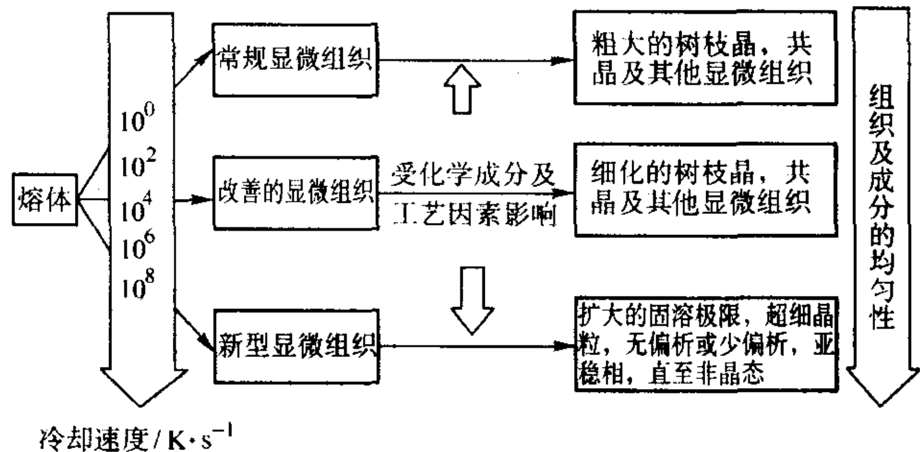


图 2.3-85　快速凝固引起的显微组织的变化

二相或夹杂物时，尺寸也相应细化，如快速凝固的奥氏体不锈钢析出的 MnS 夹杂物，尺寸比常规凝固低 2～3 个数量级。在某些共晶型合金中，当快速凝固不能遏制另一相析出时，往往会形成各种离异共晶和层片共晶，但层片的细密程度常规合金组织无法比拟（如 Al－Cu 合金，冷速 10^5 μ/s，片间距 0.06～0.08 μm）。

表 2.3-14　铝合金的固溶极限　%

合金系	平衡最大固溶极限 x_B	快速凝固固溶量 x_B	平衡共晶点成分 x_B
Al－Cu	2.53	18	17.3
Al－Si	1.78	16	11.3
Al－Mg	18.90	40	37.0
Al－Ni	<1	8	
Al－Cr	<1.2	6	
Al－Mn	<2	9	
Al－Fe	<1	6	
Al－Co	微量	5	

3) 少偏析和无偏析　在快速凝固的合金中，如处于平界面生长或溶质完全截留状态，则形成单相组织，在整个合金的宏观体积范围，不存在任何形式的偏析。实际应用的快速凝固合金，尽管显微组织仍由胞晶或胞状树枝晶组成，但由于固溶极限的扩大和晶粒的细化，胞、枝晶的分散度大大提高，显著改善了组织和化学均匀性。

4) 形成亚稳相　在快速凝固合金中，除了出现亚稳的过饱和固溶体外，还会形成其他的亚稳相。其晶体结构可能与平衡状态图上相邻的某一中间相的结构极为相似，因此可看作是快冷条件下，中间相的亚稳浓度范围扩大的结果。另一方面，也有可能形成某些在平衡相图上完全不出现的亚稳相。

5) 高的点缺陷密度　由于快速凝固，使液态金属中的大量“缺陷密度”较多地保存到了固态金属中，例如在快速凝固的 Fe－Cr20%－Ni25% 合金中，许多晶粒含有沿着 <100> r 方向分布的、相互平行的由空位环所形成的带。这些环的柏氏向量 $B = a/2<100>$，缺陷密度高于同成分正常凝固合金。

7.2.4　非晶合金

非晶合金是快速凝固应用的一个最好实例，非晶态合金是非平衡凝固的一种极限情况（冷却速度超过 10^9 K/s）。在足够高的冷却速度下，液态合金中可避免通常的结晶过程（形貌和生长），而过冷至某一温度（玻璃转化温度）以下时，其内部原子冻结在液态时所处位置附近形成非晶结构。由于是从液态连续冷却而形成的非晶固体，故也称为金属玻璃。金属玻璃并不是无定形的，它具有明确的原子程序有序结构，由于特殊的原子排列，而具有优良的综合性能，如良好的延性，易于加工。很多合金系列可形成金属玻璃，如 $Pd_{30}Si_{20}$、$Cu_{60}Zr_{40}$，$Ti_{60}Be_{40}$、$Fe_{80}B_{20}$、$Rb_{85}O_{15}$，$Ca_{70}Mg_{30}$，$La_{76}Au_{24}$ 等。

(1) 非晶形成原理

非晶的形成与熔体的黏度有较大关系。图 2.3-86 即为金属凝固时，黏度与温度的关系。图中虚曲线 1 表示在接近平衡条件（T_m－平衡熔点）进行的结晶过程，结晶完毕时。黏度上升至 10^{15} Pa·s 左右。曲线 2 表示一过冷的合金熔体，随着温度的下降（过冷度增大）黏度不断上升的情况。图中 T_g 为玻璃转化温度。显然，合金的熔点（或平衡相线温度）愈低，玻璃转化温度愈高，则愈容易在连续冷却的过程中避免结晶过程的产生，最后在 T_g 温度转变为玻璃。因此常用 T_g/T_m 的比值来衡量合金形成非晶态的倾向。一般情况，只要冷速足够大，任何一种合金熔体都有可能过冷至 T_g 而形成非晶合金。只是不同的合金，其形成非晶态的临界冷却速度 R 有较大差别。所以可用形成非晶结构的临界冷却速度来定量描述某合金形成玻璃态的能力。Davies 在晶化转变动力学研究基础上作出了纯 Ni 及三种合金 TTT 时间－温度－结晶曲线图（如图 2.3-87）。由图可知在几种合金中，由于液相线温度降低，及玻璃转化温度的提高，其结晶起始曲线与纯镍相比大大向右推迟了。图中虚线是由 TTT 曲线计算出来的 CCT 曲线。非晶态的临界冷却速度应由刚能避开 CCT 曲线的鼻子部分的冷却曲线来代表。

对硅酸盐和许多有机聚合物，鼻部位置的时间可以是几个小时，甚至几天，而纯金属所需时间小于 1 μs。一般将临界冷却速度 $R_c \leqslant 10^6$ K/s 的合金列为容易非晶化的合金。这些合金在旋铸等方式快速凝固后，可形成厚度大于 15～20 μm的非晶条带。一些工业价值较大的 Fe、Ni、Co 基非晶合金，临界过冷度大都在 3×（10^4～10^6）K/s 的范围内。

(2) 非晶合金的性能

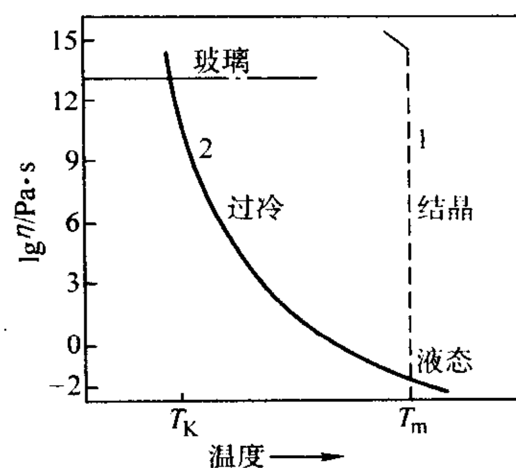


图 2.3-86　黏度与温度的关系

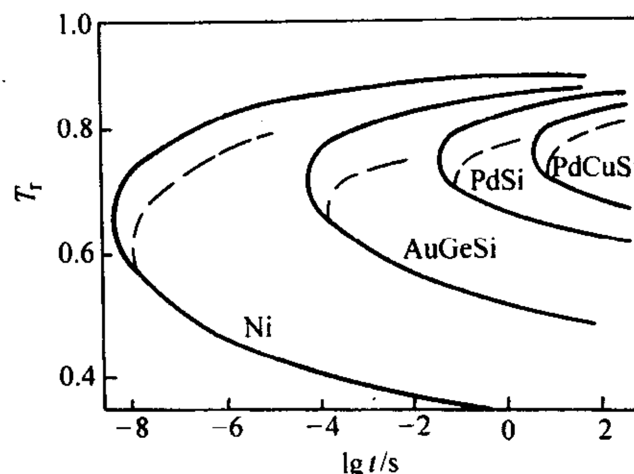


图 2.3-87　TTT 曲线（实线）RCCT 曲线（虚线）
转变开始线 $x = 10^{-6}$
合金：$Au_{78}Ge_{14}Si_8$，$Pb_{82}Si_{18}$，$Pb_{78}Cu_6Si_{16}$

金属玻璃具有一系列优良性能。在力学性能方面，它有极高的强度和硬度，如 $Fe_{80}B_{20}$ 玻璃的屈服点 σ_s 可达3 626 MPa，硬度值 HV 为 10 790 MPa，弹性模量 E 为165 730 MPa。与晶态合金比，金属玻璃的强度已接近理论值。在拉伸时，金属玻璃只表现出较小的伸长率（1.5%～2.5%），但在压缩时却表现出很高的塑性。它的撕裂能也比一般晶态合金高，表明它在具有高强度的同时也有较好的韧性。

由于非晶态合金没有晶界，也没有位错、夹杂物相等显微缺陷，因此铁、钴、镍基的金属玻璃具有十分良好的软磁性能，它们的铁心损耗仅为晶态合金的几分之一，是优良的变压器铁心，录音磁头及多种磁性器件材料。由于非晶态合金具有很小直至为零的电阻温度系数，因而可成为标准电阻及磁泡存储材料。

金属玻璃（如 $Fe_{70}Cr_{79}P_3C_7$）还具有极好的耐蚀性能，可作为最优秀的不锈钢。近年来发现非晶态合金的表面具有很高的化学活性和对化学反应的选择性，使金属玻璃有可能成为一种新型的催化剂及电极材料。此外，非晶合金还是有希望的储氢及超导材料。

编写：魏伯康（华中科技大学）

第4章　液态材料流动问题的分析方法

液体材料成形技术的应用十分广泛，包括液态金属砂型铸造、金属液态模锻、铝镁合金压铸、复合材料液态挤压成形、玻璃材料压制与吹制成形、高分子材料的注射与挤出成形等。在液体材料成形过程中所发生的各种物理和化学变化，都是在材料流动状态下进行和完成的，因而了解材料流动规律是分析其成形特点的基础。成形中的液态材料属于流体的范畴，材料的宏观运动规律是流体力学的研究对象，也是本章所讨论的主要内容。为叙述方便，本章中的液态材料统称为流体。

液体材料成形中的流动问题是多种多样的。本章主要介绍流体流动问题的基础理论与分析方法，包括影响流动的基本因素、处理流动问题的基本途径及流动分析的常用经典理论。对研究流动问题所用的方法有了一般了解之后，在处理具体的材料成形问题时，就可结合所研究的对象，抓住主要因素，略去次要因素，选择适当的方法。

1　流体的连续介质假设

从微观角度来看，流体和其他物体一样，也是由大量分子组成的。这些分子总是不停地、杂乱无章地运动着，分子之间存在间隙。因此，流体实际上并非是连续充满空间的物质。如果从分子运动入手来研究流体流动的规律，显然将十分困难，甚至难以进行。材料流动分析是研究在外力（如重力、压力、摩擦力等）作用下流体平衡和运动的规律，所研究的是大量分子的平均行为。另外，流动分析所研究的实际工程尺寸要比分子间距大得无法比拟。因此，在液态材料流动问题的研究中，将实际的由分子组成的结构用一种假想的流体模型——流体微团来代替。流体微团由足够数量的分子组成，连续充满它所占据的空间，彼此间无任何间隙，这就是1753年由欧拉首先建立的"连续介质模型"。这样就不用再去研究流体分子的运动了，而是研究模型化的连续流体介质。研究对象的转变，使我们可以将描述流体流动的一系列参数，如压强、速度、加速度、密度等在绝大多数情况下看成是连续分布的，从而可以把它们看作是坐标和时间的连续可微函数，使得在流动分析的研究中可以引用微分方程等强有力的数学工具。从所建立的流体力学基本理论与实际工程应用结果来看，引入连续介质模型是完全合理的，也是非常必要的。

但是，也要指出，使用连续介质模型也有一定的范围，在某些特殊流动的研究中，它并不适用。当所研究的工程实际尺寸与分子的自由行程为相同或相近的数量级时（如在某些微机电系统中），就不能应用连续介质作为其研究模型了，必须从微观着手进行研究。

2　影响流动特性的基本因素

影响液态材料流动问题的因素很多，加以归类，大致有：①材料的物理属性，如黏性、表面张力等；②流动空间的几何特征，如形状、尺寸等；③外界条件，如流动界面上的作用力、作用时间、流量等。下面分别加以简单说明。

2.1　材料的物理属性

流体受力将改变初始的运动状态，使流体微团发生平移、旋转和变形，力与运动、变形之间的关系，因流体的物理属性而异。如果说力是起因，运动和变形是结果，那么两者之间的联系则受流体属性的制约。在流体的物理属性之中，黏性尤为重要，当然，密度、压缩性、表面张力等也是需要注意的。

1）密度　密度是流体介质的重要属性之一，其定义为单位体积的质量，它表明材料在空间某点质量的密集程度。当流体作加速运动时，流体的密度与加速度的乘积常称为惯性力，当惯性力起重要作用时，流动问题的计算求解往往比较复杂。重力（密度与重力加速度的乘积）是作用于流体上最常见的质量力，对于有自由表面的运动颇为重要。

2）压缩性　流体密度随温度与压力而变化的特性称为流体的压缩性。实际流体都是可压缩的，但液体的压缩性一般来说很小，因此通常可作为不可压缩流体处理，但在很高的压力下，液体体积将明显减小，是否考虑流体的压缩性需视具体情况而定。根据压缩性，可将流体划分为两大类，一类是不可压缩流体，其密度为常数；另一类是可压缩流体，其密度是变化的。

3）黏性　流体黏性是影响流体运动特性的最重要的物理性质。设在两块的平行平板中间充满着不可压缩流体（如图2.4-1所示），下板固定，上板以均匀速度 U 做平行于下板的移动。此时，板间流体处于流动状态，且紧靠上板的流体薄层持有平板的速度，即流体与壁面之间无相对速度，称之为无滑动现象。随着离开上板，速度减小，至下板表面，速度为零。板间不同流体层具有不同的速度，形成如图所示的速度分布。流体作宏观运动的同时，流体层间存在随机的分子热运动。速度快的流体层（A 层）分子进入速度慢的流体层（B 层），使 B 层流体加速；而速度慢的流体层（B 层）分子进入速度快的流体层（A 层），使 A 层流体减速，从而引起流体层间的动量传递。根据牛顿第二定律，在相邻流体层接触面上必定存在着切向作用力（称为剪切力），其方向对于 A 层流体，与运动方向相反（阻力）；对于 B 层流体，则与运动方向相同（拉力）。这就是流体的内摩擦作用。流体的黏性正是内摩擦的表现，它是分子随机运动的动量交换和分子间引力作用的结果。

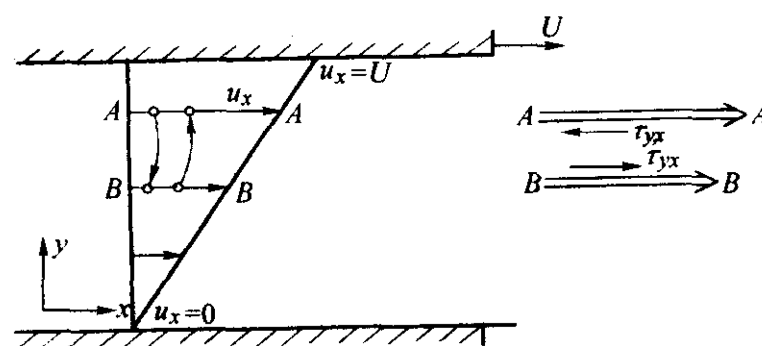


图2.4-1　平行平板间流体的直线运动

从动量传递的分析可以推知，流体中任何面上的切向力或单位作用面上的切向力即剪切应力都是作用面附近分子运动及相互作用的结果。如果该面附近各流体层速度相等，则剪切应力为零；各流体层速度不等则将产生剪切应力。流体层之间的速度变化常用单位距离上的速度变化率——速度梯度表示。流体作平行直线运动时，相邻流体层之间的剪切应力 τ_{yx} 与该处垂直于流动方向的速度梯度 $\frac{du_x}{dy}$ 成正比，即

$$\tau_{yx} = \mu \frac{du_x}{dy} \tag{2.4-1}$$

式中 μ 为动力黏度，是流体的物性常数。在工程计算中，常出现黏度与密度的比值，以 ν 表示，$\nu = \frac{\mu}{\rho}$，由于 ν 具有

运动学因次，故称为运动黏度。

2.2　流动空间的几何特征

流体运动总是在一定的几何空间内发生，对于流动现象，无论是进行物理分析还是数学描述，确立运动方程，提出边界条件等，都与流动空间的几何条件有关。因而，流动空间的几何特征是影响流体运动特征的又一个重要因素。

流动空间的几何特征通常包括它的形状、大小、方位、表面状况等。描述有关流动的物理量如果只依赖于一个坐标，则此运动称为一维运动；依赖于两个坐标称为二维运动；依赖于三个坐标则称为三维运动。在分析流动时，总是从几何形状出发，力求找到物理上的依据，以便将三维运动简化成二维甚至一维。

2.3　流体运动的动力学条件

在流体上施加作用力，可使静止的流体发生运动，或改变运动流体的速度大小和方向，这是影响流动的动力学条件。在材料成形中造成流体运动的方式很多，如给流体以压力使其在压力差的推动下运动、以边界运动或是流体中物体的运动来推动流体等。力是改变运动状态的基本原因，不同的受力情况会导致流体运动的不同规律。

3　研究流动问题的基本途径

为了得到流动过程中流速、压力等参数的变化规律，通常运用动量、能量及质量守恒原理建立流体运动的基本方程，并将影响流动的诸因素用数学关系式联系起来，这类方程的建立和求解，往往是研究流动问题的关键。

工程上的流动一般比较复杂，需要从不同的方面进行考察，如有时着眼于系统总体的变化，而有时则需要分析系统内部发生的细节。因而分析研究流体运动，建立描述流动的方程就有两种基本方法：总体衡算法和微元体衡算法。

总体衡算法：研究流体运动过程总体的性能，而不涉及流体内部的详细结构，运用物理学上守恒原理分析流体流动，解决过程前后的总变化问题。

微元体衡算法：考察系统内部流体微团运动以及在空间各点上有关物理量的相互关系，对微元体运用守恒原理，解决流场内流动参数的分布问题。

在解决工程问题时，这两种方法各有其适用的场合，是相辅相成的。此外，还经常应用实验方法观察流动特征。

3.1　总体衡算法

应用总体衡算法分析流体流动过程，就是对所考察的对象进行总衡算。为便于得到流动参数间的相互关系，总体衡算法引入了很重要的简化假定，即只考虑流体在主运动方向上的流动参数的变化，即限于一维流动。因此，对系统进行总衡算时可不涉及复杂的数学运算，无需了解过程的内在机理，可以迅速而简便地得到有关流动参数的变化规律，并给出可能存在的函数形式及参数的数量级，这对解决工程问题是很有用的。

1）控制体和控制面　运用守恒原理对运动流体做总体衡算时，需要规定明确的衡算对象，因为流体本身不能保持其位置和形状，通常只能在空间选择固定的体积即控制体作为考察对象。控制体的形状和大小可以根据流动情况和位置边界任意选取，但一经选定，其形状和位置相对于坐标系统就不再改变。控制体所围空间区域的边界面称为控制面，控制面必须是封闭的。

质量、动量和能量可以通过控制面输入和输出。于是，守恒原理可表达为：控制体内某特征量的变化速率等于该特征量进入与离开控制体的速率之差。下面以管内流体流动为例（如塑料在注射模流道系统中的流动），进行质量、能量及动量的总体衡算，并给出数学表达式。

2）质量守恒　在管流系统内，取一有限长管段，其两端的截面分别为 A_1、A_2，对应于该面积上的流体密度分别为 ρ_1、ρ_2，速度均匀且为 U_1、U_2，所选定的控制体如图2.4-2虚线所示，即为该管段内壁面的流体边界及截面 A_1 和 A_2 所包围的区域。

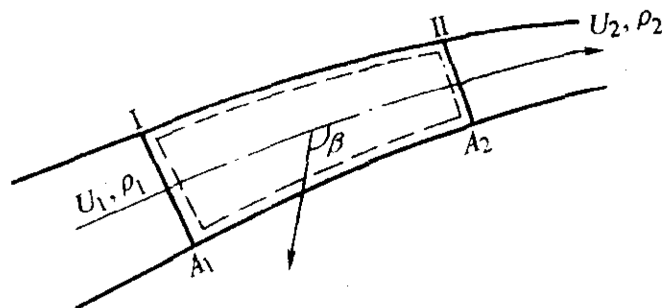


图2.4-2　一维流动

按照质量守恒原理，单位时间内流入控制面 A_1 和流出控制面 A_2 的质量之差，应等于单位时间内该控制体积内所含物质的增量，即

$$\rho_1 U_1 A_1 - \rho_2 U_2 A_2 = \frac{\mathrm{d}\int_{x_1}^{x_2} \rho A \mathrm{d}x}{\mathrm{d}t} \tag{2.4-2}$$

这就是非定常管流质量衡算的数学表达式，亦称为一维流动时的连续性方程。定常流动时，上式右端为零，则有

$$\rho_1 U_1 A_1 = \rho_2 U_2 A_2 \tag{2.4-3}$$

对于不可压缩流体，ρ 为常数，则 $U_1 A_1 = U_2 A_2$，即

$$\frac{U_2}{U_1} = \frac{A_1}{A_2} \tag{2.4-4}$$

在上式中 U_1、U_2 为相应截面上的均匀流速。实际上，由于流体具有黏性，截面上的速度是不均匀的，形成某种速度分布。在工程计算中，通常采用平均流速 $\overline{U}$ 近似表达实际上具有速度分布的流体流动速度，从而简化为一维流动。平均流速 $\overline{U}$ 按体积流量相等的原则定义，即 $\overline{U}A = \int_A u\mathrm{d}A$，所以平均流速为 $\overline{U} = \dfrac{\int_A u\mathrm{d}A}{A}$。

3）能量守恒　根据热力学第一定律（能量守恒定律），在某一过程中系统内总能量变化等于系统吸收的热量与所作功之差，即

$$\Delta E = Q - W \tag{2.4-5}$$

式中，Q 为单位质量流体所吸收的热量；W 为单位质量流体所作的功；E 为单位质量流体所具有的总能量。

在流动过程中，总能量主要包括流体动能（$\frac{U^2}{2}$）、位能（gZ）及内能（e）三种，即

$$E = e + \frac{U^2}{2} + gZ \tag{2.4-6}$$

如果流体的流动过程中不涉及系统与外界的热交换，则所进行的能量衡算为机械能衡算。对于等温系统无外功输入的不可压缩理想流体的定常流动，可得如下的基本关系式

$$\frac{\Delta U^2}{2} + g\Delta Z + \frac{\Delta P}{\rho} = 0 \tag{2.4-7}$$

即

$$\frac{U^2}{2} + gZ + \frac{P}{\rho} = 常数 \tag{2.4-8}$$

上式即为著名的柏努利方程。柏努利方程用于黏性流体时需作修正，由于黏性造成了截面上的速度分布，因此必须采用平均动能替代式，此外，还应计入黏性摩擦造成的机械能损耗 h_f，所以上式应修正为

$$\left(\frac{\overline{U_1^2}}{2}\right) + gZ_1 + \frac{P_1}{\rho} = \frac{\overline{U_2^2}}{2} + gZ_2 + \frac{P_2}{\rho} + h_f \quad (2.4\text{-}9)$$

4）动量守恒　动量守恒定律指出，物体的动量随时间的变化率等于作用在物体上外力的总和。在流动系统中对控制体做动量衡算的原则是，作用在控制体上所有外力的总和等于流体通过控制体的动量变化率。

作用在控制体上的外力通常包括重力 F_g、压力 F_p、摩擦力 F_f 以及控制面上所受到的其他外力 F_n。考察如图 2.4-2 所示的管内流动，设 x 轴与重力方向夹角为 β，摩擦力为 F_{xf}，则 x 方向动量衡算关系可写为

$$P_1A_1 - P_2A_2 - F_{xf} + \int_{x_1}^{x_2} \rho g \cos\beta A \mathrm{d}x = [\rho_2 U_2 A_2 \cdot U_2] - [\rho_1 U_1 A_1 \cdot U_1] + \frac{\mathrm{d}P_x}{\mathrm{d}t} \quad (2.4\text{-}10)$$

若是水平管内的定常流动，$\beta = 90°$，则

$$P_1A_1 - P_2A_2 - F_{xf} = \rho_1 U_1 A_1 (U_2 - U_1) = \rho_2 U_2 A_2 (U_2 - U_1) \quad (2.4\text{-}11)$$

动量守恒和机械能守恒是流体流动所必须遵循的规律，都体现了流动中参数变化的相互关系，但应用场合各有不同。应用机械能守恒定律时，首先要确定机械能的损耗 h_f，而动量守恒定律则不涉及能量损耗。因而，在不能有效地应用机械能衡算时，可以应用动量守恒定律，然而必须正确确定所有的作用力。

3.2　微元体衡算法

在解决许多工程问题时，往往需要了解空间各点速度的大小和方向以及流场的详细结构（流型）。为此需要取微元体研究力和流体运动之间的关系，并在微小空间范围内运用守恒原理，建立流体运动的微分方程式。根据质量守恒定律导出连续性方程式，由动量守恒定律导出流体运动微分方程式。

1）连续性方程　在流场中取一微元体积，如图 2.4-3 所示。单位时间沿 x 方向从微元体流出与流入的质量差为

$$\left[\rho u_x + \frac{\partial(\rho u_x)}{\partial x}\mathrm{d}x\right]\mathrm{d}y\mathrm{d}z - \rho u_x \mathrm{d}y\mathrm{d}z = \frac{\partial(\rho u_x)}{\partial x}\mathrm{d}x\mathrm{d}y\mathrm{d}z \quad (2.4\text{-}12)$$

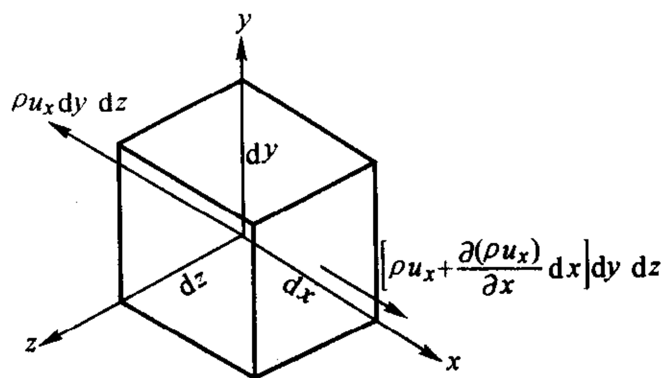


图 2.4-3　质量衡算微元体

同理可得沿 y、z 两个方向单位时间内由微元体流出与流入的流体质量差，计算单位时间内整个微元体流出与流入的流体质量差为

$$\frac{\partial(\rho u_x)}{\partial x}\mathrm{d}x\mathrm{d}y\mathrm{d}z + \frac{\partial(\rho u_y)}{\partial y}\mathrm{d}x\mathrm{d}y\mathrm{d}z + \frac{\partial(\rho u_z)}{\partial z}\mathrm{d}x\mathrm{d}y\mathrm{d}z = \left[\frac{\partial(\rho u_x)}{\partial x} + \frac{\partial(\rho u_y)}{\partial y} + \frac{\partial(\rho u_z)}{\partial z}\right]\mathrm{d}x\mathrm{d}y\mathrm{d}z \quad (2.4\text{-}13)$$

微元体中流体质量因流体密度随时间的变化而变化，微元体中流体质量的增量为：$\frac{\partial\rho}{\partial t}\mathrm{d}x\mathrm{d}y\mathrm{d}z$。根据质量守恒定律，单位时间内流出和流入微元体的流体质量差应等于单位时间内该微元体中物质的增量，故可得

$$\frac{\partial\rho}{\partial t} + \frac{\partial(\rho u_x)}{\partial x} + \frac{\partial(\rho u_y)}{\partial y} + \frac{\partial(\rho u_z)}{\partial z} = 0 \quad (2.4\text{-}14)$$

上式即为三维流动时流体的连续性方程，是流体运动的速度场所必须满足的条件。

对于定常流动，$\frac{\partial\rho}{\partial t} = 0$，式（2.4-14）为

$$\frac{\partial(\rho u_x)}{\partial x} + \frac{\partial(\rho u_y)}{\partial y} + \frac{\partial(\rho u_z)}{\partial z} = 0 \quad (2.4\text{-}15)$$

对于不可压缩流体，ρ 常数，$\frac{D\rho}{Dt} = 0$，上式成为

$$\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z} = 0 \quad (2.4\text{-}16)$$

这是连续性方程最常用的形式。

2）运动方程　将牛顿运动定律应用于流体运动，分析受力和运动加速度之间的关系，可以得到流体运动方程式。以黏性流体运动为例，作用在黏性流体微元体上的力比较复杂，如图 2.4-4 所示。

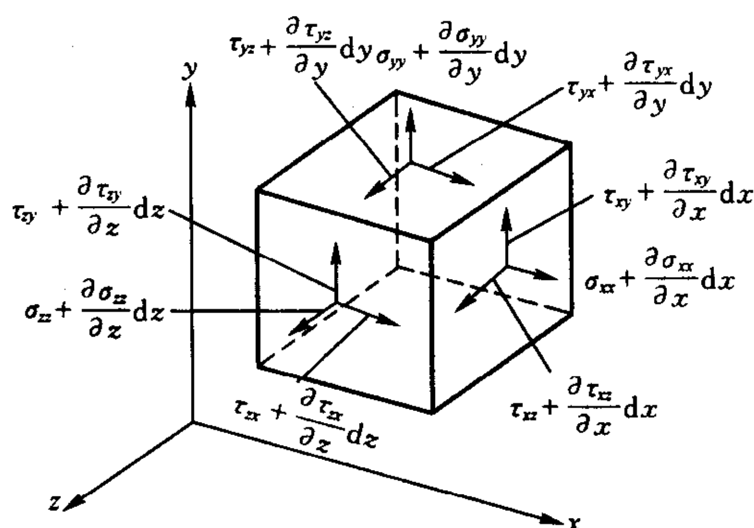


图 2.4-4　黏性流体中的微元体

根据微元体所受的力及其加速度的关系，结合考虑因黏性而引起的法向力及切向应力对运动的影响，可以得到以应力表示的黏性流体运动方程式：

$$X + \frac{1}{\rho}\left(\frac{\partial\sigma_{xx}}{\partial x} + \frac{\partial\tau_{yx}}{\partial y} + \frac{\tau_{zx}}{\partial z}\right) = \frac{Du_x}{Dt} \quad (2.4\text{-}17a)$$

$$Y + \frac{1}{\rho}\left(\frac{\partial\tau_{xy}}{\partial x} + \frac{\partial\sigma_{yy}}{\partial y} + \frac{\partial\tau_{zy}}{\partial z}\right) = \frac{Du_y}{Dt} \quad (2.4\text{-}17b)$$

$$Z + \frac{1}{\rho}\left(\frac{\partial\tau_{xz}}{\partial x} + \frac{\partial\tau_{yz}}{\partial y} + \frac{\partial\sigma_{zz}}{\partial z}\right) = \frac{Du_z}{Dt} \quad (2.4\text{-}17c)$$

式中 X、Y、Z 分别为 x、y、z 三个方向上的质量力分量。

上述以应力表示的运动方程适用于任何黏性流体，但方程数目与未知量个数不等。方程组中有三个速度分量，六个应力分量（由于应力对称，有三个应力分量不是独立的），但方程数仅四个，即三个运动方程、一个连续性方程。因而，方程是不封闭的，必须补充方程式。通常从假定应力分量和速度梯度有线性关系着手。将牛顿黏性定律推广到一般情况，可得

$$\tau_{xy} = \tau_{yx} = \mu\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right) \quad (2.4\text{-}18a)$$

$$\tau_{xz} = \tau_{zx} = \mu\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right) \quad (2.4\text{-}18b)$$

$$\tau_{zy} = \tau_{yz} = \mu\left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}\right) \quad (2.4\text{-}18c)$$

$$\sigma_{xx} = -P + 2\mu\frac{\partial u_x}{\partial x} \quad (2.4\text{-}18d)$$

$$\sigma_{yy} = -P + 2\mu\frac{\partial u_y}{\partial y} \quad (2.4\text{-}18e)$$

$$\sigma_{zz} = -P + 2\mu\frac{\partial u_z}{\partial z} \quad (2.4\text{-}18f)$$

将上述方程式代入式（2.4-17），可导出不可压缩黏体流体运动方程式为：

$$\rho\left(\frac{\partial u_x}{\partial t}+u_x\frac{\partial u_x}{\partial x}+u_y\frac{\partial u_x}{\partial y}+u_z\frac{\partial u_x}{\partial z}\right)=$$
$$\rho X-\frac{\partial P}{\partial x}+\mu\left(\frac{\partial^2 u_x}{\partial x^2}+\frac{\partial^2 u_x}{\partial y^2}+\frac{\partial^2 u_x}{\partial z^2}\right)\quad(2.4\text{-}19a)$$
$$\rho\left(\frac{\partial u_y}{\partial t}+u_x\frac{\partial u_y}{\partial x}+u_y\frac{\partial u_y}{\partial y}+u_z\frac{\partial u_y}{\partial z}\right)=$$
$$\rho Y-\frac{\partial P}{\partial y}+\mu\left(\frac{\partial^2 u_y}{\partial x^2}+\frac{\partial^2 u_y}{\partial y^2}+\frac{\partial^2 u_y}{\partial z^2}\right)\quad(2.4\text{-}19b)$$
$$\rho\left(\frac{\partial u_z}{\partial t}+u_x\frac{\partial u_z}{\partial x}+u_y\frac{\partial u_z}{\partial y}+u_z\frac{\partial u_z}{\partial z}\right)=$$
$$\rho Z-\frac{\partial P}{\partial z}+\mu\left(\frac{\partial^2 u_z}{\partial x^2}+\frac{\partial^2 u_z}{\partial y^2}+\frac{\partial^2 u_z}{\partial z^2}\right)\quad(2.4\text{-}19c)$$

这一组方程称为奈维－斯托克斯方程，是黏性流体运动时牛顿第二定律的数学表达式，又称动量微分方程。方程左端是流体微元运动的加速度和质量的乘积，即运动的惯性力，第一项为局部加速度所产生的惯性力，其余项为迁移加速度所产生的惯性力。方程右端是作用于微元体上的外力：第一项为重力（质量力），第二项为压力（表面力），第三项为黏性力（表面力）。因此，也可将此方程看作是惯性力、重力、黏性力和压力这四种力的平衡关系式。

3）黏性流体运动时的初始条件和边界条件　运动方程表达了流体运动时力与运动的关系，描述了流动特性如速度、压力等物理量所服从的普遍关系。要确定某特定流动情况下有关各量的具体关系，还必须明确运动的初始情况以及运动所处空间的边界情况。从数学上说，就是要对方程给出初始条件和边界条件，即方程的定解条件。初始条件决定了所求函数在某一给定时刻的值。边界条件决定了所求函数在流动边界上的值，它们在不同的具体问题中是各不相同的。当流体沿不可穿透的固体表面运动时，紧靠固体壁的流体像贴在固体上一样，即沿表面无滑动。这是黏性流体与理想流体在流动上的一个重要区别。当壁面运动时，流体速度与壁面速度相等。在非固定相界面上（两种不互溶的液体或气液两相），速度不一定为零，但应满足切向分速度连续、法向分速度为零、应力分量连续等条件。

4）速度分布和压力分布的计算——运动方程的求解　速度或压力的分布表示了在给定瞬间，速度或压力随空间位置的变化。求解运动方程可以得到这种函数关系。但是，由于运动方程中含有未知量的乘积项，如 $u_x\frac{\partial u_x}{\partial x}$ 等非线性项，目前还无法求出非线性偏微分方程的一般解，求解只能在一些特定的情况下进行。

如果方程的非线性项消失，则能求得准确解，例如直圆管内的层流流动。当雷诺数很小，惯性力与黏性力相比可以忽略，则可得极慢运动（爬流）的近似解。当雷诺数很大时惯性力居于主要地位，应用边界层概念简化方程，亦能得到近似解。

近年来，计算机的应用日益广泛，计算流体力学发展迅速，为应用数值计算方法（如有限差分法、有限元法）求解运动方程奠定了基础。数值解法比解析法更能适应复杂的流动问题。

3.3　实验观测法

实验是研究流体运动最基本的方法。通过实验可以验证理论计算的结果，也可以探索新的流动现象。

考察流动所采用的测试技术分为两种类型：一是使用速度、压力、力矩等的测量仪器，定量测定这些物理量，如采用毕托管、热线风速仪和激光测速仪等测量速度；二是使流场中的流动情况可视化，以便直接观察或者摄影记录，称为流场显示技术。下面对流动可视化方法作简要介绍。

用肉眼直接观察流体运动往往难以得到结果。如果采取特殊措施，例如在液流中用颜色水示踪，则流动图像立即清晰可见，这就是流动显示。流动显示需要解决两个问题，一是流动如何可视？二是观察结果如何表示？

使流动成为可视的方法大致分为两大类：示踪法和光学法。示踪法是将示踪物加入流体并随之运动，通过观察示踪物，可以了解流体本身的运动。这种显示是间接的，因此需选择适当尺寸、密度的示踪粒子，力求示踪物对流体有良好的跟随性，以便能真实反映流体流动的实际情况。示踪法主要用于观察不可压缩流体的运动，包括壁面显示法、直接注入法、化学反应法、电控制法等。光学法主要用于观察可压缩流体的运动。流体密度的变化会引起折射率变化，利用这一关系，借助于光学技术则可显示流场。光学法主要包括纹影法和干涉法，其特点是对流场无干扰、精度高。

实验观察得到的流动结果常以流线、轨线及脉线表示出来。所谓流线是指流场中每点上的切线都和该点处速度方向重合的曲线。表示流体质点位置随时间变化的曲线称为轨线。质点运动的方向必定是轨线切线的方向，因此，当质点运动描绘出轨线时，轨线和经过该质点瞬时位置的流线相切。流线表示在给定瞬时每个质点的运动；轨线表示给定质点在每一瞬时的运动，对于定常运动，轨线和流线重合。

获取流线的方法，是将大量微小示踪粒子加入流场中，然后摄影记录下这些质点在极短时间内的位置变化，通过综合所有质点的运动途径则可得到流线。将有色液体加入液态材料中，可以观察到轨线或脉线，所谓脉线是指先后通过指定点的所有流体质点的连线。如果瞬时地加入有色液体，连续观察或取长的曝光时间摄影，这样得到的是轨线。

必须注意定常运动下流线、轨线、脉线是相同的，但非定常流动时则完全不同。

4　边界层理论

实际流体都具有黏性，实验表明，当流体绕任何物体运动时，在物体表面附近的薄层流体中，黏性的影响非常显著，这一薄层称为边界层。边界层概念没有严格的数学证明，但有实验基础。边界层理论是研究黏性流体运动规律的理论基础，可用于分析阻力产生的机理，计算流体的速度分布，进而应用于传热和传质计算。

4.1　边界层概念

流体绕物体流动时，因黏性而附着于物体表面，即在物体表面上无流体滑动，流体运动速度为零，并使表面附近的流体运动受阻，速度减慢。距表面一定距离之后，速度逐渐接近流体的来流速度（即与物体表面相接触之前的流体速度）。在 Re 值很大的情况下，速度从零达到来流速度所需要的距离很短，即这一发展在紧靠表面的薄层内完成，这就表明该薄层内沿物体表面法向存在很大的速度梯度，故物体受到流体的黏性力很大。即使流体的黏性系数很小，在边界层内黏性力与惯性力仍有同样的数量级，而在距表面相当远的边界层外，流体的速度基本上不再变化，速度梯度几乎为零，黏性影响就完全可以忽略。因此，沿壁面的流动可分为两个区域：在边界层之外，黏性力作用很小，可以作为理想流体流动；在边界层内，由于黏性力作用显著，必须以黏性流动处理。边界层的边缘是两个区域的分界线。

（1）边界层发展

为说明边界层的形成和发展，下面介绍绕平板流动的情况。如图2.4-5所示，远离平板前缘的上游，流体速度为 U_0，流体一经与固体表面接触，就黏附在表面上，速度为零。随着沿表面法向距离的增加，流体的速度逐渐增大。在板的前缘（$x=0$），边界层厚度为零。沿着流动方向，边界

层开始形成。随着 x 距离的增加，流过板上的流体在壁面影响下，沿壁面法向更多的流体被阻滞，因而使边界层增厚。

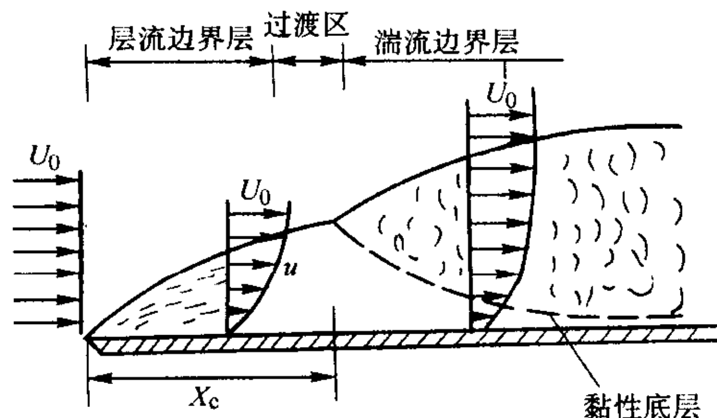


图 2.4-5　平板边界层

流体刚接触平板，边界层很薄，扰动在薄层内不易发展，不管来流是层流还是湍流，则该处边界层内的流动总是层流，称为层流边界层。当流体继续沿平板流动，只要板面足够长，经过一段距离后随着边界层厚度的迅速增加，扰动便会发展起来，边界层中流体运动由层流过渡到湍流，边界层厚度与上游层流时相比增加得更快，称为湍流边界层。但在湍流边界层内紧靠平板表面的薄层仍然具有层流特性，称为黏性底层。黏性底层的厚度较湍流边界层的厚度薄得多，但是很重要，因边壁上的速度梯度对于剪切应力的大小有着决定性的影响。

边界层由层流过渡为湍流取决于雷诺数的大小，流动状态发生转变的临界值为

$$Re_{x,cr} = \frac{U_0^x}{v} = 5\times10^5 \sim 2\times10^5 \qquad (2.4\text{-}20)$$

临界雷诺数的数值与来流中的湍流程度、表面粗糙度以及板前缘的形状等因素有关。由层流边界层向湍流边界过渡时，并不是突然转变的，其间有段过渡区。

(2) 边界层厚度的估计

在边界层中流体速度从壁面上的零值发展到边界层外边缘近似为来流速度 U_0，通常将这段法向距离定义为边界层厚度 δ。严格说来，边界层与外流之间颇难划定一条明确的界限，于是规定：$y=\delta$ 时 $u=0.99U_0$。

流体绕过几何形状简单的物体时，边界层厚度可以通过计算得到。一般可作近似估计，估计的依据是边界层中黏性力与惯性力具有相同数量级，这种近似估算是了解边界层厚度数量级的最简便的方法。

对于层流，由于 $\tau=\mu\frac{\mathrm{d}u_x}{\mathrm{d}y}$，单位体积流体的黏性力则为$\frac{\partial\tau}{\partial y}=\mu\frac{\partial^2 u_x}{\partial y^2}$，而单位体积流体的惯性力为 $\rho u_x\frac{\partial u_x}{\partial x}$。下面估计这两种力的数量级。式中变量 y 限制在边界层之内，即 $0\leqslant y\leqslant\delta$，$y$ 的数值与 δ 具有同一量级，$y\sim\delta$。变量 x 从零到L，$x\sim L$，L 是物体的特征长度。速度分量 u_x 从零至 U_0，于是 $u_x\sim U_0$。通过以上的分析可得黏性力的量级为 $u\frac{U_0}{\delta^2}$，惯性力的量级为 $\rho\frac{U_0^2}{L}$，由于两者量级相同，即有

$$u\frac{U_0}{\delta^2}\sim\rho\frac{U_0^2}{L} \qquad (2.4\text{-}21)$$

经整理可得

$$\frac{\delta}{L}\sim\frac{1}{\sqrt{ReL}} \qquad (2.4\text{-}22)$$

根据上述估计可知，δ 与 Re 数的平方根成反比。

4.2　边界层计算

边界层计算是边界层理论的重要组成部分，但为避免冗长的数学演算，这里对边界层微分方程的建立和求解仅仅给出若干主要步骤，详细的数学分析，可参考黏性流体力学方面的专著。

(1) 边界层方程

黏性流体运动方程描述了边界层内的流体运动，对于沿平板的定常二维流动，运动方程和连续方程为

$$u_x\frac{\partial u_x}{\partial x}+u_y\frac{\partial u_x}{\partial y}=-\frac{1}{\rho}\frac{\partial P}{\partial x}+v\left[\frac{\partial^2 u_x}{\partial x^2}+\frac{\partial^2 u_x}{\partial y^2}\right] \qquad (2.4\text{-}23a)$$

$$u_x\frac{\partial u_y}{\partial x}+u_y\frac{\partial u_y}{\partial y}=-\frac{1}{\rho}\frac{\partial P}{\partial y}+v\left[\frac{\partial^2 u_y}{\partial x^2}+\frac{\partial^2 u_y}{\partial y^2}\right] \qquad (2.4\text{-}23b)$$

$$\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}=0 \qquad (2.4\text{-}23c)$$

相应的边界条件为

壁面上：$y=0$，$u_x=0$，$u_y=0$；

边界层外缘：$y=\delta$，$u_x=U_0$，$\frac{\partial u_x}{\partial y}=0$；

采用量级比较方法可知，方程 (2.4-23a) 等号左边两惯性项的量级相当，为 $O\left(\frac{U_0^2}{L}\right)$，$O$ () 符号表明括号内的数值为数量级。等号右边两阶导数项量级则差别很大，$\frac{\partial^2 u_x}{\partial x^2}\sim\frac{U_0}{L^2}$，$\frac{\partial^2 u_x}{\partial y^2}\sim\frac{U_0}{\delta^2}$，两者相比，$\frac{\partial^2 u_x}{\partial x^2}$项可以忽略。在方程 (2.4-23b) 中，左端各项数量级为 $O\left(\frac{U_0^2\delta}{L^2}\right)$，对于右端的粘性项，通过数量级比较，可得$\frac{\partial^2 u_y}{\partial x^2}\ll\frac{\partial^2 u_y}{\partial y^2}$，后者的量级与左端相同，余下的 $\frac{1}{\rho}\frac{\partial P}{\partial y}$ 项，其量级也不会大于 $O\left(\frac{U_0^2\delta}{L^2}\right)$，由于大雷诺数下$\frac{\delta}{L}$是小量，故可认为$\frac{\partial P}{\partial y}=0$，即边界层中沿 y 方向无压力变化。由此可知，流体中的压力取决于外流流型，它仅仅是纵向距离 x 的函数。

经量级比较，简化后沿平壁流动的运动方程和连续方程为

$$u_x\frac{\partial u_x}{\partial x}+u_y\frac{\partial u_x}{\partial y}=-\frac{1}{\rho}\frac{\partial P}{\partial x}+v\frac{\partial^2 u_x}{\partial y^2} \qquad (2.4\text{-}24a)$$

$$\frac{\partial P}{\partial y}=0 \qquad (2.4\text{-}24b)$$

$$\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}=0 \qquad (2.4\text{-}24c)$$

这组方程称为边界层方程。它比原来的二维奈维-斯托克斯方程大大简化，可以求解的范围有所扩大。下面就平板上的流动给出它的解答。

(2) 布拉修斯解

对于沿平板的流动，纵向压力梯度$\frac{\partial P}{\partial x}=0$，布拉修斯于 1908 年首先得到这一方程的解。求解的关键在于，假定边界层中不同截面上的速度分布有相似性，即在任意 x 处的速度分布图形，通过适当的坐标变化，就可与另一 x 处的速度图形相合。选择外流速度 U_0 及边界层厚度 δ 分别作为速度及坐标 y 的尺度，使速度和距离都成为无因次量，即$\frac{u_x}{U_0}$和$\frac{y}{\delta}$，由于 δ 随 x 变化。因此，不同 x 处 y 方向距离无因次化的比尺实际上是变化的。这样，在边界层内无因次速度$\frac{u_x}{U_0}$只是无因次距离$\frac{y}{\delta}$的函数

$$\frac{u_x(y)}{U_0}=\Phi\left[\frac{y}{\delta(x)}\right] \qquad (2.4\text{-}25)$$

具有这种性质的解，称为相似解。

由于已知边界厚度 $\delta \sim \sqrt{\frac{\nu x}{U_0}}$，于是令无因次距离

$$\eta = y\sqrt{\frac{U_0}{\nu x}} \tag{2.4-26}$$

引进这一新的无因次变量，使原来的两个独立变数 (x, y) 减少为一个 η。对于方程式（2.4-24a）进行变换，则使偏微分方程变成常微分方程

$$2\frac{d^3 f(\eta)}{d\eta^3} + f(\eta)\frac{d^2 f(\eta)}{d\eta^2} = 0 \tag{2.4-27}$$

为求解上述三阶常微分方程，可将 $f(\eta)$ 表示为 η 的幂级数。满足边界条件 $\eta = 0$ 时，$u_x = u_y = 0$ 的解具有如下形式

$$f(\eta) = \frac{\alpha\eta^2}{2!} - \frac{\alpha^2}{2}\frac{\eta^5}{5!} + \frac{11}{4}\frac{\alpha^3\eta^8}{8!} - \frac{375}{8}\frac{\alpha^4\eta^{11}}{11!} + \cdots \tag{2.4-28}$$

式中 α 为数值常数，由边界层边缘的边界条件 $y \to \infty$，$u \to U_0$ 可得 $\alpha = 0.332$。

无因次速度 $\frac{u}{U_0}$ 与 η 的关系可表示为

$$\frac{u_x}{U_0} = \alpha\eta - \frac{\alpha^2}{2}\frac{\eta^4}{4!} + \frac{11}{4}\frac{\alpha^3\eta^8}{7!} - \frac{375}{8}\frac{\alpha^4\eta^{10}}{10!} + \cdots \tag{2.4-29}$$

壁面上的表面摩擦力 τ_w 有下式

$$\tau_w = \mu\left(\frac{\partial u_x}{\partial y}\right)_{y=0} = \mu\left(\frac{U_0}{\nu x}\right)^{1/2}\left(\frac{\partial u_x}{\partial \eta}\right)_{\eta=0} = \mu U_0\alpha\left(\frac{U_0}{\nu x}\right)^{1/2} = 0.332\mu U_0\left(\frac{U_0}{\nu x}\right)^{1/2} \tag{2.4-30}$$

表面摩擦系数

$$C_D = \frac{\tau_w}{1/2\rho U_0^2} = 0.664\left(\frac{U_0 x}{\nu}\right)^{-1/2} \tag{2.4-31}$$

并可根据边界层厚度的定义得到

$$\delta \approx 5.0\sqrt{\frac{\nu x}{U_0}} \text{ 即 } \frac{\delta}{x} \approx 5.0Re_x^{-1/2} \tag{2.4-32}$$

计算结果表明，摩擦阻力正比于速度的3/2次方，而极慢运动时则与速度的一次方成正比例。此外，阻力正比于长度的平方根，而不是长度的一次方。这可解释为因为平板后半部边界层较厚，壁上的剪切应力小，致使板的后半部所占总阻力的比例较前半部少。

(3) 边界层动量积分关系式

由于普朗特边界层微分方程中仍包含有非线性项，致使准确求解十分困难。为此，需寻求边界层的近似计算方法，这对湍流边界层尤为重要。边界层的近似计算以卡门导出的边界层动量积分关系式为基础。

如图2.4-6所示，在边界内取一微元体，投影至 xoy 平面的 $ABCD$，令垂直纸面的距离为1个单位。

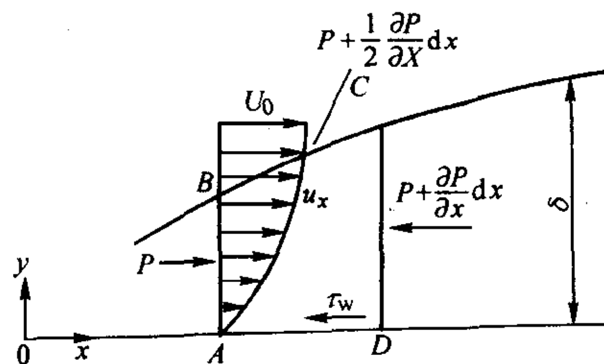


图2.4-6　边界层内控制体分析

通过 AB 平面进入微元体的质量流率和动量速率分别为 $\rho\int_0^\delta u_x dy$ 和 $\rho\int_0^\delta u_x^2 dy$，通过 CD 平面由微元体流出的质量流率和动量速率分别为 $\rho\int_0^\delta u_x dy + \rho\frac{\partial}{\partial x}(\int_0^\delta u_x dy)dx$ 和 $\rho\int_0^\delta u_x^2 dy + \rho\frac{\partial}{\partial x}(\int_0^\delta u_x^2 dy)dy$。由 BC 面进入的质量流率必然等于分别由 AB、CD 二侧流进和流出的质量流率差 $\rho\frac{\partial}{\partial x}(\int_0^\delta u_x dy)dx$，通过 BC 平面带入微元体的动量速率为 $\rho U_0\frac{\partial}{\partial x}(\int_0^\delta u_x dy)dx$。

因此，进出微元体总的动量变化率为

$$\rho dx\frac{\partial}{\partial x}(\int_0^\delta u_x^2 dy) - \rho U_0 dx\frac{\partial}{\partial x}(\int_0^\delta u_x dy) = \rho dx\frac{\partial}{\partial x}\left[\int_0^\delta (u_x - U_0)u_x dy\right] \tag{2.4-33}$$

作用于微元体的外力有沿壁面 AD 的切应力 τ_w 以及作用于 AB、CD、BC 面上的压力，即

$$P_{AB} = P\delta \tag{2.4-34a}$$

$$P_{CD} = -\left(P + \frac{\partial P}{\partial x}dx\right)(\delta + d\delta) \tag{2.4-34b}$$

$$P_{BC} = \left[P + \frac{1}{2}\frac{\partial P}{\partial x}dx\right]d\delta \tag{2.4-34c}$$

忽略二阶小量，上述各项外力的综合为 $-\left(\tau_w + \delta\frac{\partial P}{\partial x}\right)dx$。根据动量定理，作用于微元体的外力矢量和等于出入微元体的动量变化率，因此得

$$\rho dx\frac{\partial}{\partial x}\left[\int_0^\delta (u_x - U_0)u_x dy\right] = \tau_w + \delta\frac{dP}{dx} \tag{2.4-35}$$

上式即为卡门边界层动量积分关系式。该式表明，在沿流动方向无压力梯度时，边界层中的动量损失纯系壁面处摩擦应力所致。在上述推导过程中并未涉及流体的运动状态，因而卡门动量积分关系式对层流和湍流的边界层运动都是适用的。

式（2.4-35）给出了 τ_w、δ、u_x 三者之间的关系。但一个方程不能求解三个未知数，方程不封闭，必须再补充两个方程。通常是选用边界层内速度分布及切应力与边界层厚度的关系式及切应力与边界层厚度的关系式。具体推导在此不再展开。

5　相似理论与量纲分析

应用数学的方法可以得出理想和黏性流体运动的微分方程式，对于一些可以用理想流体处理的工程问题，及对于象圆管、平行平板间等有限的几种层流流动，可以得到精确求解的方法。对于大多数黏性流体流动的工程问题，由于流体流动内在结构的复杂性，难以用微分方程加以描述，或者即使能够建立微分方程式，却由于起始条件和边界条件不能用数学方法给定，目前还无法求得其精确解。

流体力学的很多基本规律都是在某些简化的前提下得到的，求解具体问题时，只能根据具体情况，做出一些假设和推断，保留主要项，忽略次要项，求得问题的近似解。但是，这些前提假设，理论分析的运动情况以及由此求得的解是否合理，只能依靠试验验证。最初，人们用直接试验的方法，解决了一些用数学方法难以解决的流动问题，但结果只能应用于与试验条件相同的流动现象，有很大的局限性。另外，随着工业的发展，对于一些大型、特殊条件下的材料成形，不少设计方法都依据于试验来获得，但又无法用实物直接进行试验研究。

所有这些，都需要采用以相似原理为基础的模型试验方法，在流动相似原理的基础上，按照一定的原则把实物原形缩小或放大，选取合适的流动介质，制成模型试验装置。根据试验测定的参数整理试验数据，得出模型中流体的流动规律，然后依据相似原理，将这些结果推广到与试验模型相似

的各种实际成形中去。为此必须解决

1）如何根据实物确定试验模型？怎样选择流动介质？

2）试验过程中需要测量哪些物理量，怎样整理试验数据才能正确反映实际工程的情况，才能得到流动的规律性？

3）怎样将模型试验的观察、测量和整理结果还原成实物，并进行推广？

上述各项需要在不同工作条件的模型和实物之间进行相互换算，这些换算比例关系的基础，就是流体的力学相似。

5.1　流体力学的相似条件

只有在几何相似的基础上，才能实现两个流动现象的力学相似。两个流动现象的力学相似必须满足下面的三个条件。

(1) 几何相似

模型液流和实物液流有相似的边界条件，一切对应的特征尺寸为同一比值，且对应的角相同。图2.4-7给出了两个液流图形，实物液流用脚标 t 表示，模型液流用脚标 m 表示，若两个液流几何相似，对应的几何长度满足

$$\frac{l_t}{l_m}=\frac{D_t}{D_m}=C_l \qquad (2.4\text{-}36)$$

式中 C_l 是相似比例常数，只有满足上述条件，液流才能几何相似。两个液流几何相似时，对应的面积和体积满足 $C_A=C_l^2$ 和 $C_V=C_l^3$。

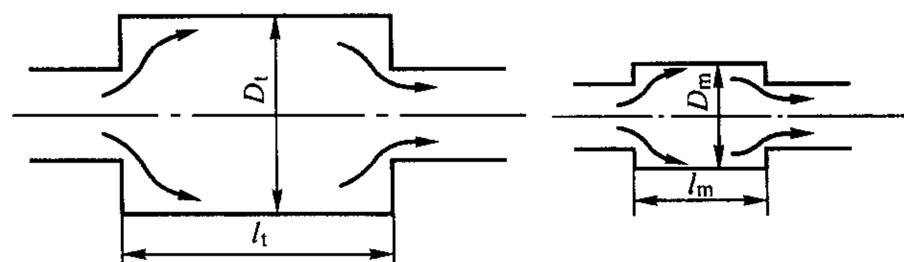


图2.4-7　液流几何相似示意图

(2) 运动相似

满足几何相似的两个液流中，若在对应瞬态，所有对应点上的速度方向一致，大小成同一比例，则两个液流运动相似。要满足上述条件，两个液流中各对应点流动状态参数变化的时间间隔应具有同一比例常数，即

$$\frac{t_{t1}}{t_{m1}}=\frac{t_{t2}}{t_{m2}}=\frac{t_{t3}}{t_{m3}}=C_t \qquad (2.4\text{-}37)$$

运动相似液流中对应点处的速度和加速度满足

$$C_v=\frac{v_{t1}}{v_{m1}}=\frac{v_{t2}}{v_{m2}}=\frac{v_t}{v_m}=\frac{l_t/t_t}{l_m/t_m}=\frac{C_l}{C_t} \qquad (2.4\text{-}38)$$

$$C_a=\frac{a_{t1}}{a_{m1}}=\frac{a_{t2}}{a_{m2}}=\frac{a_t}{a_m}=\frac{C_l}{C_t^2} \qquad (2.4\text{-}39)$$

显然，两个运动相似的液流中，同一瞬时对应的流线应当相似，对恒定流动，对应流体质点的运动轨迹几何相似，且通过对应迹线的时间成同一比例。

(3) 动力相似

运动相似的液流中，在对应的瞬时，对应点上受相同性质力的作用，力的方向相同，且各对应的同名力大小成同一比例，则两个液流动力相似。若一个液流中某流体质点上只受重力和离心力作用，另一液流对应上也必须只受重力和离心力作用，且对应流体质点上所受的重力之比等于离心力之比，因此，动力相似有

$$C_f=\frac{F_{t1}}{F_{m2}}=\frac{F_{t2}}{F_{m2}}=\frac{F_t}{F_m} \qquad (2.4\text{-}40)$$

又由牛顿定律有

$$C_f=\frac{F_t}{F_m}=\frac{M_t a_t}{M_m a_m}=\frac{\rho_t V_t a_t}{\rho_m V_m a_m}=C_\rho C_l^2 C_v^2 \qquad (2.4\text{-}41)$$

即 $\dfrac{C_f}{C_\rho C_l^2 C_v^2}=1$，于是有

$$\frac{F_t}{\rho_t l_t^2 v_t^2}=\frac{F_m}{\rho_m l_m^2 v_m^2}=Ne \qquad (2.4\text{-}42)$$

无因次数 Ne 称为牛顿数，是判别液流是否动力相似的无因次数或牛顿判据，牛顿数相等是两个流体动力学相似的必要条件。

由上述相似条件得到，几何相似是运动相似的必要条件，运动相似是动力相似的必要条件。

5.2　黏性流体流动的力学相似准数

任何相似的现象都是属于同一种类的现象，在两个力学相似流动的对应点处，流体运动的所有物理量遵循同一微分方程式，各物理量间的比值相互约束。应该从黏性流体运动的微分方程出发，寻求流体流动的一般相似判据。

设有实物和模型两个力学相似的不可压缩流体流动，对于实物流动，通过给定点的任意流体质点有如下运动方程

$$X_t-\frac{1}{\rho_t}\frac{\partial p_t}{\partial x_t}+v_t\nabla^2 v_{xt}=\frac{\partial v_{xt}}{\partial t_t}+v_{xt}\frac{\partial v_{xt}}{\partial x_t}+v_{yt}\frac{\partial v_{xt}}{\partial y_t}+v_{zt}\frac{\partial v_{xt}}{\partial z_t} \qquad (2.4\text{-}43a)$$

$$Y_t-\frac{1}{\rho_t}\frac{\partial p_t}{\partial y_t}+v_t\nabla^2 v_{yt}=\frac{\partial v_{yt}}{\partial t_t}+v_{yt}\frac{\partial v_{yt}}{\partial y_t}+v_{yt}\frac{\partial v_{yt}}{\partial z_t}+v_{zt}\frac{\partial v_{yt}}{\partial z_t} \qquad (2.4\text{-}43b)$$

$$Z_t-\frac{1}{\rho_t}\frac{\partial p_t}{\partial z_t}+v_t\nabla^2 v_{zt}=\frac{\partial v_{zt}}{\partial t_t}+v_{v_{zt}}\frac{\partial v_{zt}}{\partial y_t}+v_{yt}\frac{\partial v_{zt}}{\partial z_{yt}}+v_{zt}\frac{\partial v_{zt}}{\partial z_t} \qquad (2.4\text{-}43c)$$

对于模型流体，通过对应点的任意流体质点有运动方程

$$X_m-\frac{1}{\rho_m}\frac{\partial p_m}{\partial x_m}+v_m\nabla^2 v_{xm}=\frac{\partial v_{xm}}{\partial t_m}+v_{xm}\frac{\partial v_{xm}}{\partial x_m}+v_{ym}\frac{\partial v_{xm}}{\partial y_m}+v_m\frac{\partial v_{xm}}{\partial z_m} \qquad (2.4\text{-}44a)$$

$$Y_m-\frac{1}{\rho_m}\frac{\partial p_m}{\partial y_m}+v_m\nabla^2 v_{ym}=\frac{\partial v_{ym}}{\partial t_m}+v_{xm}\frac{\partial v_{ym}}{\partial y_m}+v_{ym}\frac{\partial v_{ym}}{\partial y_m}+v_m\frac{\partial v_{ym}}{\partial z_m} \qquad (2.4\text{-}44b)$$

$$Z_m-\frac{1}{\rho_m}\frac{\partial p_m}{\partial z_m}+v_m\nabla^2 v_{zm}=\frac{\partial v_{zm}}{\partial t_m}+v_{xm}\frac{\partial v_{zm}}{\partial x_m}+v_{ym}\frac{\partial v_{zm}}{\partial y_m}+v_m\frac{\partial v_{zm}}{\partial z_m} \qquad (2.4\text{-}44c)$$

由于两个流动几何相似和运动相似，根据相似条件有

$$x_t=C_l x_m;\ y_t=C_l y_m;\ z_t=C_l z_m;\ t_t=C_t t_m$$

$$v_{xt}=\frac{C_l}{C_t}v_{xm}=C_v v_{xm};\ v_{yt}=\frac{C_l}{C_t}v_{ym}=C_v v_{ym};$$

$$v_{zt}=\frac{C_l}{C_t}v_{zm}=C_v v_{zm} \qquad (2.4\text{-}45)$$

即有 $v_t=C_v v_m$。

对于力学相似的两个液流，对应点处的同名力有

$$p_t=C_p p_m$$

$$X_t=C_f X_m;\ Y_t=C_f Y_m;\ Z_t=C_f Z_m \qquad (2.4\text{-}46)$$

同时，在所讨论的力学相似的流动中，流体的密度和黏度系数在对应点处也满足下述关系

$$\rho_t=C_\rho\rho_m;\ \eta_t=C_\eta\eta_m \qquad (2.4\text{-}47)$$

在流体的物理量变化过程中，在对应瞬时的对应点上应满足式（2.4-45）~式（2.4-47）。上述分析中我们引入了两个流动各物理量的相似常数 C_l，C_v，C_t，C_p，C_ρ，C_f，C_η 等，这些常数与所选取的坐标和时间相关，取决于所研究流动的起始条件、边界条件及这些流动中流体的力学性质。

显然，由于两个相似流动的各物理量必须满足同一运动微分方程，因此各对应量的比值互相制约，相似倍数不能任意选取。

将各物理量的相似常数代入实物流动的运动方程中，得到

$$C_f X_m-\frac{C_p}{C_\rho C_l}\frac{1}{\rho_m}\frac{\partial p_m}{\partial x_m}+\frac{C_\eta C_v}{C_l^2}\eta_m\nabla^2 v_{xm}$$

$$= \frac{C_v}{C_t}\frac{\partial v_{xm}}{\partial t_m} + \frac{C_v^2}{C_l}\left(v_{xm}\frac{\partial v_{xm}}{\partial x_m} + v_{ym}\frac{\partial v_{xm}}{\partial y_m} + v_{zm}\frac{\partial v_{xm}}{\partial z_m}\right) \tag{2.4-48}$$

这样处理后，方程的参数变成了模型流动中的参数，与模型流动微分方程的各项参数只是系数不同。由流体动力相似得到：两个力学相似流动中，对应点上各同名力的比值相等，于是有

$$C_f = \frac{C_p}{C_\rho C_l} = \frac{C_\eta C_v}{C_l^2} = \frac{C_v}{C_t} = \frac{C_v^2}{C_l} \tag{2.4-49}$$

若将上式各项除以 C_v^2/C_l，有

$$\frac{C_f C_l}{C_v^2} = \frac{C_p}{C_\rho C_v^2} = \frac{C_\eta}{C_l C_v} = \frac{C_l}{C_t C_v} = 1 \tag{2.4-50}$$

由此可求得黏性不可压缩流体流动的相似准数，即

1) $\frac{C_v C_t}{C_l} = 1$

得

$$\frac{v_t t_t}{l_t} = \frac{v_m t_m}{l_m} = \frac{vt}{l} = St \tag{2.4-51}$$

式中 St（不变数）为斯特劳哈（Strouhal）数。若两液流相应的斯特劳哈数相等，则液流对非恒定流动惯性力而言是力学相似的。对非恒定的周期运动，反映其周期性相似。对恒定流动，则不必考虑 St 数。

2) $\frac{C_l C_v}{C_\eta} = 1$

得

$$\frac{v_t l_t}{\eta_t} = \frac{v_m l_m}{\eta_m} = \frac{vl}{\eta} = Re \tag{2.4-52}$$

式中 Re（不变数）为雷诺（Reynolds）数。若两液流对应点上的雷诺数相等，则液流对粘性力而言是力学相似的。因此，很容易理解不论何种流体，在几何相似的大小不同的圆管中流动，它们的临界雷诺数都相等，其原因是对黏滞力而言处于流动力学相似状态。

3) $\frac{C_p}{C_\rho C_v^2} = 1$

得

$$\frac{p_t}{\rho_t v_t^2} = \frac{p_m}{\rho_m v_m^2} = \frac{p}{\rho v^2} = Eu \tag{2.4-53}$$

式中 Eu（不变数）为欧拉（Euler）数。因为流体中的压强或压差往往取决于其他参数，而并不是流体固有的物理性质，所以欧拉数不是独立的相似准数，而可以是其他相似准数的函数，即它不是相似条件而是相似结果。

4) $\frac{C_v^2}{C_f C_l} = 1$

得

$$\frac{v_t^2}{F_t l_t} = \frac{v_m^2}{F_m l_m} = \frac{v^2}{Fl} = Fr \tag{2.4-54}$$

式中 Fr（不变数）为佛劳德（Froude）数。若流体所受的质量力只有重力，$F_t = F_m = g$，则 $Fr = \frac{v^2}{gl}$。Fr 数相等，表示了流动现象的重力作用相似，它反映了重力对流体的作用。

另外，对可压缩流体的流动，声速为

$$a^2 = \frac{\partial p}{\partial \rho} \tag{2.4-55}$$

对模型和实物流动中，流体质点上的相似比例常数为：

$$C_a^2 = \frac{a_t^2}{a_m^2} = \frac{\partial p_t/\partial \rho_t}{\partial p_m/\partial \rho_m} = \frac{C_p}{C_\rho} \tag{2.4-56}$$

又因为 $\frac{C_p}{C_\rho C_v^2} = \frac{C_a^2}{C_v^a} = 1$

得

$$\frac{v_t}{a_t} = \frac{v_m}{a_m} = \frac{v}{a} = Ma \tag{2.4-57}$$

式中 Ma（不变数）为马赫（Mach）数。因为 $a^2 = \frac{\partial p}{\partial \rho}$，所以声速 a 实质上为流体单位密度变化所需要的压强变化，即反映了流体的压缩性，因此 Ma 数相等就是压缩性相似。

如果黏性不可压缩流体的两个流动力学相似，那么，满足边界条件和起始条件相似的 St、Eu、Re 和 Fr 数应当相等，这就是相似准则，称 St、Eu、Re 和 Fr 数为相似准数。两个流动的相似准数相同，表示了对应点处单位质量流体上作用的力多边形几何相似。

5.3　决定性相似准数

上述得到的相似准数，是任意两个力学相似的流动应当满足的条件。在实际的液流中，流体上总有重力、黏滞力、压力和惯性力等同时作用。但是，由试验和分析知道，在流体流动的力学现象中，通常只有一到二种力起着主要作用，决定力学现象的本质，另一些力则属于次要地位，因而在任何流动现象中，都存在着决定现象本质的主要作用力。

由于实际流动的复杂性，要同时满足四个相似准数十分困难，且有些相似准数要同时满足也是不可能的，因此，在流体力学的相似理论中，一般采用近似的力学相似，即只需求得主要作用力的力学相似，其他对流动现象不起主要作用的力则忽略不计。

两个力学相似的流动，必须考虑的最少数目的相似准则数，称为决定性相似准数。两个流动力学相似，只要决定性相似准则数相同，次要的相似准则数可不必考虑，在工程实际中，决定性相似准则数通常不超过两个。

实际上，两个流动要同时满足 Re 和 Fr 准则数相等一般都是很难实现的。若两种不同流体的流动力学相似，满足 Re 数相等，$\frac{v_t l_t}{\eta_t} = \frac{v_m l_m}{\eta_m}$，有

$$\frac{v_t}{v_m} = \frac{1}{C_l}\frac{\eta_t}{\eta_m} \tag{2.4-58}$$

若同时满足 Fr 数相等，并令 $g_t = g_m$，则

$$\frac{l_t}{v_t^2} = \frac{l_m}{v_m^2} \tag{2.4-59}$$

有

$$\frac{v_t}{v_m} = \sqrt{C_l} \tag{2.4-60}$$

于是

$$C_l^{3/2} = \frac{\eta_t}{\eta_m} \tag{2.4-61}$$

若实物尺寸为 1 m，现以 $l = 0.1$ m 的模型进行试验，$C_l = 10$，则

$$\eta_m = \frac{\eta_t}{31.6} \tag{2.4-62}$$

如果实物试验的介质为水，必须选用黏性比水小 31.6 倍的流体做模型试验。在工程实际中，很难找到粘度相差这么大的两种流体，为此，必须增大模型尺寸，但这又失去了模型试验的优点。

显然，决定性相似准则数的判断选择十分重要。选择不当，设计的模型试验甚至会得到与实际完全不同的结果，所以必须在对实际流动充分观察和分析的基础上，选取决定性相似准则数。所选取的相似准则数越少，实施中越容易满足。

从上述分析可见，当同时满足两个相似准则设计试验时，模型中流体介质的选择要受模型尺寸选择的限制。当选用更多的相似准则时，将因为流体的物理量间相互制约而难以进行模型设计。因此，通常都选用决定性相似准则进行近似的模型设计和试验。

5.4　量纲分析方法

流动的各种物理现象常受到各种因素的影响，对于简单的现象可以通过简化，建立运动微分方程，求得精确解。但对大量复杂的流动现象，理论分析本身就比较困难，必须结合试验，才能使理论分析深入进行。但是，如果没有正确的

理论指导，不知需要测定哪些物理量和试验数据应该如何整理，那么，虽然能获取大量数据，却也无法找出影响现象本质的因素，使试验带有盲目性。通过量纲分析，能将影响物理现象的各种变量合理组合，将有量纲的量变成无量纲积，由于无量纲积的数量少于原来有量纲变量的数量，可使问题简化。

自然界的物理力学现象所建立的方程，不管采用哪一种单位制，都必须满足方程两边量纲的齐次性。通过量纲分析，能将无量纲积进行组合，求得一个含有特定系数的通式，待定系数的函数关系指示了试验的方向，从复杂的试验数据分析得出普遍规律，通过试验来确定待定参数。

任何物理量都是有单位的，但是量纲与单位不同。例如长度的单位可以是米、厘米或毫米，但它的量纲都是长度 $[L]$；时间的单位可以是分或秒，但它的量纲都是时间 $[T]$；速度的单位可以是米/分，或厘米/秒，但它的量纲都是 $[LT^{-1}]$，即所有这些不同的单位，都可以由某些不能用其他量纲导出的基本量纲组成。在不同的单位制中，有不同的基本量纲。国际单位制中的基本量纲为：长度 $[L]$，时间 $[T]$，质量 $[M]$，温度 $[Q]$ 和热量 $[H]$。

利用量纲分析确定力学过程的函数关系时，可按下面的步骤进行：①列出所有与该物理现象有关的变量，对现象有重要影响的变量不可丢掉，但可略去一些次要的变量，关键在于对现象物理本质的了解程度；②将这些变量的量纲用基本量纲 L、T、M 表示出来；③将变量组成某种由基本量纲表示的量纲一致的函数关系（通常为各变量指数乘积关系）；④将各量的量纲代入上面的指数乘积关系；⑤利用关系式量纲的齐次性，对各基本量纲的指数列出代数方程，联立求解方程，将解得的指数代入函数中，得到函数的具体形式；⑥试验确定所引入的无量纲常数。

下面介绍量纲分析方法中广泛应用的 π 定理（E·Buckingham 定理）。设有一个未知函数关系

$$N = f(n_1, n_2, n_3, \cdots\cdots, n_k) \tag{2.4-63}$$

在 N 和（n_1，n_2，n_3，……，n_k）共 $k+1$ 个物理量中，可能有常数，有变数，有些量可以有量纲，另一些可能无量纲。但是，这样的函数关系式与所选用的单位制无关，不同单位制只是数值不同，函数关系式不变。

为此，在所有这些变量中，首先确定出三个物理量 n_1、n_2、n_3 的量纲为基本量纲，这三个基本量纲彼此独立，其余（$k+1-3$）个物理量的量纲都可以表示为这三个基本物理量量纲的指数形式

$$[N] = [n_1]^x[n_2]^y[n_3]^z \tag{2.4-64a}$$

$$[n_i] = [n_1]^{x_i}[n_2]^{y_i}[n_3]^{z_i} \tag{2.4-64b}$$

根据物理方程两边量纲的齐次性，可以确定指数 x、y、z 和 x_i、y_i、z_i。这样，就可以将原来函数关系式中 $(k+1-3)$ 个物理量变为量纲为 1 的量，这些量纲为 1 的量与原来的变量之间相差量纲为 1 的数 π 和 π_i，即

$$N = \pi n_1^x n_2^y n_3^z \tag{2.4-65a}$$

$$n_i = \pi_i n_1^{x_i} n_2^{y_i} n_3^{z_i} \tag{2.4-65b}$$

或

$$\pi = \frac{N}{n_1^x n_2^y n_3^z} \tag{2.4-66a}$$

$$\pi_i = \frac{n_i}{n_1^{x_i} n_2^{y_i} n_3^{z_i}} \tag{2.4-66b}$$

π 及 π_i 是 N 和 n_i 以 n_1、n_2、n_3 为基本量纲下的数值，于是，在新的基本量纲下，原来的函数关系式变为

$$\frac{N}{n_1^x n_2^y n_3^z} = f\left(1,1,1,\frac{n_4}{n_1^{x_4} n_2^{y_4} n_3^{z_4}},\cdots\cdots\frac{n_k}{n_1^{x_k} n_2^{y_k} n_3^{z_k}}\right) \tag{2.4-67}$$

即：

$$\pi = f(1,1,1\pi_4\cdots\cdots\pi_k) \tag{2.4-68}$$

π 为 π_i 的无量纲函数，因此式（2.4-66）右边也必定为无量纲数，故式中右边的指数可以由分子与分母的量纲相等原则确定。

这就是 π 定理。经过上述变换，将原来 $k+1$ 个带有量纲的物理变量 N 和 n_i 间的函数关系式，变成（$k+1-3$）个无量纲数 π 和 π_i 之间的关系式。

利用 π 定理，能够指导人们从大量的试验数据中分析得出普遍规律，指出试验的正确方向，从而有目的地指导试验工作的进行。当然，一定要在大量试验的基础上，了解影响某物理量的主次因素，否则将不可能利用量纲分析得出正确的结果。

对于一些较复杂的流动现象，当只靠数学分析无法求解，或者暂时还难以找到确切的方程加以描述时，π 定理提供了利用量纲分析结合试验研究的解决方法，π 定理和前面得到的相似准则数是模型试验研究的理论基础。

现举例说明 π 定理的应用。有一直径为 d 的圆球、在黏性系数为 μ，密度为 ρ 的液体中以等速度 v 下降，试求圆球受到的阻力 F_f。

显然阻力 F_f 与 d、μ、ρ、v 等因素间的函数关系可以写为：

$$F_f = f(v,\rho,\mu,d) \tag{2.4-69}$$

首先应将上面的函数关系式写成无量纲关系式。式中独立变量有 v、ρ、μ 和 d 共四个，即 $k=4$。选用 $[M、L、T]$ 为基本量纲。所有各个变量的量纲都可以用基本量纲加以表示：$F_f\left[\frac{ML}{T^2}\right]$，$v\left[\frac{L}{T}\right]$，$d\ [L]$，$\rho\left[\frac{M}{L^3}\right]$，$\mu\left[\frac{M}{LT}\right]$。并且选择 ρ，v，d 为三个基本变量，则其余 $(k+1-3)=2$ 个变量 F_f 和 μ 的量纲可以表示为三个基本物理量量纲的指数形式，即

$$F_f = \pi\rho^x v^y d^z \tag{2.4-70}$$

$$\mu = \pi_1 \rho^{x_1} v^{y_1} d^{z_1} \tag{2.4-71}$$

由此可以列出两个无量纲数 π 和 π_1，把各个量的量纲代入，得无量纲数 π 为：

$$\pi = \frac{F_f}{\rho^x v^y d^z} = \left[\frac{ML}{T^2}\right]\left[\frac{M}{L^3}\right]^{-x}\left[\frac{L}{T}\right]^{-y}[L]^{-z} = [M^0L^0T^0] \tag{2.4-72}$$

由等式两边量纲相等的原则，可列出指数方程为

$$\begin{cases} M: -x+1=0 \\ T: y-2=0 \\ L: 3x-y-z+1=0 \end{cases} \tag{2.4-73}$$

解上面代数方程组得：$x=1$，$y=2$，$z=2$。于是有

$$\pi = \frac{F_f}{\rho v^2 d^2} \tag{2.4-74}$$

同样，对于 π_1，有

$$\pi_1 = \frac{\mu}{\rho^{x_1} v^{y_1} d^{z_1}} = \left[\frac{M}{LT}\right]\left[\frac{M}{L^3}\right]^{-x_1}\left[\frac{L}{T}\right]^{-y_1}[L]^{-z_1} = [M^0L^0T^0] \tag{2.4-75}$$

可列出指数方程为

$$\begin{cases} M: -x_1+1=0 \\ T: y_1-1=0 \\ L: 3x_1-y_1-z_1=0 \end{cases} \tag{2.4-76}$$

解上面代数方程组得：$x_1=1$，$y_1=1$，$z_1=1$。所以

$$\pi_1 = \frac{\mu}{\rho v d} = \frac{v}{vd} = \frac{1}{Re} \tag{2.4-77}$$

于是，原来的函数关系式变为

$$\frac{F_f}{\rho v^2 d^2} = f(1,1,1,\frac{1}{Re}) = f\left(\frac{1}{Re}\right) \qquad (2.4\text{-}78)$$

这样，就能够按上面的关系式进行试验。显然，经过应用量纲分析的 π 定理得到的关系式，简化并指明了试验方向。本来为了确定所要求的阻力 F_f 与四个变量 ρ，v，μ，d 的关系，必须分别在四个变量中逐次将三个量保持不变，改变另一个量进行试验测定，如先将 ρ，μ，d 设为一定值，变化不同的 v 进行一组试验测定，然后将 ρ，v，μ 设为一定值，变化不同的 d 进行试验测定，这样，若对每个变量给出10个值进行试验，总共就需要作 10^4 次的试验测定，工作量十分浩大。但是，当采用量纲分析的 π 定理后，原来五个量纲之间的关系就缩为上面的无量纲关系式，试验中只要求改变10个 Re 值测定对应的 $\frac{F_f}{\rho v^2 d^2}$ 值就可以了。实际上要使 Re 变化，只需要 ρ，μ，d 保持不变的情况下，改变圆球下降的速度 v 就可以完成试验测定，从而大大简化了试验方法，缩短了试验时间，量纲分析的方法显示了明显的优越性。

编写：周华民（华中科技大学）

第5章　固体材料的本构关系

1　概述

1.1　固体材料本构关系的一般原理

塑性成形中涉及的材料一般为固体，可以作为连续介质看待，它们遵从连续介质力学的普遍规律，其中包括连续性方程、运动方程、能量方程等。由这些方程及相应的边界条件就构成了描述塑性加工过程的偏微分方程的边值问题或边值-初值问题。但是，这些方程的数目少于未知物理量的数目。因此，为了能求得材料成形过程中涉及的连续介质运动问题的解，还必须在各物理量之间建立一些描述材料本身所固有的物理特性的补充方程，以便得到封闭的方程组。这些补充方程可以统称为本构方程。其中确定应力 $\boldsymbol{\sigma}$ 的方程称为应力应变关系；确定热流密度向量 $\boldsymbol{q}$ 的方程称为热传导规律；而热力学函数（密度 ρ、压力 p、绝对温度 T、内能 e、熵 s）之间满足的方程都称状态方程。

通常所说的本构方程是指狭义的本构方程，即应力应变关系。它是表达连续介质的应力或应力变化率与应变或应变速率之间关系的物理方程。它与材料性质、结构和变形条件有关。

不同的连续介质，或同一介质在不同的运动条件下将有不同的本构方程。每种介质在特定条件下的本构方程应由理论和实验共同建立。建立本构方程时应遵从如下的一般原理。

1）决定性原理　物体中某点的应力恒可由物体内各点的以往运动史惟一地确定。

2）局部作用原理　物体中某点处的应力完全由该质点附近任意小邻域的变形历史和状态所决定。

3）坐标系无关性　本构方程应该用张量方程表述。

4）客观性原理　本构方程是由物体性质决定的，它不应随观察者的不同而改变。因此在作相对运动的时空参考系中，本构方程具有相同的函数形式。

1.2　塑性成形分析中材料的变形和本构关系的选取

塑性成形一般涉及两类材料，即工具和工件。工具只允许发生弹性变形，如果工具出现塑性变形，则工件的尺寸将随之变化，无法保证精度。在塑性成形问题的分析中，一般可以将工具视为不变形的刚体，仅分析工件的塑性变形。如果要校核工具的强度或刚度，考察工具的变形对工件成形质量的影响，就必须将工具作为弹性体进行分析。由于这时弹性变形很小，可以用线弹性本构方程。另外，在塑性成形中有时采用橡胶、聚氨酯等高分子材料作为传力介质，这一类材料经历极大的变形以后仍然保持弹性，在变形过程中应力-应变关系一般是非线性的。对于这一类材料要采用超弹性本构方程。

在成形过程中，工件必然要发生塑性变形。通常可以采用宏观塑性力学本构方程描述工件的塑性变形。塑性变形的机理与材料本身的微观组织密切相关，同时塑性变形过程伴随着材料微观组织的演化，其规律相当复杂。因此，对于一些特定的塑性变形问题，或者为了研究塑性变形过程某些特定的方面，常常需要应用相应的更合适的本构方程。工程中使用的金属材料一般为多晶体材料，在常温下其塑性变形的主要机理是位错滑移。对于低层错能材料，孪晶也是塑性变形的重要方式。晶体塑性理论考虑晶体材料的塑性变形的几何特点，很适用于研究晶体材料的变形特点。在超塑性变形条件下，晶界的滑动成为塑性变形的主要机理。当材料的某个尺寸减小到一定程度以后，会出现尺度效应。这时材料的屈服应力会出现与该尺寸相关的变化。应变梯度塑性理论考虑了材料内秉尺寸对屈服条件的影响。塑性成形过程中材料微观组织的演化规律相当复杂。由于塑性变形，多晶体材料的晶粒会呈现与具体晶格类型和加载方式相关的择优取向，称为织构。织构将导致多晶体材料的各种物理和力学性能呈现各向异性。需要考虑各向异性对塑性成形过程的影响时可采用各向异性本构方程。塑性成形过程中由于位错之间的相互作用、各相材料变形的不均匀等原因，在材料中会形成微裂纹和微空洞，这些微缺陷会随着塑性变形而发展，直到产生宏观裂纹，导致材料的破坏。损伤力学本构方程可用来研究材料的损伤演化过程。

为了描述千差万别的工程材料在各种加工条件下的力学行为，已经提出的本构关系有着不同的表现形式。在研究实际工程问题时，应根据材料本身的特点和研究计算的目的选取适当的本构关系。在选定本构关系以后，还需要查阅有关资料或采用试验方法测定本构方程中的材料常数，以保证所采用的本构方程能准确地反映所关心的材料的力学特性。

由于材料种类以及成形条件的多样性，也由于研究手段的不断改进，材料本构关系一直是一个相当活跃的研究领域。

2　弹性本构方程

2.1　线弹性本构方程

固体材料发生弹性变形时，其应力仅取决于当前的应变状态，是应变的单值函数。当载荷卸除后，弹性变形可以恢复，即弹性变形是可逆的。当应变很小时，一般材料仅发生弹性变形。弹性应力应变关系一般可视为线性的。

2.1.1　各向同性弹性

在温度不变的条件下固体材料发生小变形时，应力分量是小变形张量分量的线性函数，即有

$$\sigma_{ij} = C^{e}_{ijkl}\varepsilon^{e}_{kl} \tag{2.5-1}$$

式中，C^{e}_{ijkl} 为弹性张量，上标“e”表示弹性。这种材料模型称为线性弹性体。上式的增量形式为

$$\mathrm{d}\sigma_{ij} = C^{e}_{ijkl}\mathrm{d}\varepsilon^{e}_{kl} \tag{2.5-2}$$

工程中所采用的金属材料一般可当作各向同性的线性弹性体，这种材料模型称为理想弹性体。这时弹性张量为四阶各向同性张量，其中仅包含两个独立参数，并可表示为

$$C^{e}_{ijkl} = 2G\left(\delta_{ik}\delta_{jl} + \frac{\nu}{1-2\nu}\delta_{ij}\delta_{kl}\right) \tag{2.5-3}$$

式中，G 为切变模量，ν 为泊松比。G 与弹性模量 E 的关系为

$$G = \frac{E}{2(1+\nu)}$$

在有限变形情况下，为了满足客观性要求，采用具有客观性的柯西应力的久曼速率张量和应变速率张量，将弹性本构方程写为：

$$\overset{\circ}{\sigma}_{ij} = C^{e}_{ijkl}d^{e}_{kl} \tag{2.5-4}$$

为了便于有限元列式，通常把应力和应变由二阶张量改写成向量，例如

$$\boldsymbol{\sigma} = [\sigma_{11}\ \sigma_{22}\ \sigma_{33}\ \sigma_{12}\ \sigma_{23}\ \sigma_{31}]^{\mathrm{T}},\ \boldsymbol{\varepsilon} = [\varepsilon_{11}\ \varepsilon_{22}\ \varepsilon_{33}\ 2\varepsilon_{12}\ 2\varepsilon_{33}\ 2\varepsilon_{31}]^{\mathrm{T}}$$

而把四阶张量 C^e 改写成矩阵

$$C^e = \frac{2G}{1-2\nu}\begin{bmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-2\nu)/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & (1-2\nu)/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-2\nu)/2 \end{bmatrix} \tag{2.5-5}$$

于是，弹性应力应变关系（即广义虎克定律）可以写成矩阵形式

$$\boldsymbol{\sigma} = \boldsymbol{C}^e\boldsymbol{\varepsilon}$$

对于矩阵运算，我们不写点积符号。在本章中，$\boldsymbol{\sigma}$ 既用于表示应力张量，也用于表示其向量写法。对于 $\boldsymbol{\varepsilon}$ 和 $\boldsymbol{C}^e$ 也有类似的约定。我们将在上下文中具体指明其含义。

弹性变形可分解为体积变化 $\varepsilon_m^e = (\varepsilon_{11}^e + \varepsilon_{22}^e + \varepsilon_{33}^e)/3$ 和形状变化 $\varepsilon_{ij}^{e'} = \varepsilon_{ij}^e - \varepsilon_m^e\delta_{ij}$，其中

$$\varepsilon_{ij}^{e'} = \frac{1}{2G}\sigma_{ij}',\ \varepsilon_m^e = \frac{1-2\nu}{E}\sigma_m$$

2.1.2 正交各向异性弹性

某些材料，例如单晶体，沿不同方向具有不同的弹性性能。其中较常见且简单的情况是正交各向异性，即它具有相互正交的主轴。考虑其中沿各主轴弹性性能相同的情况，如立方晶系的单晶体，弹性矩阵为

$$C^e = \begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix} \tag{2.5-6}$$

2.2 超弹性本构方程

橡胶等高分子材料具有高度非线性的应力-应变关系，在极大的应变（如100%或更高）下仍保持为弹性，而且在拉伸、压缩、剪切等不同的加载条件下，其应力-应变关系曲线可呈现不同的变化趋势。这类材料通常采用超弹性（hyper-elasticity）本构方程描述。

超弹性体的定义如下。如果单位质量的物体具有应变能函数 w，w 为应变张量的解析函数，且其增量等于应力所作的功，则该物体称为超弹性体。即超弹性体满足

$$w = \frac{1}{\rho_0}S_{ij}\dot{E}_{ij} \tag{2.5-7}$$

式中，ρ_0 为初始构形中材料的密度；S_{ij} 为克希荷夫应力张量；$\dot{E}_{ij}$ 为格林应变速率张量。

另一方面，由于应变能函数为应变张量的解析函数，又有

$$\dot{w} = \frac{\partial w}{\partial E_{ij}}\dot{E}_{ij} \tag{2.5-8}$$

由式（2.5-7）和式（2.5-8）得

$$\left(\frac{1}{\rho_0}S_{ij} - \frac{\partial w}{\partial E_{ij}}\right)\dot{E}_{ij} = 0 \tag{2.5-9}$$

在 $\dot{E}_{ij}$ 的各分量是完全独立的情况下，由 $\dot{E}_{ij}$ 的任意性知

$$S_{ij} = \rho_0\frac{\partial w}{\partial E_{ij}} \tag{2.5-10}$$

式（2.5-10）即为超弹性体的本构方程。

如果在变形过程中物体的体积变化可以忽略不计，即物体为不可压缩材料时，E_{ii} 的各分量不是完全独立的，而是要满足如下体积不可压缩条件

$$\varepsilon_{ii} = \frac{\partial X_m}{\partial x_i}\frac{\partial X_n}{\partial x_i}\dot{E}_{mn} = 0 \tag{2.5-11}$$

比较式（2.5-9）和式（2.5-11）知，此时有

$$\frac{1}{\rho_0}S_{ij} - \frac{\partial w}{\partial E_{ij}} = \frac{\partial X_i}{\partial x_m}\frac{\partial X_j}{\partial x_m}$$

即

$$S_{ij} = \rho_0\frac{\partial w}{\partial E_{ij}} + h\frac{\partial X_m}{\partial x_i}\frac{\partial X_n}{\partial x_i} \tag{2.5-12}$$

式（2.5-12）即为不可压缩超弹性体的本构方程。其中 h 为静水压力，由边界条件确定。

对各向同性超弹性材料，应变能密度函数可表示为格林变形张量 C 的主不变量的函数

$$w = w(I_1, I_2, I_3)$$

其中

$$\left.\begin{aligned} I_1 &= C_{ii} = 2E_{ii} + 3 \\ I_2 &= (C_{ii}C_{jj} - C_{ij}C_{ji})/2 = 2E_{ii}E_{jj} + 4E_{ii} - 2E_{ij}E_{ij} + 3 \\ I_3 &= \det\boldsymbol{C} = \det\left(\frac{\partial x_m}{\partial X_i}\frac{\partial x_m}{\partial X_j}\right) \end{aligned}\right\} \tag{2.5-13}$$

于是

$$S_{ij} = \rho_0\left[\frac{\partial w}{\partial I_1}\frac{\partial I_1}{\partial E_{ij}} + \frac{\partial w}{\partial I_2}\frac{\partial I_2}{\partial E_{ij}} + \frac{\partial w}{\partial I_3}\frac{\partial I_3}{\partial E_{ij}}\right] \tag{2.5-14}$$

其中

$$\left.\begin{aligned} \frac{\partial I_1}{\partial E_{ij}} &= 2\delta_{ij} \\ \frac{\partial I_2}{\partial E_{ij}} &= 2(\delta_{ij}\delta_{rs} - \delta_{ir}\delta_{js})(2E_{rs} + \delta_{rs}) \\ \frac{\partial I_3}{\partial E_{ij}} &= 2\frac{\partial X_i}{\partial x_m}\frac{\partial X_j}{\partial x_m}I_3 \end{aligned}\right\} \tag{2.5-15}$$

对于体积不可压缩超弹性体，不需考虑 I_3（注意到此时 $|\boldsymbol{C}| = |\boldsymbol{F}^T\boldsymbol{F}| = |\boldsymbol{F}^T|\cdot|\boldsymbol{F}| = \left(\frac{dV}{dV_0}\right)^2 = 1$），于是

$$S_{ij} = \rho_0\left[\frac{\partial w}{\partial I_1}\frac{\partial I_1}{\partial E_{ij}} + \frac{\partial w}{\partial I_2}\frac{\partial I_2}{\partial E_{ij}}\right] + h\frac{\partial X_i}{\partial x_m}\frac{\partial X_j}{\partial x_m} \tag{2.5-16}$$

对应变能密度函数给出不同定义，就定义不同的超弹性材料模型。下面是几种常用的超弹性本构模型。

1）Neo Hooke 材料模型

$$w = C_1(I_1 - 3) \tag{2.5-17}$$

2）Moony 材料模型

$$W = C_1(I_1 - 3) + C_2(I_2 - 3) \tag{2.5-18}$$

式中，C_1、C_2 为材料常数。

以上两种材料的应变能密度函数都是利用格林变形张量的不变量来定义的。由式（2.2-16）知 $C_{ii} = \lambda_i^2$（λ_i 为主伸长率），所以这两种材料的应变能密度函数也都是 λ_i 的偶次幂的函数，都属于 Rivilin 材料的范畴。Rivilin 材料的应变能密度函数还可以取为 I_1，I_2 的更高次数的多项式，随着次数（项数）的增加，本构方程与实际材料吻合得更好，但待定的材料参数也越多，计算越复杂。同时也可以看到，以上两种材料的应变能密度函数中都不包含 I_3，故均为不可压缩超弹性材料。

3）Ogden 材料模型　这种材料模型中考虑了体积的变化，其应变能密度函数为

$$w = \sum_{n=1}^{N}\frac{\mu_n}{\alpha_n}\left[J^{\frac{-\alpha_n}{3}}(\lambda_1^{\alpha_n} + \lambda_2^{\alpha_n} + \lambda_3^{\alpha_n}) - 3\right] + 4.5K(J^{\frac{1}{3}} - 1)^2 \tag{2.5-19}$$

其中，μ_n 和 α_n 为材料常数；K 为体积模量；J 为体积率且定义为

$$J = \lambda_1\lambda_2\lambda_3$$

采用式（2.5-19）进行分析时体积变化的量级应为

0.01。对于体积变化很大的情况，可用广义的可压缩 Ogden 列式

$$w = \sum_{n=1}^{N}\frac{\mu_n}{\alpha_n}[\lambda_1^{\alpha_n} + \lambda_2^{\alpha_n} + \lambda_3^{\alpha_n} - 3] + \sum_{n=1}^{N}\frac{\mu_n}{\beta_n}(1 - J^{\beta_n}) \tag{2.5-20}$$

式中，μ_n、α_n 和 β_n 为材料常数。通常 Ogden 模型中取 $N=2$ 或 $N=3$。

应该注意到，超弹性模型中的材料常数不能由实验直接测量，通常是通过对单向拉伸、单向压缩和纯剪切等试验所得到的应力－应变关系曲线进行拟合计算而间接地确定的。一些通用的有限元计算程序都可以进行这种拟合。

3 弹塑性本构方程

3.1 概述

常温下金属的塑性变形行为是与变形历史相关的，一般要采用塑性流动理论，在弹塑性本构方程中建立应力速率（或增量）与应变速率（或增量）的关系。在小变形问题中，物体构形的变化很小，欧拉描述和拉格朗日描述是一致的。在这种条件下建立的是小变形弹塑性本构方程。在有限变形问题中，变形前后物体的构形发生了明显的变化，物质坐标与空间坐标之间发生了有限改变。于是，在建立有限变形条件下的弹塑性本构方程时要求它满足客观性，即对于坐标系刚性转动的不变性。有限变形条件下的弹塑性本构方程与小变形弹塑性本构方程在形式上是相同的，只是其中的应力速率和应变速率要有客观性。

在建立本构方程时，我们先考虑小变形问题，在本构方程中建立柯西应力和小应变之间的增量关系。然后将本构方程推广到有限变形的情况，这时本构方程中要采用具有客观性的应力和应变速率。客观性应力速率可以采用柯西应力的久曼导数 $\hat{\boldsymbol{\sigma}}$，客观性应变速率可采用应变速率张量 $\boldsymbol{d}$。

将物体的弹塑性变形分解为可恢复的弹性变形和不可恢复的塑性变形两部分。在小变形时，将应变增量 $\mathrm{d}\boldsymbol{\varepsilon}$ 分解如下

$$\mathrm{d}\boldsymbol{\varepsilon} = \mathrm{d}\boldsymbol{\varepsilon}^{e} + \alpha\mathrm{d}\boldsymbol{\varepsilon}^{p}$$

相应地，在有限变形时将应变速率 d 分解为

$$\boldsymbol{d} = \boldsymbol{d}^{e} + \alpha\boldsymbol{d}^{p}$$

式中上标 e 和 p 分别表示其弹性分量和塑性分量。α 为加载因子，纯弹性变形（材料屈服前的变形和卸载过程）时 $\alpha=0$，弹塑性变形时 $\alpha=1$。

3.2 屈服准则

对金属试件缓慢地进行单向均匀加载（拉伸、压缩）时，当变形很小时，材料处于弹性状态，应力和应变之间呈线性关系。而当单向应力达到屈服点（屈服应力）σ_s 时，材料发生屈服，进入塑性状态。弹性比例极限与屈服点十分接近，一般不加区别。在塑性状态，应力应变之间呈非线性关系，卸载后会留下不可恢复的永久变形（塑性变形）。常温下金属材料发生塑性变形的机理主要是位错的滑移和孪晶，在塑性变形过程中由于位错增殖使滑移阻力增大，会发生应变硬化，使塑性变形持续下去所需的应力（称为后继屈服应力或流动应力）不断增加，如图 2.5-1 所示。由于单向拉伸中试件会形成缩颈，图中真实应力和真实应变都要由缩颈处测量所得的瞬时面积计算。由图中还可看到，真实应力随应变的增加而不断增加，这种材料称为稳定材料。

当物体中应力分布不均匀的时候，各质点不是同时发生屈服的。在一般应力状态下，只有当各应力分量之间符合一定的关系时，质点才进入塑性状态，这种关系就叫屈服准则。屈服准则的数学表达式一般可以写成

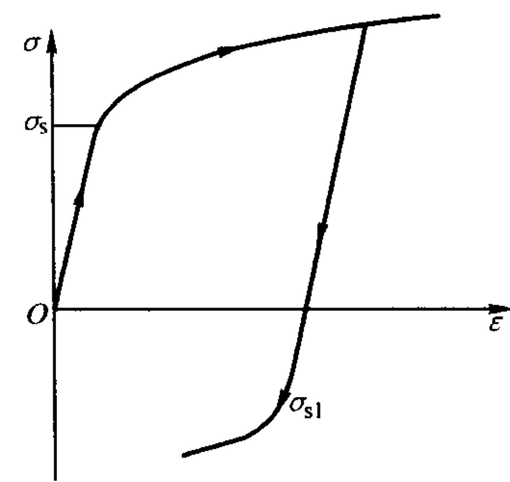


图 2.5-1 真实应力应变关系曲线

$$F = f(\sigma_{ij}) - \sigma_s = 0 \tag{2.5-21}$$

式中，$f(\sigma_{ij})$ 为应力分量的函数；σ_s 为初始屈服时的应力。

屈服准则在几何上表示应力空间中的一张曲面，称为屈服面，或屈服轨迹。当应力 σ_{ij} 位于此曲面之内，即 $F(\sigma_{ij})<0$ 时，材料处于弹性状态；当应力 σ_{ij} 位于曲面之上，即 $F(\sigma_{ij})=0$ 时，材料发生屈服而进入塑性状态。应力 σ_{ij} 不可能处于曲面之外，即不能有 $F(\sigma_{ij})>0$。

3.3 流动法则

（1）关连的流动法则

一般工程材料在发生塑性变形时，流动应力是塑性应变的增函数，这样的材料称为稳定材料。稳定材料在塑性变形中满足 $\Delta\sigma_{ij}\Delta\varepsilon_{ij}>0$。因此，若令 Y 代表后继屈服应力，则后继屈服（即加载）曲面 $F(\sigma_{ij}, Y)$ 是外凸的；并且若令应力空间和应变空间重合，则塑性应变增量（或速率）必然指向屈服面 $F(\sigma_{ij}, Y)$ 的外法线方向，即

$$\dot{\varepsilon}_{ij} = \dot{\lambda}\frac{\partial F}{\partial\sigma_{ij}} \tag{2.5-22}$$

式中，偏导数 $\partial F/\partial\sigma_{ij}$ 的几何意义就是在应力空间中屈服面 F 的外法线向量的分量；$\dot{\lambda}$ 为一标量因子。采用塑性势理论的概念，上式中的 F 可视为塑性势函数。当取屈服函数为塑性势函数时，就称为关连（或相关）的流动法则。采用不同的屈服函数，可由关连流动法则导出不同的塑性流动规律。若将塑性势函数取为屈服函数以外的其他函数，则称为非关连的流动法则。对金属材料通常采用关连的流动法则。

（2）加载－卸载准则

常温下金属材料在塑性变形中产生应变硬化，也称为加工硬化，导致后继屈服应力（也称为流动应力）增加。材料的后继屈服准则可一般地写为

$$F = f(\sigma) - Y(\bar{\varepsilon}^{p}) = 0$$

其中后继屈服应力 Y 一般视为累积塑性应变 $\bar{\varepsilon}^{p}$ 的函数，它描述了材料的加工硬化。小变形情况下，硬化材料的加载、卸载准则为

$$\left.\begin{aligned}&\text{加载：} F=0,\ \mathrm{d}F=0 \text{ 和 } \frac{\partial F}{\partial\sigma_{ij}}\mathrm{d}\sigma_{ij}>0\\&\text{卸载：} F=0,\ \mathrm{d}F<0\\&\text{中性变载：} F=0,\ \mathrm{d}F=0 \text{ 和 } \frac{\partial F}{\partial\sigma_{ij}}\mathrm{d}\sigma_{ij}=0\end{aligned}\right\} \tag{2.5-23}$$

上式中 $(\partial F/\partial\sigma_{ij})\mathrm{d}\sigma_{ij}$ 的几何意义是应力增量 $\mathrm{d}\boldsymbol{\sigma}$ 在屈服面外法线矢量 $\partial F/\partial\boldsymbol{\sigma}$ 上的投影。在加载状态，这个投影的值大于零。加载时要求满足 $F=0$ 和 $\mathrm{d}F=0$，称为一致性条件。由一致性条件可得

$$\frac{\partial F}{\partial\sigma_{ij}}\mathrm{d}\sigma_{ij} - \mathrm{d}Y = 0$$

由于加载时 $\bar{\sigma}=Y$，$\mathrm{d}\bar{\sigma}=\mathrm{d}Y$，上式又可写成

$$\frac{\partial F}{\partial \sigma_{ij}}\mathrm{d}\sigma_{ij} - \mathrm{d}\bar{\sigma} = 0$$

3.4 强化规律

(1) 两种强化假说

对于硬化材料，可以认为其后继屈服仍然服从前述的初始屈服准则，但要用后继屈服应力取代其中的初始屈服应力。后继屈服面，也称为加载面。

后继屈服轨迹的变化十分复杂。目前最常用的假说有两种。一种是各向同性硬化假说，它的要点是：①材料在硬化后仍然保持各向同性；②硬化后屈服轨迹的中心位置和形状都不变，但其大小则随变形的进行不断地扩大。另一种是随动硬化假说，其要点是：硬化后屈服轨迹的大小和形状都不变，仅在应力空间中刚性地移动。也可以将这两种假说结合起来应用。

(2) 等效应力 - 等效应变关系的函数形式

流动应力是累积塑性应变或塑性功的函数，其变化规律一般采用单一曲线假定来确定，即认为材料在简单应力状态下的应力应变关系与一般应力状态下的等效应力 - 等效应变关系是相同的。通常采用单向拉伸中真实应力和对数应变的关系来确定一般应力状态下的等效应力 - 等效应变关系。所谓等效应力 $\bar{\sigma}$ 是应力张量的一个对于坐标变换的不变量（标量），它可用来表示应力的强度。等效应变 $\bar{\varepsilon}$ 亦是如此。

工程中常用的描述等效应力 - 等效应变关系的函数形式有如下一些形式，这些方程通常仅描述塑性的等效应力 - 等效应变关系，弹性与塑性应力应变关系的协调一般可通过初始屈服应力 $\bar{\sigma}_{s0}$ 的一致性实现。

1) 理想塑性

$$\bar{\sigma} = \sigma_s$$

即忽略应变强化。

2) 线性强化

$$\bar{\sigma} = \sigma_s + K\bar{\varepsilon}$$

式中，K 为切线模量。

3) 幂函数

$$\bar{\sigma} = K\bar{\varepsilon}^n$$

式中，K 为强度系数；

n 为硬化指数。

4) 包含初应变的幂函数

$$\bar{\sigma} = K(\bar{\varepsilon}_0 + \bar{\varepsilon})^n$$

(3) 包兴格效应

一般金属材料在经过一定塑性变形后，如果受到相反方向的载荷而再次屈服，则反向屈服时的屈服应力低于初始屈服应力，而且反向屈服应力的降低程度随着原来正向应变的增加而增加，如图 2.5-1 所示。这种现象称为包兴格效应。采用随动硬化假说能较好地描述包兴格效应。

3.5 常用的塑性本构关系

(1) J_2 流动理论

假设材料是塑性各向同性的，由于塑性屈服是一种客观的物理规律，因此屈服条件可以表示为应力不变量的函数；又由于静水压力不影响材料的塑性屈服，因此屈服条件仅与应力偏量有关。于是屈服条件可以一般地表示为

$$F(J_2', J_3') = 0$$

其中 J'_2，J'_3 分别是应力偏张量 σ'_{ij} 的第二和第三不变量（注意 $J'_1 = 0$）。

对于各向同性硬化材料，可采用如下形式的密席斯屈服准则

$$F = \sqrt{\frac{3}{2}\sigma'_{ij}\sigma'_{ij}} - Y = \sqrt{3J'_2} - Y = 0 \quad (2.5\text{-}24)$$

由于该屈服准则为应力偏张量的第二不变量 J'_2 的函数，利用它推得的关连流动法则又被称为 J_2 流动理论。

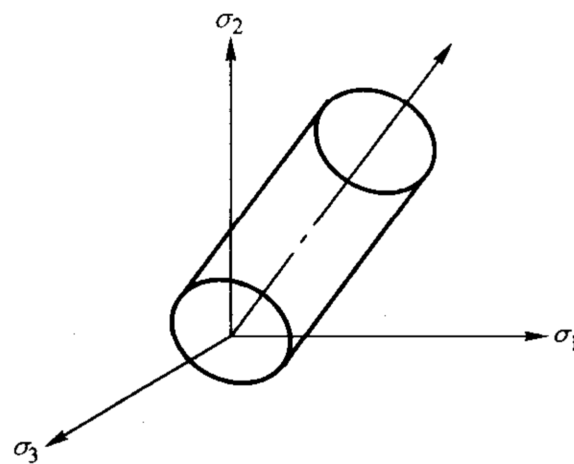


图 2.5-2 主应力空间中的密席斯屈服面

密席斯屈服准则的几何意义是主应力空间以等倾线为对称轴、以 $\sqrt{2/3}\,Y$ 为半径的圆柱面，如图 2.5-2 所示。在主应力空间中，称过原点并垂直于等倾线的平面为 π 平面，在 π 平面内密席斯服轨迹是一个圆周。

将密席斯屈服准则代入关连的流动法则得

$$\dot{\varepsilon}^p_{ij} = \dot{\lambda}\sigma'_{ij} \quad (2.5\text{-}25)$$

将上式两端自乘求和，并引入如下的等效应力 $\bar{\sigma}$ 和等效应变速率 $\dot{\bar{\varepsilon}}$

$$\bar{\sigma}^2 = \frac{3}{2}\sigma'_{ij}\sigma'_{ij},\ \dot{\bar{\varepsilon}}^2 = \frac{2}{3}\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}$$

解得

$$\dot{\lambda} = \frac{3\dot{\bar{\varepsilon}}^p}{2\bar{\sigma}}$$

将上式代入式 (2.5-25)，得到圣维南塑性流动方程

$$\dot{\varepsilon}^p_{ij} = \frac{3}{2}\frac{\dot{\bar{\varepsilon}}^p}{\bar{\sigma}}\sigma'_{ij}$$

写成增量形式为

$$\mathrm{d}\varepsilon^p_{ij} = \frac{3\mathrm{d}\bar{\varepsilon}^p}{2\bar{\sigma}}\sigma'_{ij} \quad (2.5\text{-}26)$$

(2) 随动硬化理论

假设材料的初始屈服准则为

$$F = \sqrt{3J'_2} - \sigma_{s0} = 0$$

屈服后，随着塑性变形的进行，屈服面的大小和形状不发生改变，只是在应力空间中发生刚性移动，即上式中的 σ_{s0} 为常数，这种模型称为 J_2 随动硬化模型。它可以用来描述包兴格效应。设屈服面中心移动到了 θ_{ij}，则后继屈服准则为

$$F = \sqrt{\frac{3}{2}(\sigma'_{ij} - \theta'_{ij})(\sigma'_{ij} - \theta'_{ij})} - \sigma_{s0} = 0 \quad (2.5\text{-}27)$$

式中，θ'_{ij} 是 θ_{ij} 的偏量；θ_{ij} 称为背应力。

$$\frac{\partial F}{\partial \sigma_{ij}} = \frac{3(\sigma'_{ij} - \theta'_{ij})}{2\sigma_{s0}}$$

定义等效应力为

$$\bar{\sigma}'^2 = \frac{3}{2}(\sigma'_{ij} - \theta'_{ij})(\sigma'_{ij} - \theta'_{ij})$$

当 $(\sigma'_{ij} - \theta'_{ij})\cdot\hat{\sigma}_{ij} > 0$ 且 $\bar{\sigma}' = \sigma_{s0}$ 时，$\alpha = 1$；当 $(\sigma'_{ij} - \theta'_{ij})\cdot\hat{\sigma}_{ij} \leqslant 0$ 或 $\bar{\sigma}' < \sigma_{s0}$ 时，$\alpha = 0$。

屈服面中心 θ 的移动可根据齐格勒（Ziegler）法则确定。该法侧认为，塑性变形过程中屈服面的移动速度 $\boldsymbol{\theta}$ 的方向与屈服面中心到屈服面上当前应力所在的点所构成的矢量 $\boldsymbol{\sigma} - \boldsymbol{\theta}$ 的方向一致，于是

$$\hat{\theta}_{ij} = \dot{\mu}(\sigma_{ij} - \theta_{ij})$$

式中 $\hat{\theta}$ 为 θ 的久曼导数。

由一致性条件可解得

$$\dot{\mu} = \frac{3(\sigma'_{ij} - \theta'_{ij})}{2\sigma_{s0}^2}\hat{\sigma}_{ij}$$

研究表明，将这种采用久曼速率的随动硬化本构方程用于分析简单剪切变形时，当剪切应变超过1以后，会发生切应力振荡的不合理现象。为了防止发生这种情况，可在本构方程中采用Green-Naghdi速率代替久曼速率，即用 $\boldsymbol{W}_R = \dot{\boldsymbol{R}} \cdot \boldsymbol{R}^T$ 代替旋转速率 $\boldsymbol{\omega}$。

(3) 正交各向异性材料

锻造和冲压等成形工艺中采用的坯料，一般是经轧制方法生产出来的。这些材料由于在轧制等加工过程中形成织构等原因，具有塑性各向异性，即沿不同方向其塑性性能不同。为了得到更精确的分析结果，对这些材料应该采用各向异性理论进行分析。

希尔（Hill）于1948年提出的正交各向异性材料屈服准则可表达为

$$F(\sigma_{22} - \sigma_{33})^2 + G(\sigma_{33} - \sigma_{11})^2 + H(\sigma_{11} - \sigma_{22})^2 + 2L\sigma_{23}^2 + 2M\sigma_{31}^2 + 2N\sigma_{12}^2 = 1 \tag{2.5-28}$$

式中，F、G、H、L、M、N 是各向异性参数，x_1，x_2 和 x_3 为材料各向异性主轴。令 $\bar{\sigma}$ 为等效应力，Y 为流动应力

$$\bar{\sigma} = \sqrt{\frac{3}{2}\frac{F(\sigma_{22} - \sigma_{33})^2 + G(\sigma_{33} - \sigma_{11})^2 + H(\sigma_{11} - \sigma_{22})^2 + 2L\sigma_{23}^2 + 2M\sigma_{31}^2 + 2N\sigma_{12}^2}{F + G + H}}$$

$$Y = \sqrt{\frac{3}{2(F + G + H)}}$$

可将式（2.5-28）写成如下形式

$$\bar{F} = \bar{\sigma} - Y = 0$$

根据关连流动法则

$$\mathrm{d}\varepsilon_{ij}^{\mathrm{p}} = \mathrm{d}\lambda\frac{\partial\bar{F}}{\partial\sigma_{ij}} = \mathrm{d}\lambda\frac{Y^2}{\bar{\sigma}}A_{ij}$$

其中

$$A_{11} = G(\sigma_{11} - \sigma_{33}) + H(\sigma_{11} - \sigma_{22})$$
$$A_{22} = F(\sigma_{22} - \sigma_{33}) + H(\sigma_{22} - \sigma_{11})$$
$$A_{33} = G(\sigma_{33} - \sigma_{11}) + F(\sigma_{33} - \sigma_{22})$$
$$A_{12} = A_{21} = N\sigma_{12}$$
$$A_{23} = A_{32} = L\sigma_{23}$$
$$A_{31} = A_{13} = M\sigma_{31}$$
$$\mathrm{d}\lambda = \mathrm{d}\bar{\varepsilon}^{\mathrm{p}} = \frac{YC_{mnpq}^{\mathrm{e}}A_{mn}\,\mathrm{d}\varepsilon_{pq}}{E_t + Y^2C_{mnpq}^{\mathrm{e}}A_{mn}A_{pq}}$$

其中，E_t 为等效应力-等效应变曲线的切线模量。

于是

$$\mathrm{d}\varepsilon_{ij}^{\mathrm{p}} = \frac{Y^2C_{mnpq}^{\mathrm{e}}A_{mn}\,\mathrm{d}\varepsilon_{pq}}{E_t + Y^2C_{mnpq}^{\mathrm{e}}A_{mn}A_{pq}}A_{ij}$$

对于平面应力问题，如板材成形问题，屈服准则简化为

$$\bar{F} = \frac{3}{2(F + G + H)}[(G + H)\sigma_{11}^2 + (F + H)\sigma_{22}^2 - 2H\sigma_{11}\sigma_{22} + 2N\sigma_{12}^2] - Y^2 = 0$$

上式中的各系数可利用单向拉伸中横向与厚向应变之比 $r = \varepsilon_t/\varepsilon_z$ 求得，令 r_0、r_{45} 和 r_{90} 分别代表沿轧向及与轧向夹角为45°和90°方向拉伸情况下的应变比值，σ_{s1}、σ_{s45} 和 σ_{s2} 分别为相应方向的单向拉伸屈服应力，则

$$G = \frac{1}{(1 + r_0)\sigma_{s1}^2}, F = \frac{1}{(1 + r_{90})\sigma_{s2}^2}$$

$$H = \frac{r_0}{(1 + r_0)\sigma_{s1}^2} = \frac{r_{90}}{(1 + r_{90})\sigma_{s2}^2}$$

$$N = (G + H)(\frac{1}{2} + r_{45})$$

若令 $r = r_0 = r_{45} = r_{90}$，即板材仅具有厚向异性，而在板面内是各向同性的，且 x，y 坐标方向与应力主轴方向一致，则屈服准则可进一步简化为

$$\bar{F} = \frac{3}{2(2 + r)}[(1 + r)(\sigma_1^2 + \sigma_2^2) - 2r\sigma_1\sigma_2] - Y^2 = 0$$

这时各向异性系数之间有如下关系，令 $\sigma_{s1} = \sigma_{s2} = \sigma_s$，则

$$F = G = \frac{1}{(1 + r)\sigma_s^2}, H = rF, N = (1 + 2r)F$$

$$L = M = \frac{(5 + r)}{2}F, Y^2 = \frac{3(1 + r)\sigma_s^2}{2(2 + r)}$$

由于以上屈服准则不能描述当 $r > 1$ 时，板材的双向拉伸屈服应力 σ_b 与单向拉伸屈服应力 σ_u 之比 $\sigma_b/\sigma_u < 1$ 的“异常”行为，Hill、Bassani等人又提出了形式更复杂的屈服准则，如

$$\bar{F} = |\sigma_1 + \sigma_2|^n + (1 + 2r)|\sigma_1 - \sigma_2|^n - 2(1 + r)\sigma_u^n = 0$$

不过，原来形式的屈服准则应用更方便，也更常用。

(4) 可压缩材料

在金属粉末等材料中，存在许多空隙，韧性材料在经受大塑性变形时，材料内部发生微裂纹和微空洞的形核、长大和合并。这些材料的变形过程中伴随着体积的改变。除应力偏量外，静水应力也会影响这些材料的屈服。它们的屈服准则一般地可写成如下函数形式

$$F = AJ'_2 + BJ_1^2 - CY^2 = 0 \tag{2.5-29}$$

式中，$J'_2 = (1/2)\sigma'_{ij}\sigma'_{ij}$ 为应力偏张量的第二不变量；$J_1 = \sigma_{ii}$ 为应力张量的第一不变量；Y 为基体材料的流动应力；A、B、C 为空洞体积分数 $f = V_{空洞}/(V_{基体} + V_{空洞})$ 或密度 ρ 的函数，要根据实验确定。

由关连流动法则可求得

$$\dot{\varepsilon}_{ij}^{\mathrm{p}} = \frac{3\dot{\bar{\varepsilon}}}{2\bar{\sigma}}(A\sigma'_{ij} + 2BJ_1\delta_{ij}), \dot{\varepsilon}_{kk} = \frac{9\dot{\bar{\varepsilon}}}{\bar{\sigma}}BJ_1$$

等效应力和等效应变分别定义为

$$\bar{\sigma}^2 = 3(AJ'_2 + BJ_1^2)$$

$$\dot{\bar{\varepsilon}}^2 = \frac{2}{3}\left[\frac{1}{A}\dot{\varepsilon}'_{ij}\dot{\varepsilon}'_{ij} + \frac{1}{18B}(\dot{\varepsilon}_{kk})^2\right]$$

当材料的可压缩性充分小时，$B \to 0$，同时令 $A \to 1$，$C \to 1/3$，则含空洞材料的屈服准则、等效应力和等效应变与密席斯准则情况下的定义是一致的。A、B 之间应满足关系式 $A + 3B = 1$。

当基体材料的等效应力应变关系已知时，可由塑性变形功率求得基体部分屈服应力 σ_M 的变化

$$\dot{\sigma}_M = \frac{H}{(1 - f)\sigma_M}\sigma_{ij}\dot{\varepsilon}_{ij}^{\mathrm{p}}$$

式中 H 为基体材料等效应力应变关系曲线的切线斜率。

3.6 弹塑性问题

在金属成形问题，特别是体积成形问题中，塑性变形往往大大超过弹性变形。如果同时工件受到模具很大的约束而且不考虑卸载，则可忽略弹性变形的影响而采用刚塑性材料模型。这时的本构方程就是塑性流动方程。在一些场合下，弹性变形对于材料加工过程的作用是不能忽略的，例如材料成形过程中的回弹及热处理残余应力分析等。这时必须采用弹塑性本构方程。

上面已分别给出了弹性本构方程和塑性流动法则的一些具体形式，下面根据弹塑性分解将两者结合起来，建立弹塑性问题的本构方程。小变形时的弹塑性本构方程可推导如下。取适当形式的屈服准则，使公式中的屈服应力 Y 为一次项，可使由关连流动法则表达式中的 $\mathrm{d}\lambda = \mathrm{d}\bar{\varepsilon}^{\mathrm{p}}$，即关连流动法则表达为 $\mathrm{d}\varepsilon_{kl}^{\mathrm{p}} = \mathrm{d}\bar{\varepsilon}^{\mathrm{p}}(\partial F/\partial\sigma_{kl})$，将此式代入弹性本构

方程式 (2.5-2) 得

$$d\sigma_{ij} = C^e_{ijkl}(d\varepsilon_{kl} - d\varepsilon^p_{kl}) = C^e_{ijkl}(d\varepsilon_{kl} - d\bar{\varepsilon}^p \frac{\partial F}{\partial \sigma_{kl}})$$

上式两边同乘$\partial F/\partial \sigma_{ij}$并求和得

$$\frac{\partial F}{\partial \sigma_{ij}} d\sigma_{ij} = \frac{\partial F}{\partial \sigma_{ij}} C^e_{ijkl}\left(d\varepsilon_{kl} - d\bar{\varepsilon}^p \frac{\partial F}{\partial \sigma_{kl}}\right)$$

由一致性条件 $(\partial F/\partial \sigma_{ij})\ d\sigma_{ij} = d\bar{\sigma}$，令 $H = d\bar{\sigma}/d\bar{\varepsilon}^p$，则上式可写为

$$H d\bar{\varepsilon}^p = \frac{\partial F}{\partial \sigma_{ij}} C^e_{ijkl}\left(d\varepsilon_{kl} - d\bar{\varepsilon}^p \frac{\partial F}{\partial \sigma_{kl}}\right)$$

由上式解得 $d\bar{\varepsilon}^p$ 为

$$d\bar{\varepsilon}^p = \frac{\frac{\partial F}{\partial \sigma_{ij}} C^e_{ijkl} d\varepsilon_{kl}}{H + \frac{\partial F}{\partial \sigma_{ij}} C^e_{ijkl} \frac{\partial F}{\partial \sigma_{kl}}}$$

于是得到小变形条件下关于应力增量的弹塑性本构方程的一般形式

$$d\sigma_{ij} = \left[C^e_{ijkl} - \alpha \frac{C^e_{ijmn} \frac{\partial F}{\partial \sigma_{mn}} \frac{\partial F}{\partial \sigma_{rs}} C^e_{rskl}}{H + \frac{\partial F}{\partial \sigma_{mn}} C^e_{mnrs} \frac{\partial F}{\partial \sigma_{rs}}} \right] d\varepsilon_{kl}$$

$$= (C^e_{ijkl} - \alpha C^p_{ijkl}) d\varepsilon_{kl} = C^{ep}_{ijkl} d\varepsilon_{kl} \qquad (2.5\text{-}30)$$

式中 α 为加载因子。例如，将密席斯屈服准则代入上式，就得到小变形条件下 J_2 流动理论的弹塑性本构方程

$$d\sigma_{ij} = \left(C^e_{ijkl} - \frac{2G\alpha}{g}\sigma_{ij}'\sigma_{kl}' \right) d\varepsilon_{kl} = C^{ep}_{ijkl} d\varepsilon_{kl}$$

式中，G 为切模量，$g = 2\bar{\sigma}^2\ [1 + h/\ (2G)]\ /3$。令 E_t 为单向拉伸时应力 - 应变曲线的斜率，设材料不可压缩，则 $1/H = d\varepsilon^p/d\sigma = (d\varepsilon - d\varepsilon^e)\ /d\sigma = 1/E_t - 1/E$，于是 $1/h = 3/\ (2H) = (3/2)\ (1/E_t - 1/E)$。

在有限应变条件下，应将加载 - 卸载准则，塑性流动法则和弹塑性本构方程表达式中的 $d\sigma_{ij}$ 替换为 $\hat{\sigma}_{ij}$，$d\varepsilon^p_{ij}$ 替换为 d^p_{ij}。

3.7 热弹塑性问题

如果材料加工过程中温度的变化很明显，则材料的弹性模量 E、泊松比 ν、线膨胀系数 α、屈服应力 Y 以及等效应力 - 等效应变关系曲线都是温度 T 的函数。如果在加工过程中发生相变，则要对各相分别计算并进行加权平均，还要考虑由相变造成的比容变化所产生的内应力。

(1) 热弹性本构关系

在弹性区内，应变（用向量表示）为

$$d\boldsymbol{\varepsilon} = d\boldsymbol{\varepsilon}^e + d\boldsymbol{\varepsilon}^T = d\boldsymbol{\varepsilon}^e + \boldsymbol{\alpha} dT \qquad (2.5\text{-}31)$$

式中，$d\boldsymbol{\varepsilon}^T$ 为热膨胀引起的温度应变增量，仅正应变分量不为零；$\boldsymbol{\alpha}$ 为线胀系数向量，$\boldsymbol{\alpha} = \alpha\ [1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0]^T$。

由胡克定律

$$\boldsymbol{\varepsilon}^e = (\boldsymbol{C}^e)^{-1}\boldsymbol{\sigma}$$

由于弹性矩阵依赖于温度 T，对上式微分得

$$d\boldsymbol{\varepsilon}^e = \frac{d(\boldsymbol{C}^e)^{-1}}{dT}\boldsymbol{\sigma} dT + (\boldsymbol{C}^e)^{-1} d\boldsymbol{\sigma}$$

将上式代入式 (2.5-31)，并解出 $d\boldsymbol{\sigma}$，得到

$$d\boldsymbol{\sigma} = \boldsymbol{C}^e[d\boldsymbol{\varepsilon} - (\boldsymbol{\alpha} + \frac{d(\boldsymbol{C}^e)^{-1}}{dT}\boldsymbol{\sigma})dT]$$

当弹性常数随温度的变化可忽略时，得到

$$d\boldsymbol{\sigma} = \boldsymbol{C}^e(d\boldsymbol{\varepsilon} - d\boldsymbol{\varepsilon}^T)$$

(2) 热弹塑性本构关系

在塑性区，由于流动应力 Y 依赖于温度 T，将密席斯后继屈服准则写成

$$F = \bar{\sigma} - Y\left(\int d\bar{\varepsilon}^p, T\right) = 0$$

将上式写成微分形式为

$$\frac{\partial F}{\partial \boldsymbol{\sigma}} d\boldsymbol{\sigma} = H d\bar{\varepsilon}^p + \frac{\partial Y}{\partial T} dT$$

其中

$$H = \frac{dY}{d\bar{\varepsilon}^p}$$

$$\frac{\partial F}{\partial \boldsymbol{\sigma}} = \left[\frac{\partial F}{\partial \sigma_{11}}\ \frac{\partial F}{\partial \sigma_{22}}\ \frac{\partial F}{\partial \sigma_{33}}\ \frac{\partial F}{\partial \sigma_{12}}\ \frac{\partial F}{\partial \sigma_{23}}\ \frac{\partial F}{\partial \sigma_{33}} \right]^T$$

在塑性区域内

$$d\boldsymbol{\varepsilon} = d\boldsymbol{\varepsilon}^e + d\boldsymbol{\varepsilon}^p + d\boldsymbol{\varepsilon}^T$$

把式 (2.5-31) 和塑性流动法则代入上式并解出 $d\boldsymbol{\sigma}$ 得

$$d\boldsymbol{\sigma} = \boldsymbol{C}^e\left[d\boldsymbol{\varepsilon} - \frac{\partial F}{\partial \boldsymbol{\sigma}} d\bar{\varepsilon}^p - \left(\boldsymbol{\alpha} + \frac{d(\boldsymbol{C}^e)^{-1}}{dT}\boldsymbol{\sigma}\right) dT\right]$$

将上式两端左乘$\frac{\partial F}{\partial \boldsymbol{\sigma}}$，并求和后可得

$$d\bar{\varepsilon}^p = \frac{\left(\frac{\partial F}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{C}^e d\boldsymbol{\varepsilon} - \left(\frac{\partial F}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{C}^e\left(\boldsymbol{\alpha} + \frac{d\ (\boldsymbol{C}^e)^{-1}}{dT}\boldsymbol{\sigma}\right) dT - \frac{\partial Y}{\partial T} dT}{H + \left(\frac{\partial F}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{C}^e \frac{\partial F}{\partial \boldsymbol{\sigma}}}$$

于是得到塑性区域中的增量应力应变关系

$$d\boldsymbol{\sigma} = \boldsymbol{C}^{ep}\left[d\boldsymbol{\varepsilon} - \left(\boldsymbol{\alpha} + \frac{d(\boldsymbol{C}^e)^{-1}}{dT}\boldsymbol{\sigma}\right) dT\right] + d\bar{\boldsymbol{\sigma}}^T$$

$$\approx \boldsymbol{C}^{ep}(d\boldsymbol{\varepsilon} - d\boldsymbol{\varepsilon}^T) + d\bar{\boldsymbol{\sigma}}^T \qquad (2.5\text{-}32)$$

其中 $d\bar{\sigma}^T$ 是由温度对塑性模量的影响而引起的附加应力

$$d\bar{\boldsymbol{\sigma}}^T = \frac{\boldsymbol{C}^e \frac{\partial F}{\partial \boldsymbol{\sigma}} \frac{\partial Y}{\partial T} dT}{H + \left(\frac{\partial F}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{C}^e \frac{\partial F}{\partial \boldsymbol{\sigma}}}$$

在增量计算中，$\Delta\bar{\varepsilon}^T$ 和 $\Delta\bar{\sigma}^T$ 可当作初应变和初应力处理。

(3) 相变问题

当材料中有多相共存时，各性能参数，如弹性模量 E、流动应力 Y 等，应取各相的加权平均值，即

$$A = \sum_{i=1}^{n} m_i A_i$$

式中，A 为加权平均后的参数值；n 为相数；m_i 为该时刻 i 相所占的百分比；A_i 为 i 相该参数的数值。

材料发生相变时会引起比容变化，其影响与热膨胀是类似的，可转化为线膨胀系数来计算。记 β_j 为与 j 相组织转变时比容变化对应的线膨胀系数，则

$$\beta = \sum_{j=1}^{n} m_j \beta_j$$

式中，β 为多相转变并存时的平均线胀系数，它是温度 T 的函数；n 为发生转变的相数；m_j 为在此计算步中 j 相的增量。

由此引起的应变增量为

$$\Delta\varepsilon^e_{tr} = \beta[1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0]^T \Delta T$$

(4) 加载 - 卸载准则

考虑温度对流动应力和等效应力 - 等效应变曲线的影响，将准则修改为

$$\left.\begin{aligned} &\text{加载}:(\frac{\partial F}{\partial \boldsymbol{\sigma}})^T d\boldsymbol{\sigma} + \frac{\partial F}{\partial T} dT > 0 \\ &\text{卸载}:(\frac{\partial F}{\partial \boldsymbol{\sigma}})^T d\boldsymbol{\sigma} + \frac{\partial F}{\partial T} dT < 0 \\ &\text{中性变载}:(\frac{\partial F}{\partial \boldsymbol{\sigma}})^T d\boldsymbol{\sigma} + \frac{\partial F}{\partial T} dT = 0 \end{aligned}\right\} \qquad (2.5\text{-}33)$$

4 黏塑性本构方程

受冲击载荷以及高温下成形的金属材料，其屈服应力和塑性模量随着应变速率的增加而提高，如图 2.5-3 所示，这种性质称为黏性。黏塑性材料的永久变形是与时间相关的。对于黏塑性材料，其应变速率也可分解为弹性部分和黏塑性

部分，这里我们仅分析其黏塑性部分，用上标 vp 表示。对于弹性部分的处理可参见第 3 节，我们忽略了材料的黏弹性响应。

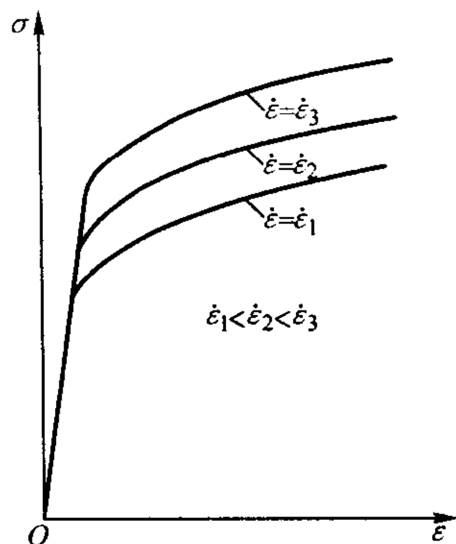


图 2.5-3　黏塑性材料的应力－应变关系曲线

4.1　一维黏塑性模型

黏塑性材料的力学模型可以一般地用图 2.5-4 所示的弹性、塑性和黏性元件的组合表示，图中的上标 e、p 和 vp 分别表示弹性、塑性和黏塑性。这种仅在塑性部分包括黏性，弹性部分不包含黏性的模型称为弹/黏塑性模型。

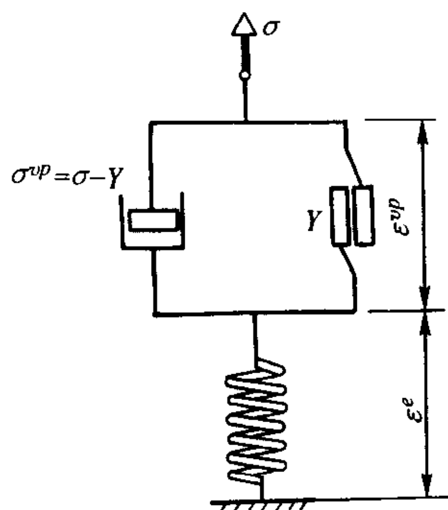


图 2.5-4　一维黏塑性模型

弹性元件的本构关系为 $\sigma = E\varepsilon$。令 Y 代表塑性元件的静态（即 $\dot{\varepsilon}$ 很小时的）屈服应力，当 $\sigma < Y$ 时，塑性元件不变形；当 $\sigma > Y$ 时，塑性元件承受的应力恒为 Y。黏性元件的本构关系为 $\sigma^{vp} = \mu\dot{\varepsilon}^{vp}$，$\mu$ 为黏性系数。

总的应力和应变速率满足如下关系

$$\sigma = \sigma^{p} + \sigma^{vp}$$

$$\dot{\varepsilon} = \dot{\varepsilon}^{e} + \dot{\varepsilon}^{vp}$$

这种材料的本构方程为

$$\left.\begin{aligned} \dot{\varepsilon} &= \frac{1}{E}\dot{\sigma}, & \text{当 } \sigma \leqslant Y \\ \dot{\varepsilon} &= \frac{1}{E}\dot{\sigma} + \frac{1}{\mu}(\sigma - Y), & \text{当 } \sigma > Y \end{aligned}\right\}$$

黏性元件承受的应力 $\sigma^{vp} = \sigma - Y$ 称为过应力，当过应力不为零时，$\dot{\varepsilon}$ 也不为零，应变 ε 将随时间而增加。

4.2　一般应力状态下的黏塑性本构方程

波兹那（Perzyna）利用塑性势理论中关连的流动法则导出了一般应力状态下的黏塑性本构方程

$$\dot{\varepsilon}_{ij}^{vp} = \gamma < \phi(F) > \frac{\partial F}{\partial \sigma_{ij}} \qquad (2.5\text{-}34)$$

式中，γ 为材料的黏性常数；F 为静态屈服函数。

记号〈　〉的含义是

$$\left.\begin{aligned} &< \phi(F) > = 0, & \text{当 } F \leqslant 0 \text{ 时} \\ &< \phi(F) > = \phi(F), & \text{当 } F > 0 \text{ 时} \end{aligned}\right\}$$

$\phi(F)$ 是过应力的函数，其具体形式要由材料试验确定，例如可取为 $\phi(F) = (\frac{F}{Y})^{n}$。当把 F 取为密席斯屈服函数时，最终可以得到

$$\dot{\varepsilon}_{ij}^{vp} = \frac{3\dot{\bar{\varepsilon}}^{vp}}{2\bar{\sigma}}\sigma_{ij}'$$

4.3　常用的黏塑性模型

黏塑性材料的流动应力是应变、应变速率和温度的函数，工程中常用的等效应力表达式有如下一些形式。

1）Backofen 模型

$$\bar{\sigma} = c\dot{\bar{\varepsilon}}^{m}$$

式中，c、m 为材料常数。

2）Rosserd 模型

$$\bar{\sigma} = k\bar{\varepsilon}^{m}\dot{\bar{\varepsilon}}^{n}$$

式中，k、m 和 n 为材料常数。

这个模型同时考虑了应变和应变速率对流动应力的影响。

3）过应力模型

$$\bar{\sigma} = Y(\bar{\varepsilon})\left[1 + \left(\frac{\dot{\bar{\varepsilon}}^{vp}}{r}\right)^{n}\right]$$

式中，$Y(\bar{\varepsilon})$ 为静态屈服应力；n、r 为材料常数。

4）幂函数

$$\bar{\sigma} = g\left(\frac{\dot{\bar{\varepsilon}}^{vp}}{\dot{a}}\right)^{m}$$

式中，g 为参考应力；$\dot{a}$ 为参考应变速率；m 为材料常数。

4.4　蠕变问题

（1）蠕变

材料在载荷作用下（例如拧紧的螺栓），经历很长的时间后，尤其是在高温环境下，即使所受应力低于屈服应力，也会发生随着时间的变化持续不断的永久变形。如果所受的载荷是由于材料本身的弹性变形而产生的（如预应力或残余应力），则随着时间的变化会持续不断地发生应力松弛，这种现象称为蠕变。蠕变与黏塑性是类似的，区别在于有无屈服点和时间的尺度不同。

一种典型的单轴蠕变曲线如图 2.5-5 所示。

初始蠕变和第三期蠕变经历的时间都很短，通常人们关心是第二期蠕变。研究退火过程中残余应力的松弛可以采用蠕变模型。

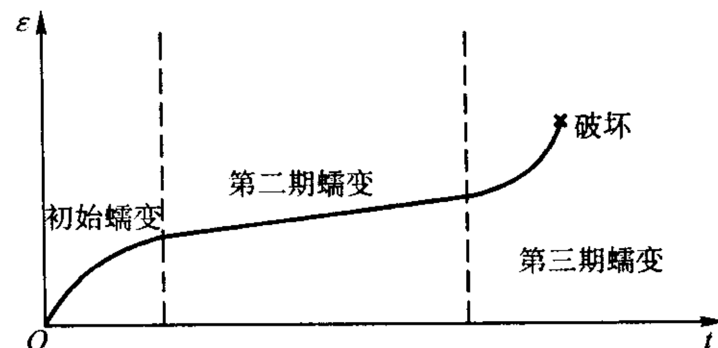


图 2.5-5　典型的单轴蠕变曲线

（2）黏弹塑性模型

在分析蠕变问题时，在弹性或塑性阶段都要考虑黏性，

这种模型称为黏弹塑性模型。

将应变增量分解如下

$$\Delta\boldsymbol{\varepsilon} = \Delta\boldsymbol{\varepsilon}^{e} + \Delta\boldsymbol{\varepsilon}^{p} + \Delta\boldsymbol{\varepsilon}^{T} + \Delta\boldsymbol{\varepsilon}^{c} \qquad (2.5\text{-}35)$$

式中，上标 e、p、T、c 分别表示该应变为弹性应变、塑性应变、温度应变和蠕变应变。

蠕变应变 $\Delta\varepsilon^{c}$ 可用蠕变应变速率 $\dot{\varepsilon}^{c}$ 表述，在计算第 i 步时

$$\Delta\boldsymbol{\varepsilon}_{i}^{c} = \dot{\boldsymbol{\varepsilon}}_{i}^{c}\Delta t_{i}(1-\theta) + \theta\dot{\boldsymbol{\varepsilon}}_{i+1}^{c}\Delta t_{i}$$

式中，Δt_i 为第 i 步的时间步长；$\dot{\boldsymbol{\varepsilon}}_{i}^{c}$、$\dot{\boldsymbol{\varepsilon}}_{i+1}^{c}$ 为当前时刻和下一时刻的蠕变速率；θ 为表示差分格式的参数，$0\leqslant\theta\leqslant1$，一般取 $\theta=1/2\sim2/3$。

设 $\dot{\bar{\varepsilon}}^{c}$ 是 $\bar{\sigma}$ 的函数，由 Taylor 展开取前两项得

$$\dot{\boldsymbol{\varepsilon}}_{i+1}^{c} = \dot{\boldsymbol{\varepsilon}}_{i}^{c} + \left.\frac{\partial\dot{\boldsymbol{\varepsilon}}^{c}}{\partial\boldsymbol{\sigma}}\right|_{i}\Delta\boldsymbol{\sigma}_{i} = \dot{\boldsymbol{\varepsilon}}_{i}^{c} + \boldsymbol{H}_{i}\Delta\boldsymbol{\sigma}_{i}$$

式中

$$\boldsymbol{H}_{i} = \left.\frac{\partial\dot{\boldsymbol{\varepsilon}}^{c}}{\partial\boldsymbol{\sigma}}\right|_{i}$$

它是一个矩阵，可根据黏塑性本构关系 $\dot{\boldsymbol{\varepsilon}}^{c}=\dot{\bar{\varepsilon}}^{c}$（$\partial F/\partial\boldsymbol{\sigma}$）$=\dot{\bar{\varepsilon}}^{c}$（$\partial\bar{\sigma}/\partial\boldsymbol{\sigma}$）求得。于是

$$\Delta\boldsymbol{\varepsilon}_{i}^{c} = \dot{\boldsymbol{\varepsilon}}_{i}^{c}\Delta t_{i} + \theta\Delta t_{i}\boldsymbol{H}_{i}\boldsymbol{\sigma}_{i} \qquad (2.5\text{-}36)$$

1）黏弹性本构方程　将式（2.5-35）、式（2.5-36）代入弹性本构方程，并令 $\Delta\boldsymbol{\varepsilon}^{p}=\boldsymbol{0}$，可解得

$$\Delta\boldsymbol{\sigma}_{i} = \hat{\boldsymbol{C}}^{e}(\Delta\boldsymbol{\varepsilon}_{i} - \Delta\boldsymbol{\varepsilon}_{i}^{T} - \dot{\boldsymbol{\varepsilon}}_{i}^{c}\Delta t_{i})$$

式中　$$\hat{\boldsymbol{C}}^{e} = [(\boldsymbol{C}^{e})^{-1} + \theta\boldsymbol{H}_{i}\Delta t_{i}]^{-1}$$

2）黏塑性本构方程　在黏弹性本构方程中计入 $\Delta\boldsymbol{\varepsilon}^{p}$，并表达为塑性流动法则的形式，可解得

$$\Delta\boldsymbol{\sigma}_{i} = \hat{\boldsymbol{C}}^{ep}(\Delta\boldsymbol{\varepsilon}_{i} - \Delta\boldsymbol{\varepsilon}_{i}^{T} - \dot{\boldsymbol{\varepsilon}}_{i}^{c}\Delta t_{i}) + \Delta\bar{\boldsymbol{\sigma}}_{i}^{T} \qquad (2.5\text{-}37)$$

式中，$\Delta\bar{\boldsymbol{\sigma}}_{i}^{T}$ 参见式（2.5-32）

而

$$\hat{\boldsymbol{C}}^{ep} = \hat{\boldsymbol{C}}^{e} - \hat{\boldsymbol{C}}^{p}$$

$$\hat{\boldsymbol{C}}^{p} = \frac{\hat{\boldsymbol{C}}^{e}\dfrac{\partial F}{\partial\boldsymbol{\sigma}}\left(\dfrac{\partial F}{\partial\boldsymbol{\sigma}}\right)^{T}\hat{\boldsymbol{C}}^{e}}{H + \left(\dfrac{\partial F}{\partial\boldsymbol{\sigma}}\right)^{T}\hat{\boldsymbol{C}}^{e}\dfrac{\partial F}{\partial\boldsymbol{\sigma}}}$$

（3）常用蠕变模型

对于不同的材料在不同条件下发生的蠕变现象，由实验确定的蠕变方程具有不同的形式，这里给出两个常用的形式。

1）单轴试验中的蠕变规律通常表示为

$$\varepsilon_{c} = A\sigma^{n}t^{m}$$

或

$$\dot{\varepsilon}_{c} = mA\sigma^{n}t^{m-1}$$

式中，m、n、A 为材料常数；t 为时间。

2）Anand 模型

$$\dot{\bar{\varepsilon}} = A_{0}(\mathrm{sh}B\bar{\sigma})^{n}e^{\frac{\Omega}{T+273}}$$

式中，$\bar{\sigma}$为等效应力；T 为温度，℃；A_0、B、Ω、n 为材料常数。

3）幂函数形式的第二期蠕变方程

$$\dot{\bar{\varepsilon}}^{c} = k_{1}\bar{\sigma}^{k_{2}}\mathrm{e}^{-\frac{k_{3}}{T}}$$

式中，k_1、k_2、k_3 为材料常数。

5　塑性细观力学本构关系

5.1　塑性细观力学的基本概念

细观又称为介观，指介于宏观（如工程结构）与微观（如原子、分子）之间的尺度，也就是多晶体中介于位错与晶粒的尺度。塑性细观力学研究在这个尺度上发生的与塑性变形有关的力学现象及有关的机理。它是连接研究原子、分子的相互作用的微观理论和研究实际的制件或试件的塑性行为的宏观塑性理论之间的中间环节。它对于揭示宏观塑性变形的机理、塑性各向异性的起因及其演化、材料内部微空洞的萌生、长大和宏观破裂等现象都有着十分重要的作用。塑性细观理论不仅为深入洞察塑性变形的机理提供了理论手段，而且在实际工程问题的分析中已经得到了初步应用。

细观力学模型的建立可分为如下步骤：

1）选择合适的尺度以表现所研究的力学现象；

2）着重考虑所研究的细观结构（如滑移应变、微空洞等），而将其周围的材料视为连续介质；

3）将上一步得到的本构方程通过平均化应用于其所在的连续介质。

5.2　晶体塑性本构方程

大多数工业用金属材料都是多晶体材料。金属晶体塑性变形的物理过程主要是位错滑移、孪晶和晶界滑动等。常温下位错滑移起着主要的作用。滑移是沿特定的晶面（原子密排面）和晶向（原子密排方向）进行的，这是这类材料塑性变形在几何学和运动学上的主要特点。这些晶面和晶向称为滑移面和滑移方向，一个滑移面和滑移方向的组合称为一个滑移系。由于在晶体内部存在着大量的位错，宏观上可以从体积平均的角度将滑移引起的位移和应变看成是均匀的、连续分布的，采用连续介质力学的方法进行处理。在后继屈服轨迹、塑性各向异性和变形诱导织构演化等方面的研究中，晶体塑性理论有其独特的优越性。我们首先介绍连续滑移模型所导出的单晶体塑性理论。然后简单介绍多晶体的处理方法。

（1）晶体塑性变形几何学

在分析单晶体的塑性变形时，如图 2.5-6 所示，对变形梯度可进行如下乘法分解

$$\boldsymbol{F} = \boldsymbol{F}^{*}\cdot\boldsymbol{F}^{p}$$

式中，$\boldsymbol{F}^{*}$ 代表弹性变形（即晶格的畸变）和刚体转动的合成；$\boldsymbol{F}^{p}$ 代表由滑移引起的塑性变形。

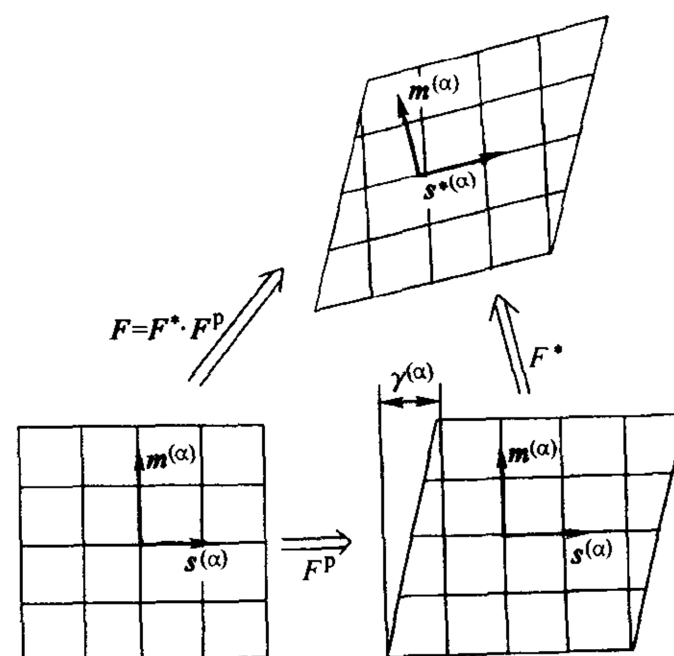


图 2.5-6　变形梯度的分解

设变形前晶体中第 α 个滑移系中沿滑移面法向和滑移方向的单位矢量分别为 $\boldsymbol{m}^{(\alpha)}$ 和 $\boldsymbol{s}^{(\alpha)}$，由于变形晶格发生畸变和刚体转动后，它们分别成为

$$\boldsymbol{s}^{*(\alpha)} = \boldsymbol{F}^{*}\cdot\boldsymbol{s}^{(\alpha)},\quad \boldsymbol{m}^{*(\alpha)} = \boldsymbol{m}^{(\alpha)}\cdot\boldsymbol{F}^{*-1}$$

$\boldsymbol{s}^{*(\alpha)}$ 和 $\boldsymbol{m}^{*(\alpha)}$ 不再是单位矢量，但仍保持正交。

与前述变形梯度的乘法分解相对应，速度梯度也可分解为分别与滑移和晶格畸变加刚体转动相对应的两部分

$$L = \dot{F}\cdot F^{-1}$$
$$= \dot{F}^{*}\cdot F^{*-1} + F^{*}\cdot \dot{F}^{p}\cdot F^{p-1}\cdot F^{*-1}$$
$$= L^{*} + L^{p}$$

其中，与滑移相对应的部分是

$$L^{p} = F^{*}\cdot \dot{F}^{p}\cdot F^{p-1}\cdot F^{*-1} = \sum_{\alpha=1}^{N}\dot{\gamma}^{(\alpha)} s^{*(\alpha)} m^{*(\alpha)}$$

式中，$\dot{\gamma}^{(\alpha)}$为切应变速率，其计算见本节第（3）小节。

另一方面，速度梯度可分解为应变速率和旋转速率之和

$$L = d + \omega$$

式中

$$d = d^{*} + d^{p}, \omega = \omega^{*} + \omega^{p}$$

上式中的塑性部分为

$$d^{p} = \sum_{\alpha=1}^{N} P^{(\alpha)}\dot{\gamma}^{(\alpha)},\ \omega^{p} = \sum_{\alpha=1}^{N} W^{(\alpha)}\dot{\gamma}^{(\alpha)}$$

式中

$$P^{(\alpha)} = \frac{1}{2}[s^{*(\alpha)}m^{*(\alpha)} + m^{*(\alpha)}s^{*(\alpha)}]$$
$$W^{(\alpha)} = \frac{1}{2}[s^{*(\alpha)}m^{*(\alpha)} - m^{*(\alpha)}s^{*(\alpha)}]$$

（2）单晶体的本构关系

滑移不影响晶体的弹性性质，单晶体的弹性本构方程为

$$\hat{\sigma}^{*} = C^{e}:d^{*}$$

式中，C^{e}为正交各向异性弹性张量，为简化计算，也可取为各向同性弹性张量；$\hat{\sigma}^{*}$为在与晶格一起旋转的直角坐标系中的柯西应力张量的变化率，为

$$\hat{\sigma}^{*} = \dot{\sigma}^{*} - \omega^{*}\cdot\sigma + \sigma\cdot\omega^{*}$$

柯西应力的久曼导数为

$$\hat{\sigma} = \dot{\sigma} - \omega\cdot\sigma + \sigma\cdot\omega$$

综合以上各式得单晶体的本构方程

$$\hat{\sigma} = C^{e}:d - \sum_{\alpha=1}^{N} R^{(\alpha)}\dot{\gamma}^{(\alpha)} \tag{2.5-38}$$

式中
$$R^{(\alpha)} = C^{e}: P^{(\alpha)} + \beta^{(\alpha)}$$
$$\beta^{(\alpha)} = W^{(\alpha)}\cdot\sigma - \sigma\cdot w^{(\alpha)}$$

（3）剪切应变速率的计算

由单晶体本构方程式（2.5-38）计算应力速率时，需要先求出各滑移系中的剪切应变速率。它是根据硬化方程来计算的。硬化方程有速率相关与速率无关两种形式。应用较为方便的是速率相关（即黏塑性）的硬化方程，常采用如下形式的幂函数

$$\dot{\gamma}^{(\alpha)} = \dot{a}^{(\alpha)}\left[\frac{\tau^{(\alpha)}}{g^{(\alpha)}}\right]\left[\left|\frac{\tau^{(\alpha)}}{g^{(\alpha)}}\right|\right]^{\frac{1}{m}-1}$$

式中，$g^{(\alpha)}$为参考切应力；$\dot{a}^{(\alpha)}$为参考切应变速率。若在整个变形过程中$\dot{\gamma}^{(\alpha)} = \dot{a}^{(\alpha)}$，则$\tau^{(\alpha)} = g^{(\alpha)}$就成为在切应变速率为$\dot{a}^{(\alpha)}$时的$\tau^{(\alpha)} - \gamma^{(\alpha)}$曲线方程；$m$为应变速率敏感指数。当$m=0$，为应变速率无关的情况。但是要注意当$m\to 0$时，计算不稳定。

当$\gamma^{(\alpha)}=0$时，令$g^{(\alpha)}$的初值为τ_0、$g^{(\alpha)}$的演化由下式确定

$$\dot{g}^{(\alpha)} = \sum_{\beta=1}^{N} h_{\alpha\beta}|\dot{\gamma}^{(\beta)}|$$

式中，$h_{\alpha\beta}$称为硬化系数，它决定了滑移系β中的滑移剪切应变对滑移系α所造成的硬化。

（4）硬化系数

晶体塑性理论中，确定硬化系数$h_{\alpha\beta}$是一项困难的任务。它是变形历史、变形温度和速度的函数，其分量也很多。对于高层错能材料，硬化主要是由滑移产生的，在实际分析中，对$h_{\alpha\beta}$可采取简化的计算方法。例如可取$h_{\alpha\beta}$为如下形式

$$h_{\alpha\beta} = h(\gamma)q_{\alpha\beta}$$

其中$\gamma = \sum_{\alpha=1}^{N}|\gamma^{(\alpha)}|$，$h(\gamma)$是累积滑移应变$\gamma$的函数，表明了硬化的总趋势，可取如下经验公式

$$h(\gamma) = \begin{cases} h_0\left(\dfrac{h_0\gamma}{n\tau_0}+1\right)^{n-1} & \text{对纯金属} \\ h_s + (h_0 - h_s)\operatorname{sech}^2\left[\dfrac{(h_0-h_s)\gamma}{\tau_s-\tau_0}\right] & \text{对合金} \end{cases}$$

而
$$q_{\alpha\beta} = q + (1-q)\ \delta_{\alpha\beta}$$

式中，h_0、h_s为初始和饱和硬化率，实际上常令$h_s=0$；n为应变硬化指数；q为潜在硬化率与自硬化率之比。

q值一般为1~1.4，可简单地取为

$$q = \begin{cases} 1 & \text{共面时} \\ 1.4 & \text{非共面时} \end{cases}$$

对非共面滑移系位错间相互作用的强度更仔细地区别处理，可以给出$q_{\alpha\beta}$的更精确的取值。

对于低层错能材料，孪晶在塑性变形中起着重要的作用，硬化规律较为复杂。

（5）多晶体塑性模型

多晶体是由大小，形状、方位各不相同的单晶晶粒集合而成的集合体。假设各相邻晶粒彼此接触，变形后不产生裂缝，则连续介质力学中的平衡方程及变形协调条件均应在集合体中满足。

从理论上说，若已知各晶粒的本构方程，则采用离散化的方法可以得到整个多晶集合体的本构关系。例如可将每个晶粒离散为一个或多个有限单元。但是塑性成形中一个制件所包含的晶粒数目极大，不能采用以上方法。

另一类方法是建立在统计平均的基础上的。假设多晶集合体是均匀的，每一点都呈现集合体的统计平均特性。多晶体中各物理量的数值可以通过对处于各个不同方位的单晶体中该物理量的数值进行体积加权平均求得。建立多晶集合体模型的方法主要有如下三种。①Sachs型模型：令集合体中所有晶粒都处于与宏观应力相同的均匀应力场中，强加了平衡条件，但在晶界处不满足协调条件。②泰勒（Taylor）型模型：令集合体中所有晶粒应变相同，都等于宏观应变，加强了协调条件，但在晶界处不满足平衡条件。③自洽理论：它是采用严格的数学、力学推导而建立起来的，能同时满足平衡条件与协调条件。其中泰勒型模型因应用方便且与实验吻合较好而应用较广泛。

晶粒方位在空间中分布的规律称为取向分布函数（ODF），一般采用欧拉角来表示。ODF可以利用由X射线衍射获得的结果经分析处理而求得，它是建立精确的多晶体模型的基础。

5.3 细观损伤理论

（1）概述

1）损伤与韧性断裂　金属材料因塑性大变形诱发的损伤破坏机制，主要是大多数弹塑性材料较易变形的基体首先在刚度较大的杂质或第二相粒子周围塑性流动并引起微空洞形核，使基体与杂质的界面破坏，或使杂质本身断裂而形成空洞。空洞随变形长大使应力场更不均匀，导致新的空洞形核或产生局部剪切带，最后空洞成片聚合而形成宏观裂纹，这就是韧性损伤破坏的过程，如图2.5-7所示。

2）损伤变量　考虑受损材料的一个微元面，其外法线为n，面积为A。该面积微元和缺陷相交的缺陷面积为A_D，$A-A_D$为微微元面的实际净面积。考虑到缺陷引起应力集中，认为有效面积$\bar{A}$比净面积更小。据此定义下述缺陷积面

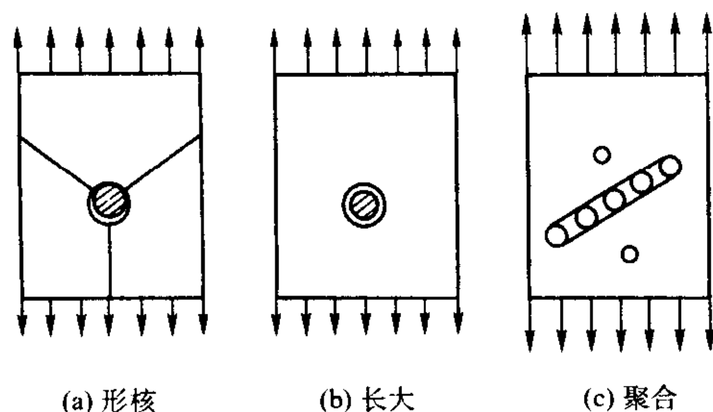


图 2.5-7 载荷作用下空洞的演化

分数 D_n 为损伤变量

$$D_n = \frac{A - \tilde{A}}{A} = 1 - \frac{\tilde{A}}{A}$$

$D_n = 0$ 表示 $\tilde{A} = A$，无损伤；$D_n = 1$，$\tilde{A} = 0$，对应于完全破坏，$0 < D_n < 1$ 表示损伤状态。为简便计，我们只讨论各向同性损伤的情况，这时 D_n 与 $\boldsymbol{n}$ 的取向无关，为一标量，简记为 D。

另一种定义损伤变量的方法是令 D 等于缺陷的体积分数f，即

$$D = 1 - \frac{\tilde{V}}{V} = 1 - \frac{\tilde{\rho}}{\rho_0} = f$$

式中，$\tilde{V}$、$\tilde{\rho}$ 为有效体积和有效密度；V、ρ_0 为名义体积和初始（无损时的）密度。

对于各向同性损伤的情况，以上这两种定义是一致的。

(2) 空洞的萌生准则

一般采用宏观意义点上的受损材料的总体应力或应变值达到临界值作为空洞萌生的准则。

1) 应力准则

$$\sigma_m + \left(\frac{2}{3} + \lambda\right)\bar{\sigma} = \sigma_c$$

式中，σ_m 为平均应力；$\bar{\sigma}$ 为等效应力；σ_c 为临界应力；λ 为用于权衡σ_m 和 $\bar{\sigma}$ 二者的参数。

2) 应变准则

$$\bar{\varepsilon} = \varepsilon_c$$

式中，$\bar{\varepsilon}$ 为全部（弹性和塑性）等效应变；ε_c 为应变的临界值。

(3) 屈服准则

Gurson 将空洞化材料看作由空洞与韧性基体所组成的集合体，再从中取出一个胞元作为集合体的统计代表进行研究。他研究了长圆柱和圆球这两种空洞模型。其中对于球型空洞模型导出的屈服准则的近似解析表达式为

$$F(\sigma, Y, f) = \left(\frac{\bar{\sigma}}{Y}\right)^2 + 2f\cosh\left(\frac{3\sigma_m}{2Y}\right) - 1 - f^2 = 0$$

式中，σ、σ_m、$\bar{\sigma}$ 为包括空洞的材料的宏观柯西应力、静水应力和等效应力；Y 为基体材料的屈服应力；f 为空洞体积分数，即前述的损伤变量 D。

Gurson 的模型是孤立地考虑了单个胞元的响应。Tvergard 考虑到空洞之间的相互作用，将 Gurson 修型修改为

$$F = \left(\frac{\bar{\sigma}}{Y}\right)^2 + 2fq_1\cosh\left(\frac{3}{2}q_2\frac{\sigma_m}{Y}\right) - (1 + q_3f^2) = 0 \tag{2.5-39}$$

式中，q_1、q_2、q_3 的建议取值为 $q_1 = 1.5$，$q_2 = 1$，$q_3 = q_1^2$。

当 $f = 0$ 时，上式退化为密席斯屈服准则。

(4) 演化方程

随着塑性变形的增加，损伤变量发生不可逆的变化。通常称损伤变量随塑性变形增长而变化的规律为演化方程，寻求演化方程是损伤力学的基本任务之一。

空洞体积分数 f（即损伤变量）的演化方程可表达为如下速率形式

$$\dot{f} = \dot{f}_{nu} + \dot{f}_{gr} + \dot{f}_{tail}$$

式中，$\dot{f}_{nu}$为由于新的空洞形核而引起的 f 的变化率；$\dot{f}_{gr}$为由于空洞长大而引起的 f 的变化率；$\dot{f}_{tail}$为由于空洞聚合而引起的 f 的变化率。

由于基体材料在塑性变形中保持体积不变，故有

$$\dot{f}_{gr} = (1 - f)\mathrm{tr}\boldsymbol{d}^p$$

$\dot{f}_{nu}$可表达为

$$\dot{f}_{nu} = A\dot{Y} + \frac{B}{3}\dot{\sigma}_m$$

式中的 A、B 可由以下方法确定。

当空洞形核取决于等效塑性应变速率时

$$A = \frac{f_n}{\sqrt{2\pi}HS}\exp\left[-\frac{1}{2}\left(\frac{\bar{\varepsilon}^p - \varepsilon_N}{S}\right)^2\right], B = 0$$

式中，ε_N 为形核速率为最大时的等效塑性应变；f_n 为空洞形核的饱和体积分数值；S 为正态分布的标准差；H 为塑性切线模量。

当空洞形核取决于界面传递的最大正应力时

$$A = B = \frac{f_n}{\sqrt{2\pi}S}\exp\left[-\frac{1}{2}\left(\frac{Y + \sigma_m - \sigma_N}{S}\right)^2\right]$$

式中，σ_N 为空洞形核速率为最大时的应力。

由于空洞聚合而导致的 $\dot{f}_{fail}$为

$$\dot{f}_{fail} = \frac{1}{\bar{H}}\frac{f_u - f_c}{\Delta\varepsilon}\dot{Y}$$

式中，f_u 为极限空洞体积分数，当 $f = f_u$ 时材料失去承受宏观应力的能力；$\Delta\varepsilon$ 为材料参数；$1/\bar{H} = 1/E_t - 1/E$；f_c 为 f 的临界值，当 $f = f_c$ 时开始发生空洞连接聚合。

(5) 流动法则

采用关连的流动法则，由上述各式可导出与式（2.5-39）相应的包含空洞的材料的宏观塑性应变速率

$$d_{ij}^p = \frac{\alpha}{\bar{G}}M_{ij}^G M_{kl}^F \hat{\sigma}_{kl}$$

其中

$$M_{ij}^G = \frac{3\sigma_{ij}'}{2Y} + \beta\delta_{ij}, M_{kl}^F = \frac{3\sigma_{ij}'}{2Y} + \mu\delta_{ij}$$

$$\beta = \frac{f}{2}q_1q_2\sinh\left(\frac{q_2\sigma_{kk}}{2Y}\right), \mu = \beta + \frac{B}{6}Y\frac{\partial F}{\partial f}$$

$$\bar{G} = \frac{Y}{2}\left\{-3\beta(1 - f)\frac{\partial F}{\partial f} - \left[\frac{\partial F}{\partial f}\left(A + \frac{1}{H}\frac{f_u - f_c}{\Delta\varepsilon}\right) + \frac{\partial F}{\partial Y}\right] \times H\frac{1}{1 - f}\left(\frac{3\sigma_{ij}'\sigma_{ij}'}{2Y^2} + \beta\frac{\sigma_{ii}}{Y}\right)\right\}$$

当空洞体积分数 f 达到一个临界值f_u 时（更一般地，当损伤因子 D 达到临界值 D_c 时），认为材料失去承载能力，即发生破裂。材料破裂时的实际空洞体积分数约为 25% 或更小。

5.4 应变梯度塑性理论

在弯曲、扭转、压印等塑性变形条件下，当试件的特征尺寸减小到一定的程度（1 ~ 10 μm 量级）以后，出现硬化率增高的现象。这种与特定尺度范围内相联系的现象称为尺度效应。出现尺度效应的尺寸称为材料的内秉尺寸，不同的金属材料其内秉尺寸也各不相同。尺度效应是与应变梯度有关的，在均匀应变状态下并不出现。经典塑性理论无法考虑尺度效应。

随着微细制造技术的发展，微细零件的加工和力学性能越来越受到重视，这就促进了应变梯度塑性理论的发展和应用。应变梯度塑性理论是建立在连续介质力学的框架中、而又能考虑尺度效应的本构关系理论，它是连接经典塑性力学

理论与原子模拟之间的必要的桥梁。由于在经典塑性力学的本构模型中不包含尺度因子，所以它不能分析尺度效应。而在应变梯度塑性理论中则引入了材料的特征尺寸和应变梯度，其中特征尺寸可以通过在该尺度范围内的各种材料试验来测定。

考虑尺度效应的一个最简单的方法是在硬化模量的计算中引入一个修正项，以考虑材料的内秉尺寸和应变梯度的影响。

Tvergard 等人对硬化模量进行如下修正。设在宏观塑性变形条件下材料的等效应力-等效应变关系为

$$\bar{\sigma} = K(\varepsilon_0 + \bar{\varepsilon}^p)^n$$

其中　ε_0——单向拉伸时的屈服应变；

$\bar{\varepsilon}^p$——塑性等效应变。

上式的增量形式为

$$\dot{\bar{\sigma}} = Kn(\varepsilon_0 + \bar{\varepsilon}^p)^{n-1}\dot{\bar{\varepsilon}} \tag{2.5-40}$$

屈服时有

$$\sigma_s = K\varepsilon_0^n = E\varepsilon_0$$

从上式可解得

$$K = E\varepsilon_0^{1-n} \tag{2.5-41}$$

将式（2.5-41）代入式（2.5-40）得

$$\dot{\bar{\sigma}} = En\left[1 + \frac{\bar{\varepsilon}^p}{\varepsilon_0}\right]^{n-1}\dot{\bar{\varepsilon}} = H\dot{\bar{\varepsilon}}$$

式中，H 为硬化模量。

考虑尺度效应时，按下式将硬化模量 H 修正为 H'

$$H' = H\left[1 + \frac{l^2\eta^2}{(\varepsilon_0 + \kappa\bar{\varepsilon}^p)^2}\right]^{1/2} \tag{2.5-42}$$

式中，l 为材料内禀尺寸；η 为最大的应变梯度分量；κ 为控制尺度效应影响程度的参数，可取为 0.1。

这种模型计算相当简单，便于使用。

编写：董湘怀（上海交通大学）

第6章 金属塑性成形中的摩擦

1 金属塑性成形中的摩擦概论

在塑性加工过程中，被加工的金属（以下简称为工件）中，位于与工模具的接触面上的任一质点，都有着相对运动或相对运动的趋势，因而在该质点上便产生了阻止其切向运动的阻力，称为（外）摩擦。在金属的塑性加工过程中，工件与工模具的接触面上的摩擦是不可避免的，如何减小或合理利用这种摩擦，是工艺设计时必须认真考虑的一个技术问题，有时甚至是工艺成败的关键。

1.1 摩擦对金属塑性加工的影响

在塑性加工过程中，摩擦对金属塑性加工的影响有正、负二面。通常负面的影响是主要的，这主要表现。

1）由于工件在接触面上的质点受到摩擦阻力的作用，流动困难，从而导致工件塑性变形不均匀，有时甚至会引发附加的拉应力，使得工件产生裂纹甚至开裂。

2）增大了塑性加工的能量和成形力。例如，在薄板拉深过程中，作用在板料法兰面上的摩擦阻力，加剧了工件的直壁部分的拉伸变薄过程。

3）使工模具摩损加剧，降低了模具的使用寿命。

4）使工件表面发生黏结、擦伤等现象，影响了产品的质量。

但是，摩擦有时候也表现出对塑性加工有利的一面。例如，模锻时利用桥部的摩擦阻力迫使坯料充满型腔；在辊锻和轧制时，依靠摩擦阻力使坯料被咬入轧辊；在充液拉深过程中，作用在工件的直壁上的高压液体增大了直壁与凸模之间的摩擦阻力，从而有效地阻止了直壁的拉伸变薄过程。

1.2 金属塑性加工时摩擦的特点

金属塑性加工中的摩擦与机械传动中的摩擦有很大差别。塑性加工中的摩擦有以下的特点。

1）这是一种伴随着金属的塑性变形而产生的摩擦。在工件的接触面上各质点的相对速度不相同，有的流动快，有的流动慢，还有的黏着不动，而且各质点的法向压力各不相同，因而各质点的摩擦也不相同。同时，工件的接触面上还会不断出现新的质点和表面，摩擦也将随之变化。

2）这是一种在高压下的摩擦，压力一般在500 MPa，而钢零件冷挤压时的压力可高达2 500 MPa，比之在重载荷下的轴承工作压力20～40 MPa要大得多，使得原先涂在工件表面上的润滑剂有很大一部分被挤掉。

3）在许多情况下，这又是一种高温下的摩擦，例如热锻时温度通常在1 000℃以上，导致工件的表面会产生氧化皮，使摩擦变得更加复杂。

1.3 目前计算摩擦应力的二个常用公式

对于金属塑性加工中的摩擦进行定量描述，是塑性成形力学中最不成熟的理论问题之一。目前，人们大都采用以下二种近似方法来描述塑性加工中的摩擦应力。

1）库仑摩擦定律　库仑摩擦定律认为，接触面上的摩擦应力 τ_f 为

$$\tau_f = \mu\sigma_n \tag{2.6-1}$$

式中，μ 是库仑摩擦系数；σ_n 是作用在质点上法向应力。

库仑摩擦定律通常用来分析金属的冷变形，例如板料冲压、冷挤压等。

2）常摩擦力条件　常摩擦力条件假设接触面上的摩擦应力 τ_f 为

$$\tau_f = mK = m\frac{\sigma_s}{\sqrt{3}} = \mu'\sigma_s \tag{2.6-2}$$

式中　m 为摩擦因子；K 为剪切屈服应力；$\mu' = m/\sqrt{3}$ 为修正的摩擦因子。

对于服从Mises屈服准则的材料，有 $K=\sigma_s/\sqrt{3}$，而 $\sigma_s = \bar{\sigma}$ 是真实应力。常摩擦力条件主要用于分析体积成形，如锻造、挤压等。

由于接触质点的摩擦应力的方向，与该质点在接触面上的相对速度 v_R 的方向相反。于是，在一般情况下，摩擦应力 τ_f 的表达式为

$$\boldsymbol{\tau}_f = -\tau_f\frac{\boldsymbol{v}_R}{|\boldsymbol{v}_R|} \tag{2.6-3}$$

2 塑性成形中摩擦的分类及机理

2.1 塑性成形中摩擦的分类

金属在塑性成形时，根据坯料与工模具的接触表面之间的润滑状态的不同，可以把摩擦分为三种类型，即干摩擦、边界摩擦和流体摩擦。

1）干摩擦　在金属塑性变形过程中，干摩擦是指不加任何润滑剂的摩擦。这时，变形金属与工模具可以视为直接接触，如图2.6-1a所示。

2）流体摩擦　当变形金属与工模具表面之间有一层完整或比较完整且比较厚的润滑油薄膜时，两表面完全被润滑油薄膜隔开，此时的润滑状态称为流体润滑，这种状态下的摩擦称为流体摩擦，如图2.6-1b所示。流体摩擦与干摩擦有着本质上的区别，其摩擦特征与所加润滑剂的性质（黏度）和相对速度梯度有关，而与接触表面的状态无关。

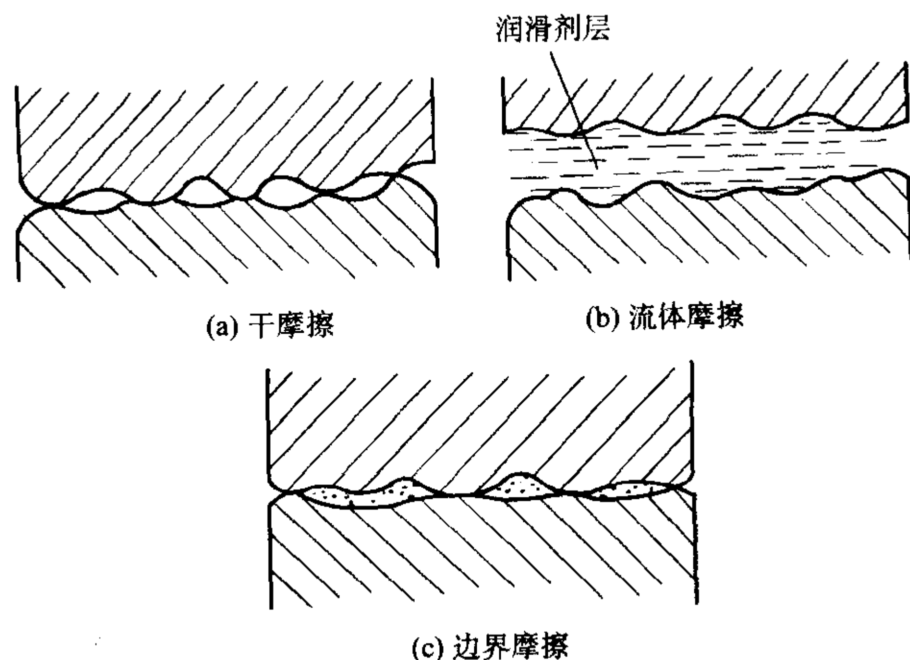


图2.6-1　摩擦分类示意图

3）边界摩擦　当坏料与工具模具之间的接触表面上加润滑剂时，随着接触压力的增加，坯料表面凸起部分被压平，润滑剂被挤入大量离散的微小凹坑中，被封存在里面（图2.6-1c），这时在压平部分与模具之间存在一层极薄的润滑膜，其厚度约为0.1 μm，接触表面就处在被这种大量离散

的微小薄膜隔开的状态，这种润滑状态称为边界润滑，这种状态下产生的摩擦称为边界摩擦，如图2.6-1c所示。

大多数金属塑性成形中的摩擦属于边界摩擦。

在实际生产中，上述三种摩擦不是截然分开的，虽然在塑性加工中多半属于边界摩擦，但有时会出现所谓的混合摩擦，即半干摩擦与半流体摩擦。半干摩擦是边界摩擦与干摩擦的混合状态；半流体摩擦是边界摩擦与流体摩擦的混合状态。

2.2　摩擦机理

塑性成形过程中摩擦的性质是复杂的，目前关于摩擦机理有三种学说。

1）表面凹凸学说　此学说认为摩擦是由于接触面上的凹凸形状引起的。因为所有经过机械加工的表面并非绝对平坦光滑的，都有不同程度的微观凸牙和凹坑。当凸凹不平的两个表面相互接触时，在压力的作用下，一个表面的“凸牙”可能会插入另一个表面的“凹坑”，产生机械咬合（见图2.6-2）。这样的接触表面在外力作用下产生相对运动时，相互咬合的凸牙部分或被切断，或使其产生剪切变形。此时摩擦力表现为这些凸牙被切断或产生剪切变形时的阻力。根据这一观点，相对接触的表面越粗糙，微“凸牙”和“凹坑”就越大，相对运动时的摩擦力就越大。降低接触表面的粗糙度，或者涂抹润滑剂以填补表面凹坑，都可起到减小摩擦的作用，对于普通粗糙程度的表面来说，这种观点已得到实践的验证。

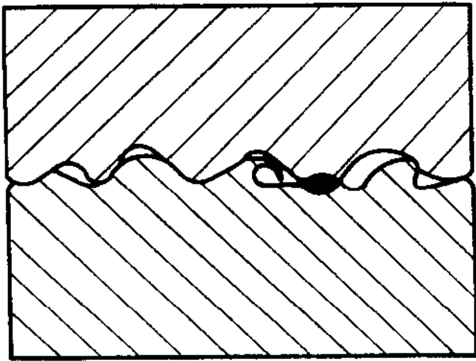


图2.6-2　接触表面凹凸不平形成机械咬合

2）分子吸附学说　当两个接触表面非常光滑时，摩擦力不但不降低，反而会提高，这一现象无法用凹凸学说来解释。这就产生了分子吸附学说，认为摩擦产生的原因是由于接触表面上分子之间相互吸引的结果。物体表面越光滑，实际接触面积就越大，分子吸附力就越强，则摩擦力也就越大。

3）黏着理论　这一理论认为，当两个表面接触时，接触面上某些接触点处压力很大，以致发生粘接或焊合，当两表面产生相对运动时，粘接点被切断而产生相对滑动。

现代摩擦理论认为，摩擦力不仅包含有接触面上的机械咬合所产生的阻力，而且包含有真实接触表面分子吸附作用所产生的黏合力及切断粘接点所产生的阻力。对于流体摩擦来说，摩擦力主要表现为润滑剂层与层之间的流动阻力。

3　影响摩擦因数的主要因素

塑性成形中的摩擦因数通常是指接触面上的平均摩擦因数。影响摩擦因数的因素很多，其主要因素分述如下。

3.1　金属的种类和化学成分

金属的种类和化学成分对摩擦因数影响很大。由于金属表面的硬度、强度、吸附性、原子扩散能力、导热性、氧化速度、氧化膜的性质以及与工模具金属分子之间相互结合力等都与化学成分有关，因此不同种类的金属及不同化学成分的同一类金属，摩擦因数是不同的。黏附性较强的金属通常具有较大的摩擦因数，如铅、铝、锌等。一般来说，材料的硬度、强度越高，摩擦因数就越小，因而凡是能提高材料的硬度、强度的化学成分都可使摩擦因数减小。对于黑色金属，随着碳的含量的增加，摩擦因数有所降低，如图2.6-3所示。

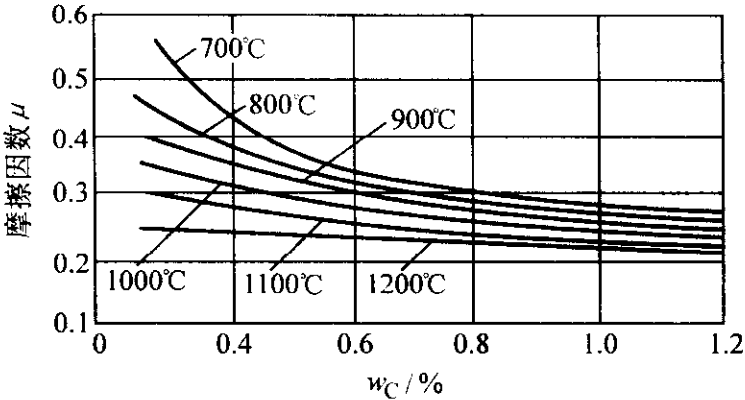


图2.6-3　钢的含碳量对摩擦因数的影响

3.2　模具的表面状态

一般来说，工模具表面越光滑，即表面凸凹不平程度越轻，这时机械咬合效应就越弱，因而摩擦因数越小。若接触表面都非常光滑，分子吸附作用增加，反而会引起摩擦因数增加，但这种情况在塑性成形中并不常见。

模具表面粗糙度在各个方向不同时，则各方向的摩擦因数亦不相同。实验证明，沿着加工方向的摩擦因数比垂直于加工方向的摩擦因数约小20%。

3.3　接触面上的单位压力

单位压力较小时，表面分子吸附作用不明显，摩擦因数保持不变，和正压力无关。当单位压力增大到一定数值后，接触表面的氧化膜被破坏，润滑剂被挤掉，这不但增加了真实接触面积，而且使坯料和模具接触面之间分子吸附作用增强，从而使摩擦因数随单位压力的增大而上升，当上升到一定程度后又趋于稳定，如图2.6-4所示。

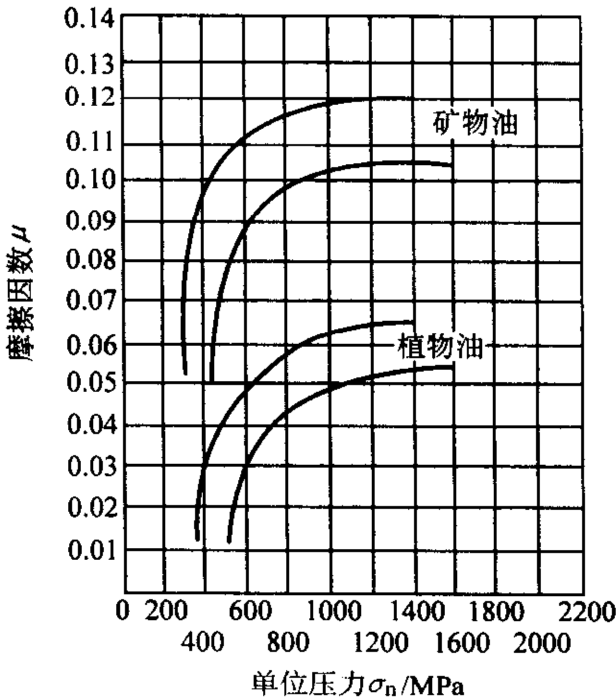


图2.6-4　正压力对摩擦因数的影响

3.4　变形温度

变形温度对摩擦因数的影响很复杂。因为变形温度变化时，材料的强度、硬度及接触面上的氧化膜的性能等都会发生变化。一般认为，变形温度较低时，摩擦因数随变形温度升高而增大，到某一温度，摩擦因数达到最大值，此后，随变形温度继续升高而降低，如图2.6-5所示。

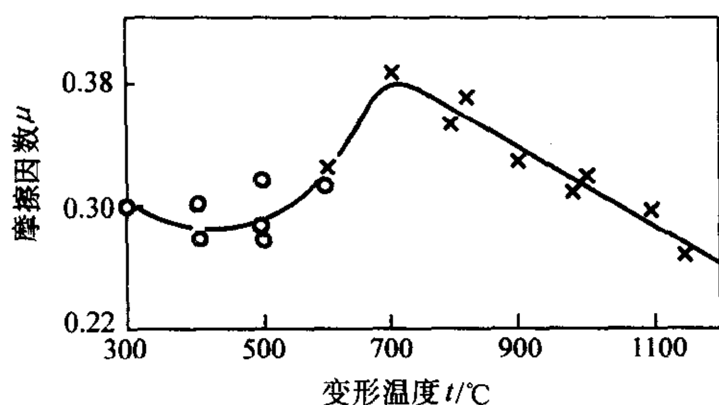


图 2.6-5　热轧时温度对碳钢摩擦因数的影响

3.5　变形速度

许多实验结果表明，摩擦因数随变形速度增加而有所下降。例如，锤上镦粗时的摩擦因数要比同样条件下压力机上镦粗时小20%～25%左右。摩擦因数降低的原因与摩擦状态有关。在干摩擦时，由于变形速度的增大，接触表面凸凹不平的部分来不及相互咬合。在边界润滑条件下，由于变形速度增加，可使润滑油膜的厚度增加，并较好地保持在接触面上，从而减少了金属坯料与模具的实际接触面积，使摩擦因数下降。图2.6-6表示冷轧铝试样时轧制速度对摩擦因数的影响。从图中可知，当变形速度增大时，不论何种润滑条件，摩擦因数都下降。

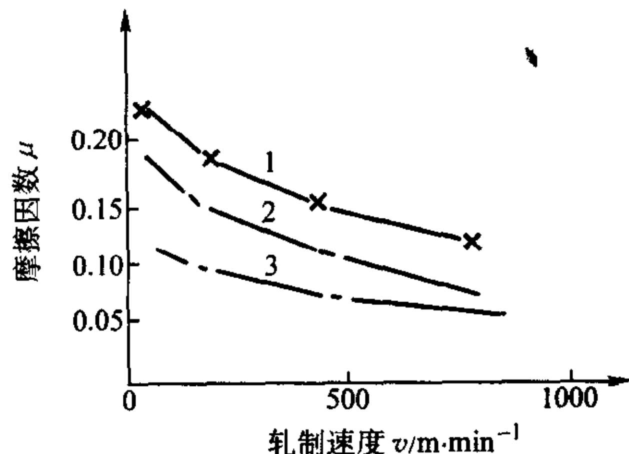


图 2.6-6　纯铝轧制时轧制速度对摩擦因数的影响
1—压下率60%，润滑油中无添加剂；
2—压下率60%，润滑油中加入酒精；
3—压下率25%，润滑油中加入酒精

4　塑性成形中的润滑

润滑是减小摩擦对塑性成形过程不良影响的最有效的措施。润滑的目的是降低接触面间的摩擦力；提高模具寿命（减小磨损、冷却模具）；提高产品质量（提高内部组织均匀性）；降低变形抗力；提高金属充满模膛的能力等。为了实现上述目的，必须选用合适的润滑剂。

4.1　塑性成形中对润滑剂的要求

塑性成形中使用的润滑剂一般应符合下述要求。

1）应有良好的耐压性能。由于塑性成形在高压下进行，因此要求润滑剂在高压作用下，润滑膜仍能吸附在接触表面上，保持良好的润滑状态。

2）应有良好的耐热性。热加工用的润滑剂在使用时应不分解、不失效。

3）应有冷却模具的作用。为了降低模具的温度，避免模具过热，提高模具寿命，要求润滑剂有冷却作用。

4）应无腐蚀作用。润滑剂不应对金属坯料和模具有腐蚀作用。

5）应无毒。润滑剂应对人体无毒、无害，不污染环境。

6）应使用、清理方便，并考虑其来源丰富，价格便宜等因素。

4.2　塑性成形中常用的润滑剂

用于塑性成形的润滑剂分液体润滑剂和固体润滑剂两大类。

（1）液体润滑剂

这类润滑剂主要包括各种矿物油、植物油、动物油、乳液和有机化合物液体等。由于动植物油（猪油、牛油、鲸油、蓖麻油、棕榈油等）中含有脂肪酸，与金属能起化学作用，附着力强，润滑性能好，但由于其化学性能不稳定，因而长期储存易变质。常用的矿物油是全损耗系统用油，它的化学稳定性好，价格便宜，来源充足。但形成稳定润滑膜的张力较差，着火点低。全损耗系统用油常与其他固体润滑剂配合使用，主要用于冷成形。乳液是由矿物油、乳化剂、石蜡、肥皂和水组成的水包油或者油包水的乳状稳定混合物，它除了具有润滑作用外，还对模具有较强的冷却作用。有机化合物液体如乙醇、十八醇等，具有游离的极性分子，与金属的结合力强，使用效果良好。它们还经常作为一种活性剂，与其他润滑剂配合使用，以提高润滑效果。

（2）固体润滑剂

这类润滑剂在常温下呈固态。根据它在高温时存在的形式，又可分为干性固体润滑剂和软（熔）化固体润滑剂。

1）干性固体润滑剂　这类润滑剂在变形过程中不改变自身的聚集状态，如石墨、二硫化钼、云母等。

① 石墨　石墨具有六方晶系的层状结构，同一层原子间的结合力比层与层间的结合力大，因而层与层之间容易滑移。当石墨作为润滑剂处于模具与金属坯料之间时，金属坯料与模具接触面间所表现的摩擦实质上是石墨层与层之间的内摩擦，这种内摩擦力比金属坯料与模具直接接触时的摩擦力要小得多，从而起到润滑作用。石墨还具有良好的导热性和热稳定性，它在540℃以上才开始氧化，因而常用作热锻、温挤的润滑剂。石墨的摩擦系数 μ 值在0.050～0.19之间。

② 二硫化钼　二硫化钼与石墨一样，具有六方晶系层状结构，其润滑原理与石墨相同，所以摩擦因数也较小，一般在0.12～0.15范围内。二硫化钼开始氧化的温度为400℃左右，比石墨低，因而在较高下度下使用时，润滑效果不如石墨。

石墨和二硫化钼是目前塑性成形中最常用的固体润滑剂，使用时可制成水剂或油剂，其比例大致是1:1。为了防止石墨、二硫化钼氧化，提高其高温下的润滑性能，常加入三氧化二硼，常用着火点较高的变压器油或炮油作为油剂。

2）软（熔）化型固体润滑剂　此类润滑剂在工作温度越过其软（熔）化点时会变软或熔化，但不燃烧，不会逸出有害气体。属于这一类的润滑剂有玻璃、珐琅、天然矿物及各种无机盐等。

① 玻璃　在高温塑性成形时，常用玻璃做润滑剂，它具有以下特点。

a）玻璃熔化后，液态玻璃包在坯料表面上，使坯料不与模具直接接触，从而起到润滑作用。同时，由于玻璃的导热性差，因而坯料的降温减少，模具也可避免过热。

b）玻璃熔化后的黏度随温度的升高而降低，不同成分的玻璃有不同的黏度-温度特性。因此可根据塑性成形的温度和所需的黏度，选择合适的玻璃成分。

c）玻璃的使用温度范围很广，在450～2 200℃范围内都可采用玻璃做润滑剂。

d）玻璃不与金属起化学作用，使用时可以粉状、薄片状或网状单独使用，也可与其他润滑剂混合作用，有着良好的润滑效果。

玻璃润滑剂的摩擦因数很小，一般在0.4～0.06之间。

② 珐琅 珐琅以粉末、含水悬浊液及酒精悬浊液的形式使用。工件变形后，可用酸洗、碱洗或其他一些专门方法除掉制品表面上的珐琅。

③ 无机盐类 无机盐类是天然的或者合成的结晶物质。在塑性成形中常用的无机盐润滑剂有盐酸盐、磷酸盐、硝酸盐等。

此外，皂类、蜡类等有机盐和硬酯酸钠、硬酯酸锌及一般肥皂也常用来作润滑剂。固体润滑剂的使用状态可以是粉末，但多数是制成糊剂或悬浮液。

4.3 润滑剂中的添加剂

为了提高润滑剂的润滑、耐磨、防腐等性能，需要在润滑剂中加入少量的活性物质，这种活性物质总称为添加剂。

润滑剂中的添加剂，一般应易溶于机油，热稳定性要好，且应具有良好的物理和化学性能。常用的添加剂有油性剂、极压剂、抗磨剂和防锈剂等。

油性剂是指天然脂、醇、脂肪酸等物质。这些物质都含有羧（COOH）类的活性基，由于它和金属表面的物理吸附作用，在金属表面形成润滑膜，起润滑和减磨作用。润滑剂中加入油性剂以后，可使摩擦因数减小。但油性剂形成的润滑膜只能在温度不高、压力较低的条件下起润滑作用，当温度较高或压力增大时，油膜容易被破坏。这时，在润滑剂中需加入极压剂。

极压剂是一种含硫、磷、氯的有机化合物，如氯化石蜡、硫化烯烃等。在高温、高压下，极压剂发生分解，分解后的产物与金属表面起化学反应，生成熔点低、吸附性强、具有片状结构的氯化铁和硫化铁等薄膜，在较高压力和较高温度下仍然起润滑作用。

抗磨剂常用的有硫化棉子油、硫化鲸鱼油等。这些物质可以分解出自由基与金属表面起化学反应生成抗腐蚀、减磨损的润滑油膜。

防锈剂常用的有石油磺酸钡，当它加入润滑剂后，在金属表面形成吸附膜，起隔水、防锈作用。

润滑剂中加入适当的添加剂后，可使摩擦系数降低，变形程度提高，金属黏模现象减少，使产品表面质量得到改善，因此目前广泛采用有添加剂的润滑剂。

塑性成形用润滑剂中常用的添加剂及其添加量可参考表 2.6-1。

表 2.6-1 润滑剂中常用的添加剂及其添加量

种类	作 用	化合物名称	添加量
油性剂	形成油膜，减小摩擦	长链脂肪酸、油酸	0.1% ~ 1%
极压剂	防止接触表面黏合	有机硫化物、氯化物	5% ~ 10%
抗磨剂	形成保护膜，防止磨损	磷酸酯	5% ~ 10%
防锈剂	防止润滑剂生锈	羧酸、酒精	0.1% ~ 1%
乳化剂	使油乳化，稳定乳液	硫酸、磷酸酯	约 3%
流动点下降剂	防止低温时油中石蜡固化	氯化石蜡	0.1% ~ 1%
黏度剂	提高润滑油黏度	聚甲基丙烯酸等聚合物	2% ~ 10%

4.4 塑性成形时的润滑方法

在金属塑性成形中，人们正逐渐采用压缩空气喷溅方法施加润滑剂。此法涂层均匀，便于机械化、自动化，劳动条件和润滑效果都较好。此外，还可结合加工的具体情况，采用以下方法。

(1) 特种流体润滑法

这种方法常用于线材拉拔，如图 2.6-7 所示，在模具人口处加一个套管，套管与坯料之间的间隙很小，并充满润滑液体。当坯料从套管中高速通过时，如模具的锥角合适且表面光洁，坯料就可把润滑剂带人模具内，金属坯料与模具之间就可得到流体润滑膜。

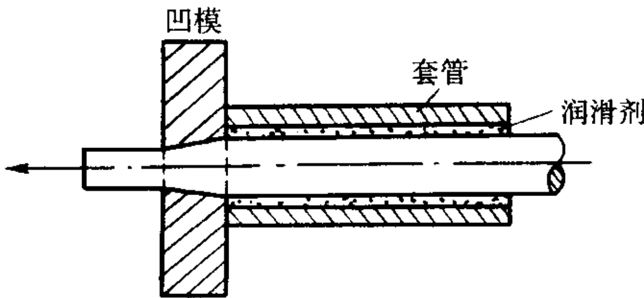


图 2.6-7 拉拔时流体强制润滑

(2) 表面磷化-皂化处理

冷挤压钢制零件时，接触面上的压力往往高达 2 000 ~ 2 500 MPa，这时需将坯料表面进行磷化处理，即在坯料表面上用化学方法制成一层磷酸盐或草酸盐薄膜，这种磷化膜由细小片状的无机盐结晶组成，呈多孔状态，对润滑剂有吸附作用。磷化膜的厚度约在 10 ~ 20 μm 之间，它与金属表面结合很牢，而且有一定的塑性，在挤压时能与钢一起变形。磷化处理后的坯料须进行润滑处理，常用的有硬脂酸钠、肥皂等，故称为皂化。磷化 - 皂化处理方法出现之后，大大推动了钢的冷挤压工艺的发展。

(3) 表面镀软金属

当加工变形抗力高的金属时，变形力大，一般的润滑剂很容易从接触表面挤出，使摩擦因数增大。在这种情况下，可在坯料表面电镀一薄层软金属，如铜或锌。这层镀层能与坯料金属很好结合，并且镀层软金属变形抗力低，延伸性好，在变形过程中，可将坯料金属与模具隔开，起润滑剂的作用。这种方法的缺点是成本高。

5 不同塑性成形条件下的摩擦因数

以下介绍在不同塑性成形条件下摩擦因数的一些数据，可供使用时参考。

1）热锻时的摩擦因数 详见表 2.6-2。

表 2.6-2 热锻时的摩擦因数

材料	坯料温度/℃	摩擦因数 μ				
		无润滑	炭末	全损耗系统用油 + 石墨		
45 钢	1 000 1 200	0.37 0.43	0.18 0.25	0.29 0.31		
锻铝	400	无润滑	气缸油 + 10% 石墨	胶体石墨	精制石蜡 + 10% 石墨	精制石蜡
		0.48	0.09	0.10	0.09	0.16

2）磷化处理后冷锻时的摩擦因数 详见表 2.6-3。

表 2.6-3 磷化处理后冷锻时的摩擦因数

压力 /MPa	摩擦因数 μ			
	无磷化膜	磷酸锌	磷酸锰	磷酸镉
7	0.108	0.013	0.085	0.034
35	0.068	0.032	0.070	0.069
70	0.057	0.043	0.057	0.055
140	0.07	0.043	0.066	0.055

3）冷拉延时的摩擦因数见表2.6-4。

表2.6-4 冷拉延时的摩擦因数

材 料	摩擦因数 μ		
	无润滑	矿物油	油+石墨
08钢	0.20～0.25	0.15	0.08～0.10
12Cr18Ni8Ti	0.30～0.35	0.25	0.15
铝	0.25	0.15	0.10
杜拉铝	0.22	0.16	0.08～0.10

4）热挤压时的摩擦因数 钢热挤压（用玻璃做润滑剂）时，$\mu=0.025\sim0.050$。有色金属热挤压时的摩擦因数见表2.6-5。

表2.6-5 热挤压时的摩擦因数

润滑	摩擦因数 μ					
	铜	黄铜	青铜	铝	铝合金	镁合金
无润滑	0.25	0.18～0.27	0.27～0.29	0.28	0.35	0.28
石墨+油	比上面相应数值降低0.030～0.035					

5）热轧时的摩擦因数 咬合时 $\mu=0.3\sim0.6$；轧制过程中 $\mu=0.2\sim0.4$。

6）拉拔时的摩擦因数 拉拔低碳钢 $\mu=0.05\sim0.07$；拉拔铜及铜合金 $\mu=0.05\sim0.08$；拉拔铝及铝合金 $\mu=0.07\sim0.11$；拉拔黄铜丝 $\mu=0.04\sim0.11$。

编写：李赞（华中科技大学）

第 7 章　金属塑性成形分析的近似解析法

金属塑性成形问题是一个非常复杂的非线性问题。求解这一问题所需的方程有应力平衡微分方程、应变速率与速度的关系方程、本构方程、屈服条件及初值与边值条件等，总数有十多个。在塑性成形的数值模拟、即有限元分析的理论、方法和相关软件开发和应用以前，人们主要通过三个近似的解析方法来求解金属塑性成形问题，这三个解析方法就是主应力法、滑移线法和上限法。

一般说来，这三个解析方法都只能分析比较简单的金属塑性成形问题，而且在求解过程中又引进了不少假设，所求得的解析解都有一定的误差。从这个意义上讲，这三个方法都属于半定量、半定性的分析方法。但是，由这三个方法求得的近似解析解能够直观地描述主要工艺参数对金属塑性成形过程的影响，且影响的趋势与工艺实际相一致，是可靠的，而这又是有限元的离散的数值解无法办到的。因此，在数值模拟技术非常成熟的今天，这三个近似方法仍有存在的价值。

1　主应力法及其应用

1.1　主应力法的基本思想

主应力法是求解金属塑性成形问题的一种简便近似方法。它的基本思想就是通过引进一些假设，建立新的能求解的常微分方程形式的应力平衡微分方程。它的具体过程如下。

1）沿着模具的作用力的方向选取一个基元块（一个宏观尺度、二个微分尺度），或选取一个单元体（三个微分尺度）。在每个面元上画出相应的应力，并假设在每个面元上应力均匀分布。

2）沿某一方向写出静力平衡方程，展开并忽略高阶微量，得一应力平衡微分方程。在主应力法中，每个应力字母都表示绝对值而非代数值。列方程时，每个应力字母前面的正、负号，由该应力的指向与所选取的方向的指向是一致还是相反来决定。

3）将正应力视为主应力，通常认为沿模具作用方向的正应力的绝对值为最大，且根据正应力的指向来确定它是拉应力还是压应力，由此确定近似的 σ_1 与 σ_3。然后代入用参数表示的 Mises 屈服准则

$$\sigma_1 - \sigma_3 = \beta\sigma_s \qquad (2.7\text{-}1)$$

中，建立近似塑性条件。再将近似塑性条件代入应力平衡微分方程中，把该方程进一步简化为只含一个未知应力的常微分方程。

4）积分求解该常微分方程，得到含有一个积分常数的表达式。

5）根据外力边值条件，确定上述积分常数。

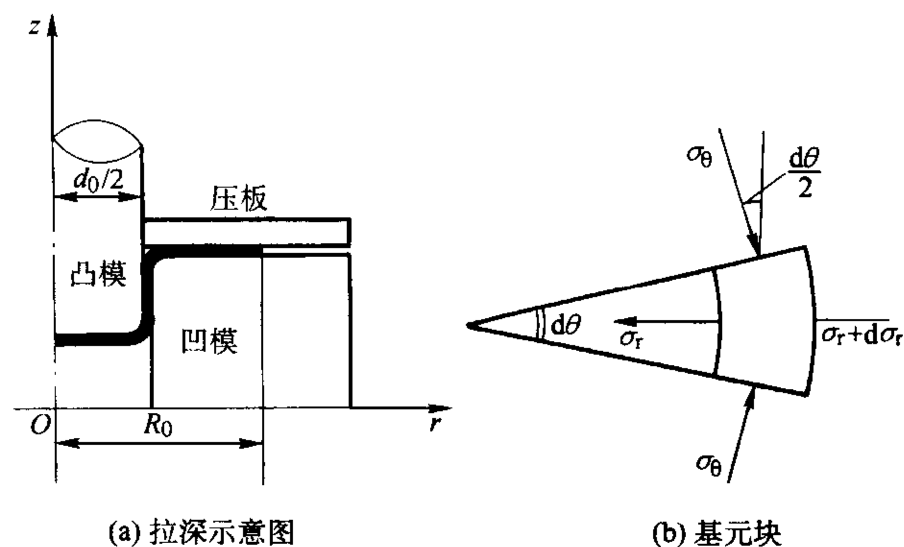
1.2　主应力法的应用举例

例 1　圆板坯拉深为圆筒件，如图 2.7-1a 所示，设板厚为 t，材料的真实应力为 σ。不计接触面上的摩擦，且忽略凹模口处的弯曲效应，试用主应力法求图示瞬间的拉深力。

解　在工件的凸缘部分取一扇形基元体，如图 2.7-1b 所示。沿负的径向的静力平衡方程为

$$\sigma_r r\mathrm{d}\theta t - (\sigma_r + \mathrm{d}\sigma_r)(r + \mathrm{d}r)\mathrm{d}\theta t - 2\sigma_\theta \sin\frac{\mathrm{d}\theta}{2}\mathrm{d}rt = 0 \qquad (a)$$

展开并略去高阶微量，可得



(a) 拉深示意图　　(b) 基元块

图 2.7-1　圆筒件拉深

$$\mathrm{d}\sigma_r = -(\sigma_r + \sigma_\theta)\frac{\mathrm{d}r}{r} \qquad (b)$$

由于 σ_r 是拉应力，σ_θ 是压应力，故 $\sigma_1 = \sigma_r$，$\sigma_3 = -\sigma_\theta$，得近似塑性条件为

$$\sigma_1 - \sigma_3 = \sigma_r + \sigma_\theta = \beta\sigma \qquad (c)$$

联解式（b）、（c），就有

$$\sigma_r = -\beta\bar{\sigma}\ln r + C \qquad (d)$$

式中的 $\bar{\sigma}$ 为 σ 的积分中值。当 $R = R_0$ 时，$\sigma_r = 0$，得 $C = \beta\bar{\sigma}\ln R_0$。最后得拉深力为

$$P = \pi d_0 t\beta\bar{\sigma}\ln\frac{2R_0}{d_0}$$

例 2　设有一球壳，外径为 R_0，内径为 r_0，如图 2.7-2a 所示。材料的真实应力为 σ，用液压胀形，试用主应力法求内压强 p。

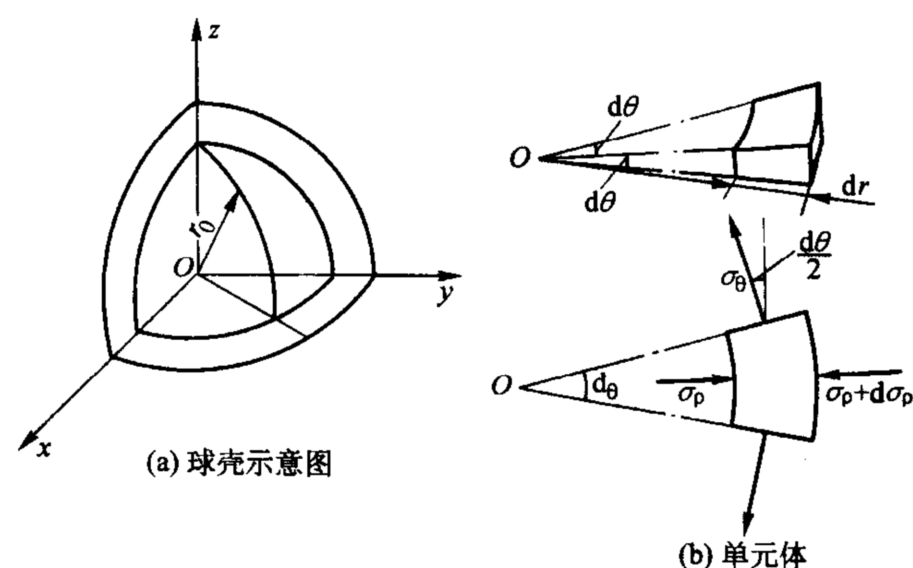


图 2.7-2　球壳液压胀形

解　用四个二二夹角为 $\mathrm{d}\theta$ 的大圆和二个同心球面切取一个单元体，如图 2.7-2b 所示。

沿径向的静力平衡方程式为

$$\sigma_\rho r^2\mathrm{d}\theta^2 - (\sigma_\rho + \mathrm{d}\sigma_\rho)(r + \mathrm{d}r)^2\mathrm{d}\theta^2 - 4\sigma_\theta \sin\frac{\mathrm{d}\theta}{2}r\mathrm{d}\theta\mathrm{d}r = 0 \qquad (a)$$

展开且忽略高阶微量，有

$$\frac{\mathrm{d}\sigma_\rho}{\mathrm{d}r} + 2\frac{\sigma_r + \sigma_\theta}{r} = 0 \qquad (b)$$

近似塑性条件为

$$\sigma_1 - \sigma_3 = \sigma_\theta - (-\sigma_\rho) = \sigma_\rho + \sigma_\theta = \beta\sigma \qquad (c)$$

故
$$d\sigma_\rho = -2\beta\bar{\sigma}\frac{dr}{r} \tag{d}$$
积分上式，可得
$$\sigma_\rho = -\beta\bar{\sigma}\ln r + C \tag{e}$$
当 $r=R_0$ 时，$\sigma_\rho=0$，得 $C=2\beta\bar{\sigma}\ln R_0$。最后得压强 p 为
$$p = \sigma_{\rho/r=r_0} = 2\beta\bar{\sigma}\ln\frac{R_0}{r_0}$$

2　塑性成形问题的滑移线解法

滑移线法是求解理想刚塑性材料的平面应变问题的精确理论，只不过在具体求解时，所给出的滑移线场都是近似的曲线族，尤其是将实际工艺抽象为理想刚塑性体的平面应变问题是一种近似表述，故本书仍称之为一种近似方法。

2.1　平面应变的特性

如果在塑性变形时，刚塑性变形体中的所有质点都仅在与某一直线（设为 z 轴）垂直的平面上流动，且质点的速度与 z 坐标无关，则称为平面应变状态。上述的这类平面称为塑性流动平面。于是有
$$\dot{u}_x = \dot{u}_x(x,y), \dot{u}_y = \dot{u}_y(x,y), \dot{u}_z = 0 \tag{2.7-2}$$
得
$$\dot{\varepsilon}_z = \dot{\gamma}_{zx} = \dot{\gamma}_{zy} = 0,\ \dot{\varepsilon}_x = -\dot{\varepsilon}_y \tag{2.7-3a}$$
及
$$\tau_{zx} = \tau_{zy} = 0,\ \sigma_z = \sigma_2 = \frac{1}{2}(\sigma_x+\sigma_y) = \frac{1}{2}(\sigma_1+\sigma_3) = \sigma_m \tag{2.7-3b}$$

由式（2.7-3b）可知，σ_z 主平面平行于 $O-xy$ 坐标平面，而 σ_1、σ_3 主平面垂直于 $O-xy$ 坐标平面。

2.2　滑移线的定义

由最大切应力理论可知，最大切应力成对出现，它所在的平面与 σ_1 面及 σ_3 面的夹角均为45°，且与 σ_2 面垂直。在平面应变问题中，由于 $\tau_{zx}=\tau_{zy}=0$，故最大切应力与 σ_2 面平行。因此，最大切应力与 σ_1 方向及 σ_3 方向的夹角均为45°，且与 $O-xy$ 坐标平面平行。

设 $O-xy$ 平面为塑性流动平面，则该平面上任一点的最大切应力都位于该平面上。在 $O-xy$ 平面上，将无限接近的最大切应力连接起来，可得二族正交曲线，曲线上任一点的切线方向即为该点处的最大切应力的方向。这二族正交的曲线族称为滑移线，它们分布在塑性变形区内，形成滑移线场。这二族曲线分别称为 α 线与 β 线，定义如下：由 σ_1 的方位线顺时针转45°到达的滑移线称为 α 线，而由 σ_3 的方位线顺时针转45°到达的滑移线称为 β 线。α 线与 β 线的正向按右旋规则人为规定，即由 α 线的正向逆时针转过90°到达 β 线的正向（见图2.7-3）。其实际方位可以由应力莫尔圆确定（见图2.7-4）。

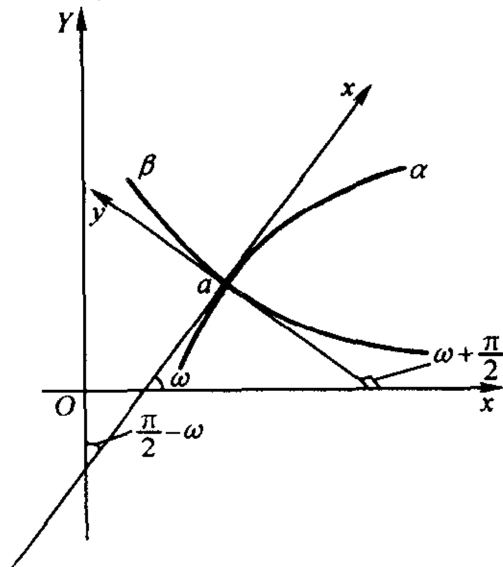


图2.7-3　滑移线与 x 轴的夹角

在塑性变形区任取一点 a，如图2.7-3所示。设 a 点处沿 α 线的切线与 x 轴的夹角为 ω，设 β 线的切线与 x 轴的夹角为 φ，且以 x 轴作为计算夹角的起始线，则有
$$\left.\begin{aligned}&\varphi = \omega + \frac{\pi}{2}\\&\frac{dy}{dx} = \tan\omega,沿\ \alpha\ 线\\&\frac{dy}{dx} = \tan\varphi = -\cot\omega,沿\ \beta\ 线\end{aligned}\right\} \tag{2.7-4}$$

2.3　滑移线场的基本方程

由应力莫尔圆可知，各应力分量与平均应力 σ_m、ω 的关系如下
$$\begin{aligned}\sigma_x &= \sigma_m - K\sin2\omega\\\sigma_y &= \sigma_m + K\sin2\omega\\\tau_{xy} &= K\cos2\omega\end{aligned} \tag{2.7-5a}$$
于是，各主应力与 σ_m、K 的关系为
$$\left.\begin{aligned}&\sigma_1 = \sigma_m + K, \sigma_3 = \sigma_m - K\\&\sigma_1 - \sigma_3 = 2K\end{aligned}\right\} \tag{2.7-5b}$$
上二式中，K 为剪切屈服应力。不难验证，由式（2.7-5a）、式（2.7-5b）确定的应力张量恒满足 Mises 屈服准则。

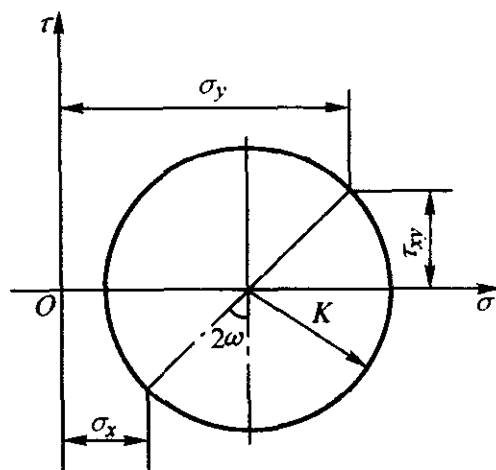


图2.7-4　应力莫尔圆

在以滑移线为坐标曲线的正交曲线坐标系中，应力平衡微分方程为
$$\left.\begin{aligned}&\frac{\partial\sigma_m}{\partial S_\alpha} - 2K\frac{\partial\omega}{\partial S_\alpha} = 0,沿\ \alpha\ 线\\&\frac{\partial\sigma_m}{\partial S_\beta} + 2K\frac{\partial\omega}{\partial S_\beta} = 0,沿\ \beta\ 线\end{aligned}\right\} \tag{2.7-6}$$

将式（2.7-6）的第一式对 α 积分，第二式对 β 积分，可得
$$\left.\begin{aligned}&\sigma_m - 2K\omega = \xi(\beta_0),沿\ \alpha\ 线\\&\sigma_m + 2K\omega = \eta(\alpha_0),沿\ \beta\ 线\end{aligned}\right\} \tag{2.7-7}$$
式（2.7-7）称为汉基（Hencky）应力方程。

2.4　滑移线场的特征

1）沿线特征　沿同一条滑移线上，平均应力的变化与滑移线的转角成正比，比例系数为 $2K$，即有
$$\left.\begin{aligned}&\sigma_{ma} - \sigma_{mb} = 2K(\omega_a - \omega_b),沿\ \alpha\ 线\\&\sigma_{ma} - \sigma_{mb} = -2K(\varphi_a - \varphi_b),沿\ \beta\ 线\end{aligned}\right\} \tag{2.7-8}$$

2）跨线特征　设同一族滑移线（例如 β 族）的两条滑移线 β_1 和 β_2，与另一族的两条滑移线 α_1 和 α_2 相交（参看图2.7-5），则两对交点之间的 $\Delta\omega$ 与 $\Delta\sigma_m$ 保持不变，即有
$$\left.\begin{aligned}&\Delta\omega = \omega_{1,1} - \omega_{2,1} = \omega_{1,2} - \omega_{2,2} = 常数\\&\Delta\sigma_m = \sigma_{m(1,1)} - \sigma_{m(2,1)} = \sigma_{m(1,2)} - \sigma_{m(2,2)} = 常数\end{aligned}\right\} \tag{2.7-9}$$

2.5　常见滑移线的类型

1）二族正交直线，为均匀应力场。

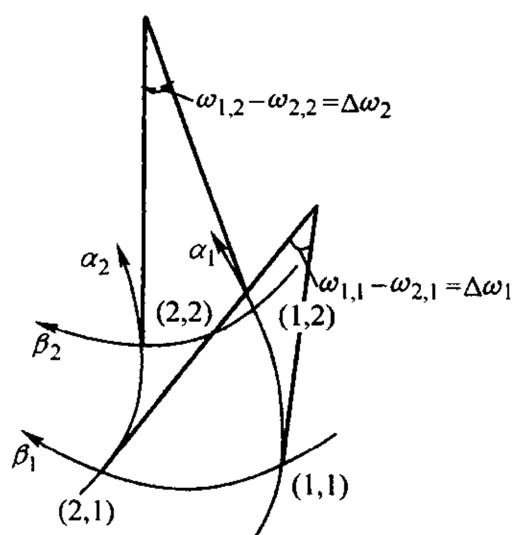


图 2.7-5　滑移线节点切线间的交角

2）一族滑移线为直线，另一族为与直线正交的曲线，为简单应力场。常见的简单应力场有①由直线族与同心圆族组成，圆心为奇点，应力不确定。②一族为某一包络线的切线，另一族为渐开线（见图 2.7-5）。

简单应力场只能与均匀应力场相邻，且交界线必定是一条公共滑移线。

2.6　应用实例

应用滑移线法求解刚塑性体的平面应变问题，首先须画出滑移线场，并找出 σ_1 与 σ_3 的作用线为已知的一点，由此来确定 α 线与 β 线，再正确运用汉基方程的推论。以下举二个应用实例。

例 1　求光滑冲头压入半无限体时的单位流动压力 p。

解　半无限体是指锻件的宽度比冲头的宽度大得多，且冲头的长度比宽度大得多，可视为平面应变。

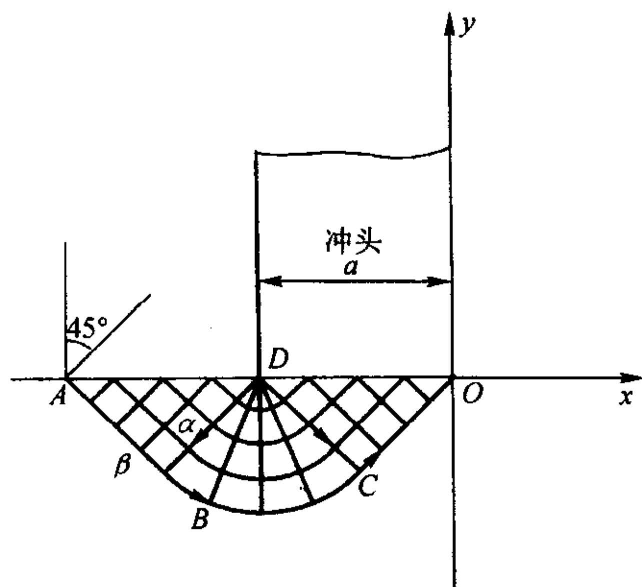


图 2.7-6　平面冲头压入半无限体

在图示的点 A 处，有 $\sigma_y=0$，$\sigma_x<0$，故 $\sigma_1=\sigma_y=0$，$\sigma_3=\sigma_x$，将直线在 A 点处由 Oy 轴的方向顺时针转 45°沿线，得 α 线，进而确定 β 线，如图 2.7-6 所示。由

$$\sigma_1-\sigma_3=2K,(\sigma_1+\sigma_3)/2=\sigma_m$$

得

$$\sigma_3=-2K,\ \sigma_{mA}=-K$$

及

$$\varphi_A=-\pi/4$$

在点 C 处，有 $\varphi_C=\pi/4$。于是，沿 β 线可有

$$\sigma_{mA}-\sigma_{mC}=-2K(\varphi_A-\varphi_C)$$

得

$$\sigma_{mC}=-K\ (1+\pi)$$

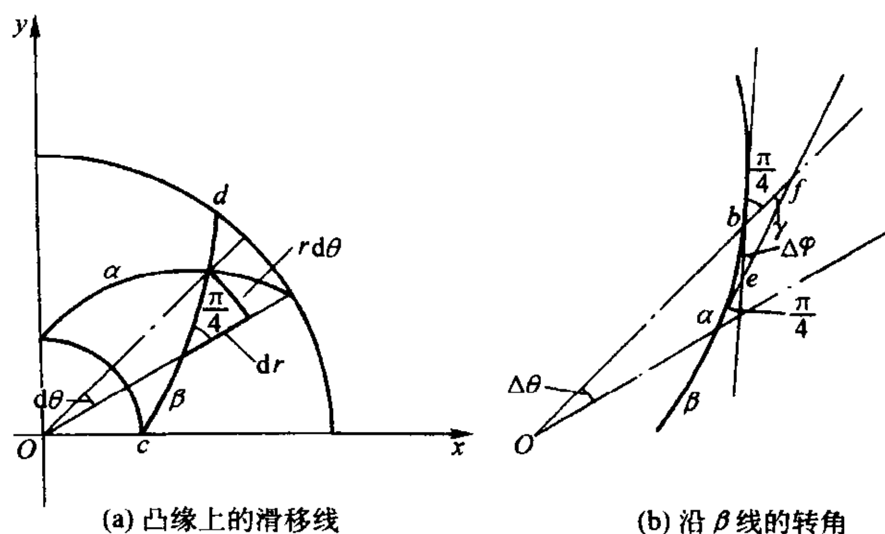
由于 ΔDCO 为均匀应力场，故 $\sigma_{mD}=\sigma_{mC}$。由 $\overline{CO}$ 线逆时针转 45°，可知在点 C 处沿 y 向的应力为 $\sigma_{yC}=\sigma_{3C}$，即

$$\sigma_{yC}=\sigma_{mC}-K=-K(2+\pi)$$

得

$$p=\frac{1}{2a}\left|\sigma_{mC}\right|2a=K\ (2+\pi)$$

例 2　不计摩擦应力及弯曲效应，试用滑移线法求图 2.7-7a 所示的圆板坯拉深为圆筒件的单位流动压力。



(a) 凸缘上的滑移线　　(b) 沿 β 线的转角

图 2.7-7

解　如果在拉深过程中，工件的板厚 t 不变，可设 $\dot{u}_z=0$，为平面应变问题。由于 $\tau_{r\theta}=0$，可知 σ_r 与 σ_θ 都是主应力，且 $\sigma_r>0$，$\sigma_\theta<0$，得 $\sigma_1=\sigma_r$，$\sigma_3=\sigma_\theta$。由此可得 α、β 线如图 2.7-7a 所示。于是，在任一点处，滑移线与径线的夹角均为 $\pi/4$，可有

$$\frac{r\mathrm{d}\theta}{\mathrm{d}r}=\tan\frac{\pi}{4}=1,\mathrm{d}\theta=\frac{\mathrm{d}r}{r}$$

得

$$\theta=\ln r+C$$

由 Δaof 与 Δbef 分别可得

$$\Delta\theta+\gamma=\frac{\pi}{4},\Delta\varphi+\gamma=\frac{\pi}{4}$$

式中的 $\Delta\varphi$ 是沿 β 线上过 a 点的切线与过 b 点的切线的夹角。故

$$\Delta\varphi=\Delta\theta$$

得

$$\Delta\varphi_{ab}=\Delta\theta_{ab}=\theta_b-\theta_a=\ln r_b-\ln r_a=\ln\frac{r_b}{r_a}$$

在外缘点 d 处，$\sigma_1=\sigma_r=0$，$\sigma_3=\sigma_\theta=-2K$，得 $\sigma_{md}=-K$。于是有

$$\sigma_{md}-\sigma_{mc}=-2K(\varphi_d-\varphi_c)=-2K\Delta\varphi_{cd}=-2K\ln\frac{r_d}{r_c}$$

得

$$\sigma_{mc}=\sigma_{md}+2K\ln\frac{r_d}{r_c}=-K+2K\ln\frac{D_0}{d_0}$$

$$\sigma_{rc}=\sigma_{1c}=\sigma_{mc}+K=2K\ln\frac{D_0}{d_0}$$

式中的 D_0 为凸缘的外径。于是，单位流动压力为

$$p=\frac{\left|\sigma_{rc}\right|\pi d_0t}{\pi d_0t}=2K\ln\frac{D_0}{d_0}$$

3　求解塑性成形问题的上限法

用动可容速度场近似代替精确速度场求解塑性成形问题的方法称为上限法。上限法是一种能比较全面地求解金属塑性成形问题的近似方法，由它所确定的载荷是真实载荷的上界。上限法的理论基础是虚功率原理。

3.1　虚功率原理的新形式

在分析金属塑性成形的过程中，包括工件自重在内的体积力通常被忽略，而作用在工件上的外力只有工件与模具接触面上的表面力。由此，虚功率原理的一般形式为

$$\int_V\sigma_{ij}\dot{\varepsilon}_{ij}\mathrm{d}V+\int_V\rho a\cdot\dot{u}\mathrm{d}V=\int_{s_c}p\cdot\dot{u}\mathrm{d}S\quad(2.7\text{-}10)$$

式中，ρ 为工件的密度；$\dot{u}$ 与 a 分别为工件质点的速度与加速度。

按照塑性加工时模具的运动属性，可将模具分为二类。

1）运动模具，即作功模具，它与工件的接触面记为 S_{cl}。

2）静止模具，不作功，它与工件的接触面记为 S_{c2}。

任取工件接触面上的一个质点，由运动学可知，该接触质点的速度（指绝对速度）$\dot{u}$ 可表示为

$$\dot{\boldsymbol{u}} = \boldsymbol{v}_e + \boldsymbol{v}_R = \boldsymbol{v}_D + \boldsymbol{v}_R \tag{a}$$

式中，$\boldsymbol{v}_e$ 为接触质点的牵连速度，它等于模具的速度 $\boldsymbol{v}_D$；$\boldsymbol{v}_R$ 为接触质点对于模具表面的相对速度，它位于模具表面在该点处的切平面上。

已知作用在工件的任一接触质点上的表面力强度 p 为

$$\boldsymbol{p} = \boldsymbol{p}_n + \boldsymbol{\tau}_f \tag{b}$$

式中，$\boldsymbol{p}_n$ 为正压力，沿着模具表面在该点处的法线方向；$\boldsymbol{\tau}_f$ 为摩擦应力，位于模具表面在该点处的切平面上，且与相对速度 $\boldsymbol{v}_R$ 方向相反，于是有

$$\boldsymbol{\tau}_f \cdot \boldsymbol{v}_R = -|\boldsymbol{\tau}_f|\ |\boldsymbol{v}_R| = -\tau_f|\boldsymbol{v}_R| \tag{c}$$

在接触面 S_{c1} 上，有

$$\boldsymbol{p} \cdot \dot{\boldsymbol{u}} = \boldsymbol{p} \cdot \boldsymbol{v}_D + \boldsymbol{p} \cdot \boldsymbol{v}_R = \boldsymbol{p} \cdot \boldsymbol{v}_D - \tau_f|\boldsymbol{v}_R| \tag{d}$$

已知在 S_{c2} 上，有 $\boldsymbol{v}_D = 0$，故在 S_{c2} 上有

$$\boldsymbol{p} \cdot \dot{\boldsymbol{u}} = -\tau_f|\boldsymbol{v}_R| \tag{e}$$

及

$$\int_V \rho \boldsymbol{a} \cdot \dot{\boldsymbol{u}} \mathrm{d}V = \int_V \frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}\rho \dot{u}^2\right)\mathrm{d}V = \frac{\mathrm{D}}{\mathrm{D}t}\int_V \frac{1}{2}\rho \dot{u}^2 \mathrm{d}V \tag{2.7-11}$$

于是，式（2.7-10）的等价形式为

$$\int_{S_{c1}} \boldsymbol{p} \cdot \boldsymbol{v}_D \mathrm{d}S = \int_V \sigma_{ij}\dot{\varepsilon}_{ij}\mathrm{d}V + \frac{\mathrm{D}}{\mathrm{D}t}\int_V \frac{1}{2}\rho \dot{u}\mathrm{d}V + \sum_{i=1}^{2}\int_{S_{ci}} \tau_f|\boldsymbol{v}_R|\mathrm{d}S \tag{2.7-12}$$

上式的物理意义为，在金属塑性变形的任一瞬时，运动模具所作的功率，等于工件的弹塑性变形功率、动能变化率及接触面上的摩擦所消耗的功率之和。对于金属塑性加工而言，在 S_{c1} 上 $\boldsymbol{v}_D$ 是常矢量，故有

$$\int_{S_{c1}} \boldsymbol{p} \cdot \boldsymbol{v}_D \mathrm{d}S = \left(\int_{S_{c1}} p\mathrm{d}S\right) \cdot v_D = P \cdot v_D \tag{f}$$

式中，P 是塑性加工时的成形力。于是，式（2.7-12）又可写为

$$P \cdot \boldsymbol{v}_D = \int_V \sigma_{ij}\dot{\varepsilon}_{ij}\mathrm{d}V + \frac{\mathrm{D}}{\mathrm{D}t}\int_V \frac{1}{2}\rho \dot{u}\mathrm{d}V + \sum_{i=1}^{2}\int_{S_{ci}} \tau_f|\boldsymbol{v}_R|\mathrm{d}S \tag{2.7-13}$$

在分析塑性加工问题时，由于一般情况下惯性力远小于成形力，所以式（2.7-13）中动能变化率一项一般忽略不计。

3.2　间断概念、动可容速度场

从这一小节开始，以下的讨论仅限于刚塑性材料。

一般说来，金属塑性变形时，变形体内的应力与速度的变化都是连续的。但是，若在某一点或某一曲面片的一个微小邻域内，应力或速度的变化很剧烈时，可简化为间断问题来处理。例如，对于沿曲面片的间断问题，就是设某个力学量的值在曲面片的两侧发生了跳跃。

动可容速度场是用上限法求解金属塑性成形问题的切入点。如果速度场 $\dot{\boldsymbol{u}}^*$ 满足以下三个条件：①速度边界条件，即在 Su 上，有 $\dot{\boldsymbol{u}}^* = \bar{\boldsymbol{u}}$；②体积连续性条件，不产生重叠和开裂；③体积不可压缩条件；则称 $\dot{\boldsymbol{u}}^*$ 为运动学容许的速度场，简称为动可容速度场。但是，与这速度场 $\dot{\boldsymbol{u}}^*$ 对应的应力场 $\boldsymbol{\sigma}_{ij}^*$ 不一定满足应力平衡微分方程和外力边界条件。所谓对应，是指应力场 σ_{ij}^* 与由 $\dot{\boldsymbol{u}}^*$ 产生的应变速率场 $\dot{\varepsilon}^*_{i,j} = (\dot{u}^*_{i,j} + u^*_{j,i})/2$ 满足塑性力学中的本构方程，对于刚塑性问题，就是满足 Mises 本构方程。

3.3　应力间断

设 S_L 为应力间断面，将变形体 V 沿 S_L 面分成 V_1 与 V_2 二部分，如图 2.7-8 所示。任取一个跨越间断面 S_L 的微元体，在微元体上，切向正应力 σ_t^+ 与 σ_t^- 不连续，即 $\sigma_t^+ \neq \sigma_t^-$ 但法向应力 σ_n 与切应力 τ_n 必定连续，即

$$\tau_n^+ + \tau_n^- = 0, \sigma_n^+ + \sigma_n^- = 0$$

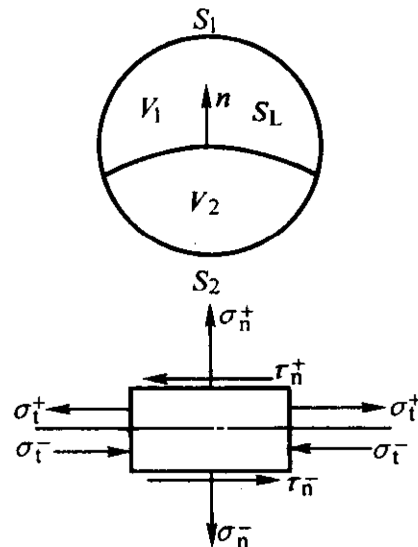


图 2.7-8　应力间断面

同时，切向正应力本身保持平衡。

在有应力间断面的条件下，虚功率方程式（2.7-13）保持不变。

3.4　速度间断

设 S_D 为速度间断面，将变形体 V 沿 S_D 面分成 V_1 与 V_2 二部分。由于塑性变形时沿 S_D 面不开裂也不重叠，故有

$$\dot{u}_n^{(1)} = \dot{u}_n^{(2)} \tag{a}$$

但切向速度可以发生间断，如图 2.7-9 所示，$\dot{u}_t^{(1)} \neq \dot{u}_t^{(2)}$。

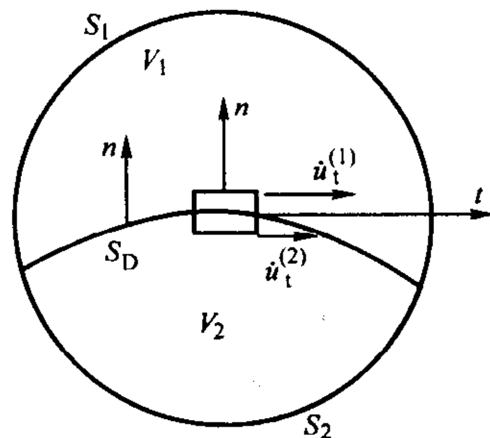


图 2.7-9　速度间断面

在速度间断的条件下，有

1）沿 S_D 面的切应力 τ 达到剪切屈服应力 K。

2）沿 S_D 面多消耗的功率为 $\int_{S_D} K|V_t|\mathrm{d}S$，式中的 $|V_t|$ 为沿 S_D 面两侧的切向速度的跳跃量，即

$$|V_t| = |\dot{u}_t^{(1)} - \dot{u}_t^{(2)}| \tag{b}$$

于是，当有速度间断面 S_D 时，式（2.7-13）变为

$$P \cdot v_D = \int \sigma_{ij}\dot{\varepsilon}_{ij}\mathrm{d}V + \sum_i \int_{S_{ci}} \tau_f|v_R|\mathrm{d}S + \sum_i \int_{S_{Di}} K|V_t|\mathrm{d}S \tag{2.7-14}$$

上式右边的最后一项表示因内摩擦多消耗的功率。

3.5　上限法基本方程

设有一刚塑性变形体 V（现时构形），σ_{ij} 是真实的应力场，$\bar{\sigma}$ 是真实的等效应力场。又设 $\dot{u}^*$ 是任一动可容速度场，$\dot{\varepsilon}^*_{ij}$ 与 σ_{ij}^* 是互相对应的应变速率场与应力场。由最大塑性功原理可得

$$\overline{\sigma\dot{\varepsilon}} = \sigma_{ij}^{*}\dot{\varepsilon}_{ij}^{*} \geqslant \sigma_{ij}\dot{\varepsilon}_{ij}^{*} \quad (a)$$

于是有

$$\int_V \overline{\sigma\dot{\varepsilon}}\,\mathrm{d}V + \sum_i \int_{S_{ci}} \tau_f |v_R^*| \mathrm{d}S + \sum_i \int_{S_{Di}} K|v_t^*| \mathrm{d}S \geqslant \int_V \sigma_{ij}\dot{\varepsilon}_{ij}^*\,\mathrm{d}V + \sum_i \int_{S_{ci}} \tau_f |v_R^*| \mathrm{d}S + \sum_i \int_{S_{Di}} K|v_t^*| \mathrm{d}S \quad (b)$$

已知 $v_D^* = v_D = \dot{u}_0$，记与动可容速度场 $\dot{u}$ 对应的近似载荷为 P^*。

一般说来，在塑性加工过程中，有 $P \cdot v_D = P^* \dot{u}_0$。于是，由式（a）可得上限法的基本方程为

$$P^*\dot{u}_0 = \int_V \overline{\sigma\dot{\varepsilon}}\,\mathrm{d}V + \sum_i \int_{S_{ci}} \tau_f |v_R^*| \mathrm{d}\sigma + \sum_j \int_{S_{Dj}} K|v_t^*| \mathrm{d}S = \int_V \sigma\,\dot{\bar{\varepsilon}}^*\,\mathrm{d}V + \sum_i \int_{S_{ci}} \tau_f |v_R^*| \mathrm{d}\sigma + \sum_i \int_{S_{Dj}} K|v_t^*| \mathrm{d}S \quad (2.7\text{-}15)$$

式中，σ 是材料的真实应力。显然，由上式求得的近似载荷 P^* 是真实载荷 P 的一个上界。

3.6 应用实例

用上限法求解金属塑性成形问题，关键是要找出一个动可容速度场。动可容速度场有两种模式：①离散的速度场，即刚性块滑移模式；②连续的或分块连续的速度场。关于刚性块滑移模式，一方面它只适用于平面应变问题，并且在这一模式下，物体的应变速率和塑性变形功率处处为零，对于金属塑性成形来说过于简单，故本章不予介绍，只介绍用连续速度场求解金属塑性加工问题的两个实例。

例1 平行砧板间的圆柱体热镦（见图2.7-10），上模的速率为 $\dot{u}_0$，摩擦应力 $\tau_f = \mu'\sigma$，σ 是材料的真实应力，试用上限法求单位流动压力 p。

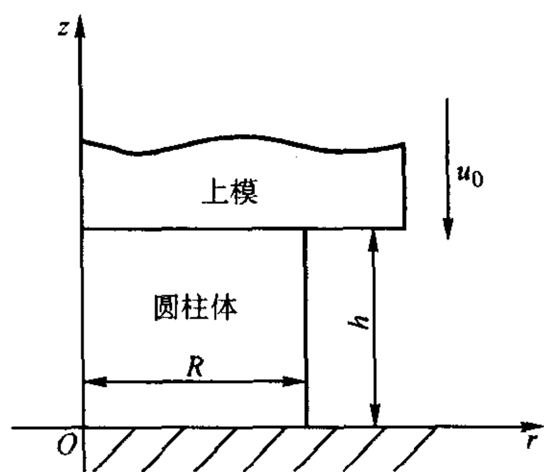


图 2.7-10 圆柱体热镦

解 设圆柱体的一种连续的速度场为

$$\left.\begin{aligned} u_r &= \frac{r}{2h}u_0 \\ u_z &= -\frac{z}{h}u_0 \end{aligned}\right\} \quad (a)$$

由此可知，当 $r=0$ 时，$u_r=0$；$z=0$ 时，$u_z=0$；$z=h$ 时，$\dot{u}_z = -\dot{u}_0$。满足速度边界条件。

由式（a）可得

$$\dot{\varepsilon} = \frac{\partial u_r}{\partial r} = \frac{\dot{u}_0}{2h}, \dot{\varepsilon}_\theta = \frac{\dot{u}_r}{r} = \frac{\dot{u}_0}{2h}, \dot{\varepsilon}_z = \frac{\partial \dot{u}_z}{\partial z} = -\frac{\dot{u}_0}{h}$$

及
$$\dot{\varepsilon}_{rz} = \frac{1}{2}\left(\frac{\partial \dot{u}_r}{\partial z} + \frac{\partial \dot{u}_z}{\partial r}\right) = 0$$

于是
$$\dot{\varepsilon}_r + \dot{\varepsilon}_\theta + \dot{\varepsilon}_z = 0$$

因此，式（a）是一种运动学容许的速度场，且有

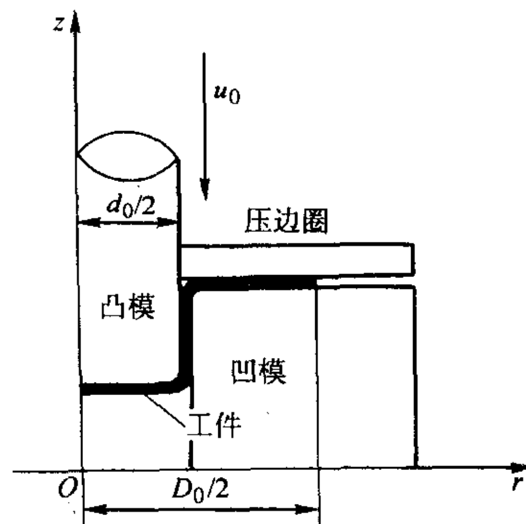
$$\dot{\bar{\varepsilon}} = \sqrt{\frac{2}{3}(\dot{\varepsilon}_r^2 + \dot{\varepsilon}_\theta^2 + \dot{\varepsilon}_z^2 + 2\dot{\varepsilon}_{rz}^2)} = \frac{u_0}{h} \quad (b)$$

由此，列出上限法基本方程为

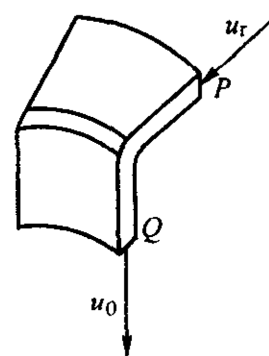
$$p\pi R^2 \dot{u}_0 = \int_V \sigma\dot{\bar{\varepsilon}}\,\mathrm{d}V + 2\int_{S_c} \tau_f |v_R| \mathrm{d}\sigma = \sigma\frac{\dot{u}_0}{h}(\pi R^2 h) + 2\int_0^R \mu'\sigma\frac{r}{2h}\dot{u}_0 2\pi r\mathrm{d}r$$

得
$$p = \sigma\left(1 + \frac{2\mu' R}{3h}\right)$$

例2 设在圆板坯拉深为圆筒件的过程中（参看图2.7-11），板厚 t 保持不变，材料的真实应力为 σ。设冲模下行的速率为 $\dot{u}_0$，不计接触面上的摩擦，且忽略凹模口的弯曲效应，试用上限法求拉深时的单位流动压力 p。



(a) 圆板拉深为圆筒



(b) 体积不可压缩示意图

图 2.7-11 圆筒件拉深

解 在凸缘处任取一点 P，径向速度记为 $\dot{u}_r$；又在侧壁上任取一点 Q，则有 $\dot{u}_{zQ} = \dot{u}_0$。于是，在时间增量 $\mathrm{d}t$ 内由体积不可压缩条件可得

$$\pi d_0 t\dot{u}_0 \mathrm{d}t = 2\pi rt(-\dot{u}_r)\mathrm{d}t$$

得
$$\dot{u}_r = -\frac{d_0}{2r}\dot{u}_0$$

由于板厚不变，可设 $\dot{u}_z = 0$。由此，凸缘区域的一种动可容速度场为

$$\begin{aligned} \dot{u}_r &= -\frac{d_0}{2r}\dot{u}_0 \\ \dot{u}_z &= 0 \end{aligned} \quad (a)$$

于是有

$$\dot{\varepsilon}_r = \frac{\partial \dot{u}_r}{\partial r} = \frac{d_0}{2r^2}\dot{u}_0$$

$$\dot{\varepsilon}_\theta = \frac{\dot{u}_r}{r} = -\frac{d_0}{2r^2}\dot{u}_0$$

及
$$\dot{\varepsilon}_{rz} = \frac{1}{2}\left(\frac{\partial \dot{u}_r}{\partial z} + \frac{\partial \dot{u}_z}{\partial r}\right) = 0$$

由此可得

$$\dot{\varepsilon}_r + \dot{\varepsilon}_\theta + \dot{\varepsilon}_z = 0$$

及
$$\dot{\bar{\varepsilon}} = \sqrt{\frac{2}{3}(\dot{\varepsilon}_r^2 + \dot{\varepsilon}_\theta^2)} = \frac{2}{\sqrt{3}}\frac{\dot{u}_0}{r^2} \quad (b)$$

于是，上限法基本方程为

$$p\pi d_0 t\dot{u}_0 = \int_V \sigma\dot{\bar{\varepsilon}}\,\mathrm{d}V = \int_{d_0/2}^{D_0/2} \sigma\frac{d_0\dot{u}_0}{\sqrt{3}r^2}2\pi r\mathrm{d}rt$$

得
$$p = \frac{2}{\sqrt{3}}\bar{\sigma}\ln\frac{D_0}{d_0}$$

式中，$\bar{\sigma}$ 是由积分中值定理得到的 σ 在凸缘上的一个中值。

将本节例2的结果与用滑移线法得出的相应结果（见2.6节例2）相比较，可知若令 $K = \bar{\sigma}/\sqrt{3}$，则二者是完全一致的。

编写：李赞（华中科技大学）

参考文献

1　罗惕乾主编．流体力学．北京：机械工业出版社，1999

2　吴树森主编．材料加工冶金传输原理．北京：机械工业出版社，2001

3　鲁德洋主编．冶金传输基础．西安：西北工业大学出版社，1991

4　吴树森．金属基复合材料凝固界面的颗粒行为的研究进展．材料导报，1998，12（5）：1～5

5　Wu S S，Nakae H，Kanno T，You Y．Zone-melted unidirectional solidification of particulate dispersed composites．Journal of Materials Science，2001，36：225～229

6　杨世铭编．传热学．北京：高等教育出版社，1994

7　Kou S．Transport Phenomena and Materials Processing．New York：JOHN WILEY & SONS，INC.，1996

8　[日] 中江秀雄．铸造工学．东京：产业图书，1995

9　吴长春主编．冶金热力学．北京：机械工业出版社，1993

10　吴树森，中江秀雄．铝基复合材料中颗粒在凝固界面的行为．金属学报，1998，34（9）：939～944

11　Wu S S，Nakae H．Nucleation effect of alumina in Al-Si/Al_2O_3 composites．Journal of Materials Science letters，1999，18：321～323

12　董若璟主编．铸造合金熔炼原理．北京：机械工业出版社，1991

13　戴永年，杨斌编著．有色金属材料的真空冶金．北京：冶金工业出版社，2000

14　李尚健主编．金属塑性成形过程模拟．北京：机械工业出版社，1999

15　边秀房等．铸造金属遗传学．济南：山东科学技术出版社，1999

16　胡汉起．金属凝固原理．第二版．北京：机械工业出版社，2000

17　常国威等．金属凝固过程中的晶体生长与控制．北京：冶金工业出版社，2002

18　沈定钊等．铁铸冶金．北京：冶金工业出版社，1995

19　戴干策，任德呈，范自晖编著．化学工程基础——流体流动、传热及传质．北京：中国石化出版社，1991

20　陈卓如，金朝铭，王成敏，包纲编．工程流体力学．北京：高等教育出版社，1992

21　刘庄，吴肇基，吴景之，张毅著．热处理过程的数值模拟．北京：科学出版社，1996

22　李国琛，M．耶纳著．塑性大应变微结构力学．第二版．北京：科学出版社，1998

23　[日] 富田佳宏．数值弹塑性力学．東京：養賢堂，1990

24　Tvergaard V，Niordson C. Nonlocal plasticity effects on interaction of different size voids. International Journal of Plasticity，2004，20：107～120

25　俞汉清，陈金德．金属塑性成形原理．北京：机械工业出版社，1999

26　姚若浩．金属压力加工中的摩擦与润滑．北京：冶金工业出版社，1990

27　陈森灿，叶庆荣．金属塑性加工原理．北京：清华大学出版社，1992

28　肖景荣，姜奎华．冲压工艺学．北京：机械工业出版社，1990

29　林治平．上限法在塑性加工工艺中的应用．北京：中国铁道出版社，1991

第 3 篇

材料成形数值模拟

主 编 柳玉起

编 写 王新云 周华民

周建新 柳玉起

审 稿 中国材料工程大典编委会

中国机械工程学会
中国材料研究学会
中国材料工程大典编委会

第1章　有限差分法

有限差分法是数值求解微分问题的一种重要工具，很早就有人在这方面作了一些基础性的工作。到了1910年，L. F. Richardson在一篇论文中论述了Laplace方程、重调和方程等的迭代解法，为偏微分方程的数值分析奠定了基础。但是在电子计算机问世前，研究重点在于确定有限差分解的存在性和收敛性。这些工作成了后来实际应用有限差分法的指南。20世纪40年代后半期出现了电子计算机，有限差分法得到迅速的发展，在很多领域（如传热分析、流动分析、扩散分析等）取得了显著的成就，对国民经济及人类生活产生了重要影响，积极地推动了社会的进步。

有限差分法在材料成形领域的应用较为普遍，与有限元法一起成为材料成形计算机模拟技术的主要的两种数值分析方法。目前材料加工中的传热分析（如铸造成形过程的传热凝固、塑性成形中的传热、焊接成形中的热量传递等）、流动分析（如铸件充型过程，焊接熔池的产生、移动，激光熔覆中的动量传递等），都可以以有限差分方式进行模拟分析。特别是在流动场分析方面，与有限元相比，有限差分法有独特的优势，因此目前进行流体力学数值分析，绝大多数都是基于有限差分法。另外，一向被认为是有限差分法的弱项——应力分析，目前也取得了长足进步。一些基于差分法的材料加工领域的应力分析软件纷纷推出，从而使得流动、传热、应力统一于差分方式下。

可以预见，随着计算机技术的飞速发展，有限差分法将得到更为广泛的应用，可以为材料成形过程提供全面的、有效的指导。

本章主要讲述有限差分的一些基本知识，包括差分原理及逼近误差，差分方程，截断误差和相容性，收敛性与稳定性以及Lax等价定理等，这些仅仅是有限差分的入门知识，为后续章节的研究奠定基础。

1　差分原理及逼近误差

1.1　差分原理

设有 x 的解析函数 $y=f(x)$，从微分学知道函数 y 对 x 的导数为

$$\frac{dy}{dx}=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x} \qquad (3.1\text{-}1)$$

dy、dx 分别是函数及自变量的微分，$\frac{dy}{dx}$ 是函数对自变量的导数，又称微商。相应地，上式中的 Δy、Δx 分别称为函数及自变量的差分，$\frac{\Delta y}{x}$ 为函数对自变量的差商。

在导数的定义中 Δx 是以任意方式趋近于零的，因而 Δx 是可正可负的。在差分方法中，Δx 总是取某一小的正数。这样一来，与微分对应的差分可以有3种形式：

向前差分　$\Delta y=f(x+\Delta x)-f(x) \qquad (3.1\text{-}2)$

向后差分　$\Delta y=f(x)-f(x-\Delta x) \qquad (3.1\text{-}3)$

中心差分　$\Delta y=f(x+\frac{1}{2}\Delta x)-f(x-\frac{1}{2}\Delta x) \qquad (3.1\text{-}4)$

上面谈的是一阶导数，对应的称为一阶差分。对一阶差分再作一阶差分，所得到的称为二阶差分，记为 $\Delta^2 y$。以向前差分为例，有

$$\begin{aligned}\Delta^2 y&=\Delta(\Delta y)\\&=\Delta[f(x+\Delta x)-f(x)]\\&=\Delta f(x+\Delta x)-\Delta f(x)\\&=[f(x+2\Delta x)-f(x+\Delta x)]-[f(x+\Delta x)-f(x)]\\&=f(x+2\Delta x)-2f(x+\Delta x)+f(x)\end{aligned} \qquad (3.1\text{-}5)$$

依此类推，任何阶差分都可由其低一阶的差分再作一阶差分得到。例如 n 阶前差分为

$$\begin{aligned}\Delta^n y&=\Delta(\Delta^{n-1}y)\\&=\Delta[\Delta(\Delta^{n-2}y)]\\&=\Delta\{\Delta\Delta[\Delta(\Delta y)]\}\\&=\Delta\{\Delta\Delta[\Delta(f(x+\Delta x)-f(x)]\}\end{aligned} \qquad (3.1\text{-}6)$$

n 阶的向后差分、中心差分的型式类似。

函数的差分与自变量的差分之比，即为函数对自变量的差商。如一阶向前差商为

$$\frac{\Delta y}{\Delta x}=\frac{f(x+\Delta x)-f(x)}{\Delta x} \qquad (3.1\text{-}7)$$

一阶向后差商为

$$\frac{\Delta y}{\Delta x}=\frac{f(x)-f(x-\Delta x)}{\Delta x} \qquad (3.1\text{-}8)$$

一阶中心差商为

$$\frac{\Delta y}{\Delta x}=\frac{f(x+\frac{1}{2}\Delta x)-f(x-\frac{1}{2}\Delta x)}{\Delta x} \qquad (3.1\text{-}9)$$

或

$$\frac{\Delta y}{\Delta x}=\frac{f(x+\Delta x)-f(x-\Delta x)}{2\Delta x} \qquad (3.1\text{-}10)$$

二阶差商多取中心式，即

$$\frac{\Delta^2 y}{\Delta x^2}=\frac{f(x+\Delta x)-2f(x)+f(x-\Delta x)}{(\Delta x)^2} \qquad (3.1\text{-}11)$$

当然，在某些情况下也可取向前或向后的二阶差商。

以上是一元函数的差分与差商。多元函数 $f(x, y, \cdots)$ 的差分与差商也可以类推。如一阶向前差商为

$$\frac{\Delta f}{\Delta x}=\frac{f(x+\Delta x,y,\Lambda)-f(x,y,\Lambda)}{\Delta x} \qquad (3.1\text{-}12)$$

$$\frac{\Delta f}{\Delta y}=\frac{f(x,y+\Delta y,\Lambda)-f(x,y,\Lambda)}{\Delta y} \qquad (3.1\text{-}13)$$

1.2　逼近误差

由导数（微商）和差商的定义知道，当自变量的差分（增量）趋近于零时，就可由差商得到导数。因此在数值计算中常用差商近似代替导数。差商与导数之间的误差表明差商逼近导数的程度，称为逼近误差。由函数的Taylor展开，可以得到逼近误差相对于自变量差分（增量）的量级，称为用差商代替导数的精度，简称为差商的精度。

现将函数 $f(x+\Delta x)$ 在 x 的 Δx 邻域作Taylor展开：

$$f(x+\Delta x)=f(x)+\Delta x f'(x)+\frac{(\Delta x)^2}{2!}f''(x)+\frac{(\Delta x)^3}{3!}f'''(x)+\frac{(\Delta x)^4}{4!}f^{\mathrm{IV}}(x)+O((\Delta x)^5) \qquad (3.1\text{-}14)$$

$$\begin{aligned}\frac{f(x+\Delta x)-f(x)}{\Delta x}&=f'(x)+\frac{f''(x)}{2!}\Delta x\frac{f'''(x)}{3!}(\Delta x)^2+\frac{f^{\mathrm{IV}}(x)}{4!}(\Delta x)^3+O((\Delta x)^4)\\&=f'(x)+O(\Delta x)\end{aligned} \qquad (3.1\text{-}15)$$

这里符号 O（　）表示与（　）中的量有相同量级的量。上式表明一阶向前差商的逼近误差与自变量的增量同量级。我们把 $O(\Delta x^n)$ 中 Δx 的指数 n 作为精度的阶数。这里 $n=1$，故一阶向前差商具有一阶精度。由于 Δx 是个小量，因此阶数越大精度越高。

又

$$f(x-\Delta x)=f(x)-\Delta x f'(x)+\frac{(\Delta x)^2}{2!}f''(x)-\frac{(\Delta x)^3}{3!}f'''(x)+\frac{(\Delta x)^4}{4!}f^{\mathrm{IV}}(x)+O((\Delta x)^5) \quad (3.1\text{-}16)$$

$$\frac{f(x)-f(x-\Delta x)}{\Delta x}=f'(x)+O(\Delta x)$$

一阶向后差商也具有一阶精度。

将 $f(x+\Delta x)$ 与 $f(x-\Delta x)$ 的 Taylor 展开式相减可得

$$\frac{f(x+\Delta x)-f(x-\Delta x)}{2\Delta x}=f'(x)+O((\Delta x)^2) \quad (3.1\text{-}17)$$

可见一阶中心差商具有二阶精度。

将 $f(x+\Delta x)$ 与 $f(x-\Delta x)$ 的 Taylor 展开式相加可得

$$\frac{f(x+\Delta x)-2f(x)+f(x-\Delta x)}{\Delta x^2}=f''(x)+O((\Delta x)^2) \quad (3.1\text{-}18)$$

这说明二阶中心差商的精度也为二阶。

在掌握了用 Taylor 展开分析差商精度的方法后，再回过来谈一谈函数差分和差商的定义。由于差分和差商是微分和导数的近似表达式，所以不必局限于前面的定义，而可予以扩充。设有函数 $f(x)$，自变量 x 的增量为 Δx，若取

$$x=x_i+j\Delta x,\ j=0,\ \pm1,\ \pm2,\ \Lambda \quad (3.1\text{-}19)$$

对应的函数值为 $f(x_i+j\Delta x)$，则 $f(x)$ 在 x_i 处的 n 阶差分可表达为

$$\Delta^n f(x_i)=\sum_{j=-J_1}^{J_2}c_j f(x_i+j\Delta x) \quad (3.1\text{-}20)$$

$$c_j=\frac{n!\,a_j}{\sum\limits_{j=-J_1}^{J_2}a_j j^n} \quad (3.1\text{-}21)$$

式中，c_j 为给定系数；J_1 和 J_2 是两个正整数。当 $J_1=0$ 时，称为向前差分；当 $J_2=0$ 时，称为向后差分；当 $J_1=J_2$ 且 $|c_j|=|c_{-j}|$ 时，称为中心差分。函数的 n 阶差分与自变量的 n 阶差分之比为 n 阶差商，可用 Taylor 展开分析其逼近误差 $O(\Delta x^m)$。显然，$m\leqslant0$ 的差商及其对应的差分是不恰当的。当且 a_j 为表 3.1-1～表 3.1-6 中所列的数值时，可得 $m>0$。其中表 3.1-1 和表 3.1-2 的 $m=1$，即此两表对应差商的精度是一阶的；表 3.1-3～表 3.1-5 的 $m=2$，即这些表对应差商的精度是二阶的；表 3.1-6 的 $m=4$，即此表对应差商的精度是四阶的。从这些表可以看出，一般地说，随着差分阶数的增大和对应差商精度的提高，差分表达式所包含的项数将增多。

以上的差分是用 $f(x_i\pm j\Delta x)$ 求得的，这表示是以等距离 Δx 向前、向后进行差分的。在有些情况下要求自变量的增量本身是变化的，如图 3.1-1 中的 Δx_{i-2}、Δx_{i-1}、Δx_i 和 Δx_{i+1}，是不相等的，相应的差分和差商就是不等距的。

表 3.1-1 a_j 值 ($m=1$)

n	j = 0	1	2	3	4
	a_j				
1	-1	1			
2	1	-2	1		
3	-1	3	-3	1	
4	1	-4	6	-4	1

表 3.1-2 a_j 值 ($m=1$)

n	j = -4	-3	-2	-1	0
	a_j				
1				-1	1
2			1	-2	1
3		-1	3	-3	1
4	1	-4	6	-4	1

表 3.1-3 a_j 值 ($m=2$)

n	j = 0	1	2	3	4	5
	a_j					
1	-3	4	-1			
2	2	-5	4	-1		
3	-5	18	-24	14	-3	
4	3	-14	26	-24	11	-2

表 3.1-4 a_j 值 ($m=2$)

n	j = -5	-4	-3	-2	-1	0
	a_j					
1				1	-4	3
2			-1	4	-5	2
3		3	-14	24	-18	5
4	-2	11	-24	26	-14	3

表 3.1-5 a_j 值 ($m=2$)

n	j = -2	-1	0	1	2
	a_j				
1		-1	0	1	
2		1	-2	1	
3	-1	2	0	-2	1
4	1	-4	6	-4	1

表 3.1-6 a_j 值 ($m=4$)

n	j = -3	-2	-1	0	1	2	3
	a_j						
1		1	-8	0	8	-1	
2		-1	16	-30	16	-1	
3	1	-8	13	0	-13	8	-1
4	-1	12	-39	56	-39	12	-1

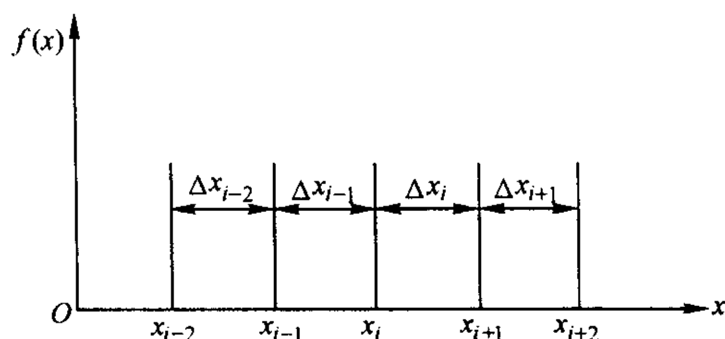


图 3.1-1 变距离差分

下面列出一些不等距的差商供参考：

一阶向后差商 $\dfrac{f(x_i)-f(x_i-\Delta x_{i-1})}{\Delta x_{i-1}}$ (3.1-22)

一阶中心差商 $\dfrac{f(x_i+\Delta x_i)-f(x_i-\Delta x_{i-1})}{\Delta x_i+\Delta x_{i-1}}$ (3.1-23)

二阶向后差商

$$2[f(x_i)\Delta x_{i-2}-f(x_i-\Delta x_{i-1})(\Delta x_{i-2}+\Delta x_{i-1})+f(x_i-\Delta x_{i-1}-\Delta x_{i-2})\Delta x_{i-1}]/[\Delta x_{i-2}\Delta x_{i-1}(\Delta x_{i-2}+\Delta x_{i-1})] \quad (3.1\text{-}24)$$

二阶中心差商

$$2[f(x_i+\Delta x_i)\Delta x_{i-1}-f(x_i)(\Delta x_{i-1}+\Delta x_i)+f(x_i-\Delta x_{i-1})\Delta x_i]/[\Delta x_{i-1}\Delta x_i(\Delta x_{i-1}+\Delta x_i)] \quad (3.1\text{-}25)$$

以上都是一阶精度的。二阶精度的差商如下：

一阶向后差商

$$\{f(x_i)[(\Delta x_{i-2}+\Delta x_{i-1})^2-\Delta x_{i-1}^2]-f(x_i-\Delta x_{i-1})(\Delta x_{i-2}+\Delta x_{i-1})^2+f(x_i-\Delta x_i-\Delta x_{i+1})\Delta x_{i-1}^2\}/[\Delta x_{i-2}\Delta x_{i-1}(\Delta x_{i-2}+\Delta x_{i-1})] \quad (3.1\text{-}26)$$

一阶中心差商

$$[f(x_i+\Delta x_i)\Delta x_{i-1}^2+f(x_i)(\Delta x_i^2-\Delta x_{i-1}^2)-f(x_i-\Delta x_{i-1})\Delta x_i^2]/\Delta x_{i-1}\Delta x_i(\Delta x_{i-1}+\Delta x_i) \quad (3.1\text{-}27)$$

基于不等距离（一维）、不等规格（二维、三维）网格的有限差分方法已广泛应用于目前的商品化软件中，一般称为变网格技术。变网格技术可以在保证计算精度的前提下，有效地提高计算速度。

2 差分方程、截断误差和相容性

从上节所述可知，差分相应于微分，差商相应于导数。只不过差分和差商是用有限形式表示的，而微分和导数则是以极限形式表示的。如果将微分方程中的导数用相应的差商近似代替，就可得到有限形式的差分方程。现以对流方程

$$\frac{\partial\zeta}{\partial t}+\alpha\frac{\partial\zeta}{\partial x}=0 \quad (3.1\text{-}28)$$

为例，列出对应的差分方程。

用差商近似代替导数时，首先要选定 Δx 和 Δt，称为步长。然后在 $x-t$ 坐标平面上用平行于坐标的两族直线：

$$x_i=x_0+i\Delta x,\quad i=0,\ 1,\ 2,\ \Lambda \quad (3.1\text{-}29)$$

$$t_n=n\Delta t,\quad n=0,\ 1,\ 2,\ \Lambda \quad (3.1\text{-}30)$$

划分出矩形网格，如图 3.1-2。通常空间步长 Δx 取为相等的，而时间步长 Δt 与 Δx 以及 α 有关，当 Δx 和 α 为常数时，Δt 也取常数。直线 $t=t_n$ 称为第 n 层。网格交叉点称为结点。

网格划定后，就可针对某一结点，例如图 3.1-2 中的结点（x_i，t_n），用差商近似代替导数。现用 $()_i^n$ 表示 () 内函数在（x_i，t_n）点的值（有时括号可省略），则对流方程在（x_i，t_n）点为

$$\left(\frac{\partial\zeta}{\partial t}\right)_i^n+\alpha\left(\frac{\partial\zeta}{\partial x}\right)_i^n=0 \quad (3.1\text{-}31)$$

这里 α 是作常数处理的。若 α 是 x 的函数，则应该用 α_i。

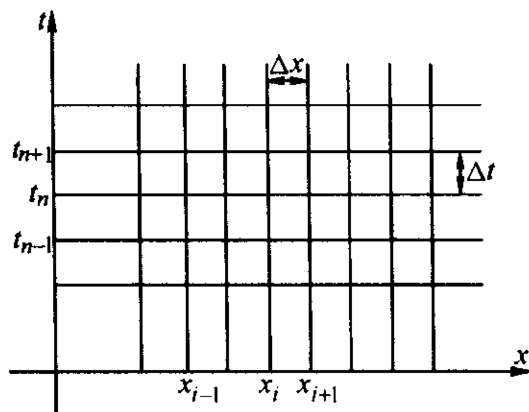


图 3.1-2 差分网格

若时间导数用一阶向前差商近似代替，即

$$\left(\frac{\partial\zeta}{\partial t}\right)_i^n\approx\frac{\zeta_i^{n+1}-\zeta_i^n}{\Delta t} \quad (3.1\text{-}32)$$

空间导数用一阶中心差商近似代替，即

$$\left(\frac{\partial\zeta}{\partial t}\right)_i^n\approx\frac{\zeta_{i+1}^n-\zeta_{i-1}^n}{2\Delta x} \quad (3.1\text{-}33)$$

则在（x_i，t_n）点的对流方程就可近似地写作

$$\frac{\zeta_i^{n+1}-\zeta_i^n}{\Delta t}+\alpha\frac{\zeta_{i+1}^n-\zeta_{i-1}^n}{2\Delta x}=0 \quad (3.1\text{-}34)$$

这就是对应的差分方程。

按照前面关于逼近误差的分析知道，用时间向前差商代替时间导数时的误差为 $O(\Delta t)$，用空间中心差商代替空间导数时的误差为 $O((\Delta x)^2)$，因而对流方程与对应的差分方程之间也存在一个误差，它是

$$R_i^n=O(\Delta t)+O((\Delta x)^2)=O(\Delta t,(\Delta x)^2) \quad (3.1\text{-}35)$$

这也可由 Taylor 展开得到。因为

$$\frac{\zeta(x_i,t_n+\Delta t)-\zeta(x_i,t_n)}{\Delta t}+\alpha\frac{\zeta(x_i+\Delta x,t_n)-\zeta(x_i-\Delta x,t_n)}{2\Delta x}$$

$$=\left(\frac{\partial\zeta}{\partial t}\right)_i^n+\frac{1}{2}\left(\frac{\partial^2\zeta}{\partial t^2}\right)_i^n\Delta t+\Lambda+\alpha\left[\left(\frac{\partial\zeta}{\partial x}\right)_i^n+\frac{1}{3!}\left(\frac{\partial^3\zeta}{\partial t^3}\right)_i^n(\Delta x)^2+\Lambda\right]$$

$$=\left(\frac{\partial\zeta}{\partial t}+\alpha\frac{\partial\zeta}{\partial x}\right)_i^n+O(\Delta t,(\Delta x)^2) \quad (3.1\text{-}36)$$

这种用差分方程近似代替微分方程所引起的误差，称为截断误差。这里误差量级相当于 Δt 的一次式、Δx 的二次式。若已知 Δt 与 Δx 的关系，例如 $\dfrac{\Delta t}{\Delta x}$ = 常数，则 $R_i^n=O(\Delta t,(\Delta x)^2)=O(\Delta t)$，精度为一阶。在一般情况下，则说对 Δt 精度为一阶，对 Δx 精度为二阶。

一个与时间相关的物理问题，应用微分方程表示时，还必须给定初始条件，从而形成一个完整的初值问题。对流方程的初值问题为

$$\begin{cases}\dfrac{\partial\zeta}{\partial t}+\alpha\dfrac{\partial\zeta}{\partial x}=0\\ \zeta(x,0)=\bar{\zeta}(x)\end{cases} \quad (3.1\text{-}37)$$

这里 $\bar{\zeta}(x)$ 为某已知函数。同样，差分方程也必须有初始条件：

$$\begin{cases}\dfrac{\zeta_i^{n+1}-\zeta_i^n}{\Delta t}+\alpha\dfrac{\zeta_{i+1}^n-\zeta_{i-1}^n}{2\Delta x}=0\\ \zeta_i^0=\bar{\zeta}(x_i)\end{cases} \quad (3.1\text{-}38)$$

初始条件是一种定解条件。如果是初边值问题，定解条件中还应有适当的边界条件。差分方程和其定解条件一起，称为相应微分方程定解问题的差分格式。

上述初值问题的差分格式可改写为

$$\begin{cases}\zeta_i^{n+1}=\zeta_i^n-\alpha\dfrac{\Delta t}{2\Delta x}(\zeta_{i+1}^n-\zeta_{i-1}^n)\\ \zeta_i^0=\bar{\zeta}(x_i)\end{cases} \quad (3.1\text{-}39)$$

我们称它为 FTCS 格式。

从 FTCS 格式可见，若已知第 n 层上（x_{i-1}，t_n）、（x_i，

t_n）和（x_{i+1}，t_n）点处函数ζ的值，立即可算出第$n+1$层上（x_i，t_{n+1}）点处函数ζ的值。由于在第0层（初始层）函数ζ的值是已给定的，故可逐层计算。为了直观起见，可用图3.1-3a表示FTCS格式的计算方式。差分方程是由图中"⊗"点列出，图中"○"表示计算所涉及的结点。这种图称为格式图。

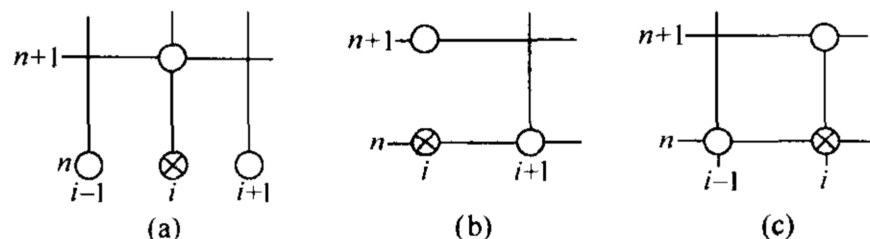


图3.1-3 差分格式

FTCS格式是采用时间向前差分、空间中心差分得来的。若时间和空间都用向前差分，则得

$$\begin{cases}\dfrac{\zeta_i^{n+1}-\zeta_i^n}{\Delta t}+\alpha\dfrac{\zeta_{i+1}^n-\zeta_i^n}{\Delta x}=0\\ \zeta_i^0=\overline{\zeta}\ (x_i)\end{cases} \tag{3.1-40}$$

或改写成

$$\begin{cases}\zeta_i^{n+1}=\zeta_i^n-\alpha\dfrac{\Delta t}{\Delta x}\ (\zeta_{i+1}^n-\zeta_i^n)\\ \zeta_i^0=\overline{\zeta}\ (x_i)\end{cases} \tag{3.1-41}$$

这是FTFS格式，其格式图如图3.1-3b。

若采用时间向前差分、空间向后差分，则得到FTBS格式：

$$\begin{cases}\dfrac{\zeta_i^{n+1}-\zeta_i^n}{\Delta t}+\alpha\dfrac{\zeta_i^n-\zeta_{i-1}^n}{\Delta x}=0\\ \zeta_i^0=\overline{\zeta}\ (x_i)\end{cases} \tag{3.1-42}$$

或

$$\begin{cases}\zeta_i^{n+1}=\zeta_i^n-\alpha\dfrac{\Delta t}{\Delta x}\ (\zeta_i^n-\zeta_{i-1}^n)\\ \zeta_i^0=\overline{\zeta}\ (x_i)\end{cases} \tag{3.1-43}$$

其格式图如图3.1-3c。

FTCS格式的截断误差为

$$R_i^n=O(\Delta t,(\Delta x)^2) \tag{3.1-44}$$

FTFS和FTBS格式的截断误差为

$$R_i^n=O(\Delta t,\Delta x) \tag{3.1-45}$$

3种格式对Δt都有一阶精度。

一般说来，若微分方程为

$$D(\zeta)=f \tag{3.1-46}$$

式中，D为微分算子，f为已知函数，而对应的差分方程为

$$D_\Delta(\zeta)=f \tag{3.1-47}$$

式中，D_Δ为差分算子，则截断误差为

$$R=D_\Delta(\phi)-D(\phi) \tag{3.1-48}$$

这里ϕ为定义域上某一足够光滑的函数，当然也可以取微分方程的解ζ。

如果当Δx、$\Delta t\to0$时，差分方程的截断误差的某种范数$\|R\|$也趋近于零，即

$$\lim_{\substack{\Delta x\to0\\ \Delta t\to0}}\|R\|=0 \tag{3.1-49}$$

则表明从截断误差的角度来看，此差分方程是能用来逼近微分方程的，通常称这样的差分方程和相应的微分方程相容（一致）。如果当Δx、$\Delta t\to0$时，截断误差的范数不趋于零，则称为不相容（不一致），这样的差分方程不能用来逼近微分方程。

以上只考虑了方程，但从整个问题来看，还应考虑定解条件。若微分问题的定解条件为

$$B(\zeta)=g \tag{3.1-50}$$

式中，B为微分算子，g为已知函数，而对应的差分问题的定解条件为

$$B_\Delta(\zeta)=g \tag{3.1-51}$$

式中，B_Δ为差分算子，则截断误差为

$$r=B_\Delta(\phi)-B(\phi) \tag{3.1-52}$$

只有方程相容，定解条件也相容，即

$$\lim_{\substack{\Delta x\to0\\ \Delta t\to0}}\|R\|=0\ 和\ \lim_{\substack{\Delta x\to0\\ \Delta t\to0}}\|r\|=0 \tag{3.1-53}$$

整个问题才相容。Δx、$\Delta t\to0$的情况有两种：一是各自独立地趋于零，这是无条件相容；另一是Δx与Δt之间在某种关系（譬如要求$\dfrac{\Delta t}{\Delta x}=K$）下同时趋于零，这种情况下的相容为条件相容。

从截断误差的分析知道，FTCS、FTFS和FTBS格式都具有相容性。这3种格式都只涉及两个时间层的量。此外，若知道第n层的ζ，可由一个差分式子直接算出第$n+1$层的ζ，故称这类格式为显式格式。总起来，以上3种格式都属于一阶精度、二层、相容、显式格式。

这3种格式也有不同的特性，如有的不能用来作实际计算，这将在下面介绍稳定性时谈到。

以上介绍中将一点（x_i，t_n）的函数值，如函数ζ在这点的值，有时写为$\zeta(x_i,t_n)$，有时写为ζ_i^n，以后还会遇到这类情况。通常认为，$\zeta(x_i,t_n)$表示连续函数$\zeta(x,t)$在（x_i，t_n）点的值；而ζ_i^n没有"连续"的含义，只是表示某离散点（x_i，t_n）处的ζ值。因此$\zeta(x_i,t_n)$可以作为微分方程的解在（x_i，t_n）点的值，而ζ_i^n则作为差分问题（代数方程）的解。

3 收敛性与稳定性

3.1 收敛性

所谓相容性，是指当自变量的步长趋于零时，差分格式与微分问题的截断误差的范数是否趋于零，从而可看出是否能用此差分格式来逼近微分问题。然而，方程（无论是微分方程或是差分方程）是物理问题的数学表达形式，其目的是为了借助数学的手段来求问题的解。因此，除了必须要求差分格式能逼近微分方程和定解条件（表明这两种数学表达方法在形式上是一致的）外，还进一步要求差分格式的解（精确解）与微分方程定解问题的解（精确解）是一致的（表明这两种数学表达方法的最终结果是一致的）。即当步长趋于零时，要求差分格式的解趋于微分方程定解问题的解。我们称这种是否趋于微分方程定解问题的解的情况为差分格式的收敛性。更明确地说，对差分网格上的任意结点（x_i，t_n），也是微分问题定解区域上的一固定点，设差分格式在此点的解为ζ_i^n，相应的微分问题的解为$\zeta(x_i,t_n)$，二者之差为

$$e_i^n=\zeta_i^n-\zeta(x_i,t_n) \tag{3.1-54}$$

称为离散化误差。如果当Δx、$\Delta t\to0$时，离散化误差的某种范数$\|e\|$趋近于零，即

$$\lim_{\substack{\Delta x\to0\\ \Delta t\to0}}\|e\|=0 \tag{3.1-55}$$

则说明此差分格式是收敛的，即此差分格式的解收敛于相应微分问题的解，否则不收敛。与相容性类似，收敛又分为有条件收敛和无条件收敛。

粗看起来，似乎只要差分格式逼近微分问题（Δx、$\Delta t\to0$）时，$\|R\|\to0$，$\|r\|\to0$，其解就应该一致；也就是说，似乎相容性能保证收敛性。其实并不一定如此。这是因为在分析截断误差时，是以差分格式与微分问题有同一个解$\zeta(x,t)$为基础（或以定解域内某足够光滑的函数ϕ为基

础），并对此函数分别在（x_i，t_n）点的邻域作 Taylor 展开的，其中所有的 ζ、$\partial\zeta/\partial t$、$\partial\zeta/\partial x$ 等都是指同一个函数及其各阶导数，所以最后得到的截断误差 R、r 实质上是当差分问题与微分问题有同一解时两种方程、两种定解条件之间的误差。R、r 并不能真正表示两种方程、两种定解条件之间的误差，因此，相容性不能保证收敛性，不能保证二者解的一致。但若没有相容性就更不能得到二者解的一致，故相容性是收敛性的必要条件，有人称相容性是形式上的逼近。

相容性不一定能保证收敛性，那么对于一定的差分格式，其解能否收敛到相应微分问题的解？答案是差分格式的解收敛于微分问题的解是可能的。至于某给定格式是否收敛，则要按具体问题予以证明。下面以一个差分格式为例，讨论其收敛性。

微分问题

$$\begin{cases}\dfrac{\partial\zeta}{\partial t}+\alpha\dfrac{\partial\zeta}{\partial x}=0\\ \zeta(x,0)=\bar{\zeta}(x)\end{cases}\tag{3.1-56}$$

的 FTBS 格式为

$$\begin{cases}\dfrac{\zeta_i^{n+1}-\zeta_i^n}{\Delta t}+\alpha\dfrac{\zeta_i^n-\zeta_{i-1}^n}{\Delta x}=0\\ \zeta_i^0=\bar{\zeta}(x_i)\end{cases}\tag{3.1-57}$$

在某结点（x_i，t_n）微分问题的解为 $\zeta(x_i,t_n)$，差分格式的解为 ζ_i^n，则离散化误差为

$$e_i^n=\zeta_i^n-\zeta\ (x_i,\ t_n)\tag{3.1-58}$$

按照截断误差的分析知道

$$\frac{\zeta(x_i,t_n+\Delta t)-\zeta(x_i,t_n)}{\Delta t}+\alpha\frac{\zeta(x_i,t_n)-\zeta(x_i-\Delta x,t_n)}{\Delta x}=O(\Delta x,\Delta t)\tag{3.1-59}$$

以 FTBS 格式中的第一个方程减去上式得

$$\frac{e_i^{n+1}-e_i^n}{\Delta t}+\alpha\frac{e_i^n-e_{i-1}^n}{\Delta x}=O(\Delta x,\Delta t)\tag{3.1-60}$$

或写成

$$\begin{aligned}e_i^{n+1}&=e_i^n-\alpha\frac{\Delta t}{\Delta x}(e_i^n-e_{i-1}^n)+\Delta tO(\Delta x,\Delta t)\\&=(1-\alpha\frac{\Delta t}{\Delta x})e_i^n+\alpha\frac{\Delta t}{\Delta x}e_{i-1}^n+\Delta tO(\Delta x,\Delta t)\end{aligned}\tag{3.1-61}$$

若条件 $\alpha\geqslant0$ 和 $\alpha\dfrac{\Delta t}{\Delta x}\leqslant1$ 成立，即 $0\leqslant\alpha\dfrac{\Delta t}{\Delta x}\leqslant1$，则

$$\begin{aligned}|e_i^{n+1}|&\leqslant(1-\alpha\frac{\Delta t}{\Delta x})|e_i^n|+\alpha\frac{\Delta t}{\Delta x}|e_{i-1}^n|+\Delta tO(\Delta x,\Delta t)\\&\leqslant(1-\alpha\frac{\Delta t}{\Delta x})\max_i|e_i^n|+\alpha\frac{\Delta t}{\Delta x}\max_i|e_i^n|+\Delta tO(\Delta x,\Delta t)\end{aligned}\tag{3.1-62}$$

式中 $\max\limits_i|e_i^n|$ 表示在第 n 层所有结点上 $|e|$ 的最大值。

由上式知，对一切 i 有

$$|e_i^{n+1}|\leqslant\max_i|e_i^n|+\Delta tO(\Delta x,\Delta t)\tag{3.1-63}$$

故有

$$\max_i|e_i^{n+1}|\leqslant\max_i|e_i^n|+\Delta tO(\Delta x,\Delta t)\tag{3.1-64}$$

于是

$$\begin{aligned}&\max_i|e_i^1|\leqslant\max_i|e_i^0|+\Delta tO(\Delta x,\Delta t)\\&\max_i|e_i^2|\leqslant\max_i|e_i^1|+\Delta tO(\Delta x,\Delta t)\\&\max_i|e_i^n|\leqslant\max_i|e_i^{n-1}|+\Delta tO(\Delta x,\Delta t)\end{aligned}\tag{3.1-65}$$

综合得

$$\max_i|e_i^n|\leqslant\max_i|e_i^0|+n\Delta tO(\Delta x,\Delta t)\tag{3.1-66}$$

由于初始条件给定函数 ζ 的初值，初始离散化误差 $e_i^0=0$，并且 $n\Delta t=t_n$ 是一有限量，因而

$$\max_i|\ e_i^n\ |\leqslant O\ (\Delta x,\ \Delta t)\tag{3.1-67}$$

可见本问题 FTBS 格式的离散化误差与截断误差具有相同的量级。最后得到

$$\lim_{\substack{\Delta x\to0\\ \Delta t\to0}}\ (\max_i|e_i^n|\ )\ =0\tag{3.1-68}$$

这样就证明了，当 $0\leqslant\alpha\dfrac{\Delta t}{\Delta x}\leqslant1$ 时，本问题的 FTBS 格式收敛。这种离散化误差的最大绝对值趋于零的收敛情况称为一致收敛。

此例介绍了一种证明差分格式收敛的方法，同时表明了相容性与收敛性的关系：相容性是收敛性的必要条件，但不一定是充分条件，还可能要求其他条件，如本例就是要求 $0\leqslant\alpha\dfrac{\Delta t}{\Delta x}\leqslant1$。

3.2 稳定性

首先介绍一下差分格式的依赖区间、决定区域和影响区域。还是以初值问题

$$\begin{cases}\dfrac{\partial\zeta}{\partial t}+\alpha\dfrac{\partial\zeta}{\partial x}=0\\ \zeta(x,0)=\bar{\zeta}(x)\end{cases}\tag{3.1-69}$$

为例。先看 FTCS 格式，如图 3.1-4a，欲计算第二层 p 点的函数值，必先知道第一层上 a、b、c 这 3 点的函数值，故说 p 点的解依赖于 a、b、c 这 3 点的解。而 a 点的解又依赖于第 0 层（初值线）上 A、d、e 的初值，b 点的解依赖于 d、e、f 的初值，c 点的解依赖于 e、f、B 的初值。因此 p 点的解依赖于初值线 AB 段上所有结点的初值，故称 AB 段上所有结点为 p 点的依赖区间。又三角形 pAB 区域内任一结点的依赖区间都包含在 AB 之内，即该区域内任一结点上的解都由 AB 段上某些结点的初值所决定，而与 AB 以外结点的初值无关，故称此三角形区域为 AB 区间所决定的区域。这里为方便起见，是以第二层的 p 点为例的，事实上对任意层的任一结点，都在初始层上有一对应的依赖区间，而初始层的任一区间都有一对应的决定区域。

FTFS 格式和 FTBS 格式的依赖区间分别为图 3.1-4b 和图 3.1-4c 中的 AB 线段上的全部结点，图中阴影部分为 AB 所决定的区域。

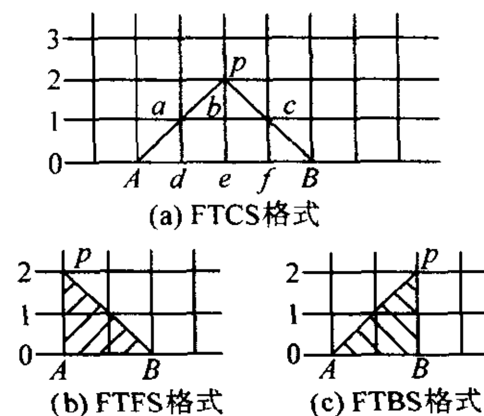


图 3.1-4 差分格式的依赖区间

随着时间的推移，一点函数值将影响以后某些结点的解。如图 3.1-5，设 p 为第 n 层的某结点，当用 FTCS 格式计算第 $n+1$ 层上的结点值时，a、b、c 这 3 点的解必须用到 p 点的函数值，在第 $n+2$ 层上则有更多点的解受 p 点函数值的影响。所有受 p 点函数值影响的结点总和为 p 点的影响区域，如图 3.1-5 中阴影所示区域。

由上可知，同一微分问题，当采用不同差分格式时，其依赖区间、决定区域和影响区域可以是不一样的。依赖区间、决定区域和影响区域是由差分格式本身的构造所决定的，并与步长比 $\Delta t/\Delta x$ 有关。

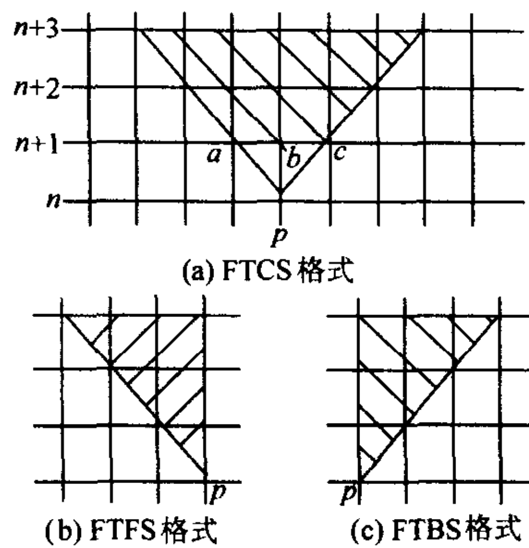


图3.1-5　差分格式的影响区域

例如微分问题

$$\begin{cases} \dfrac{\partial \zeta}{\partial t}+\dfrac{\partial \zeta}{\partial x}=0 \\ \zeta(x,0)=0 \end{cases} \tag{3.1-70}$$

其解为零，即$\zeta(x,t)=0$。若用FTBS格式计算，且计算中不产生任何误差，则结果也是零，即

$$\left.\begin{aligned} \zeta_i^n=0, n=0,1,2,\Lambda \\ i=0,\pm1,\pm2,\Lambda \end{aligned}\right\} \tag{3.1-71}$$

假设在第k层上的第j点，由于计算误差得到$\zeta_j^k=\varepsilon$。不妨设$k=0$，$j=0$，$\varepsilon=1$，即相当于FTBS格式写成

$$\begin{cases} \zeta_i^{n+1}=\zeta_i^n-\dfrac{\Delta t}{\Delta x}(\zeta_i^n-\zeta_{i-1}^n) \\ \zeta_0^0=1 \\ \zeta_i^0=0,\quad i\neq 0 \end{cases} \tag{3.1-72}$$

现分别取$\dfrac{\Delta t}{\Delta x}=1/2$，1和2，列表计算如下：

1) $\dfrac{\Delta t}{\Delta x}=\dfrac{1}{2}$

4	0	0	0	0	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$
3	0	0	0	0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	0
2	0	0	0	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	0	0
1	0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
0	0	0	0	0	1	0	0	0	0
n / ζ_i^n / i	-4	-3	-2	-1	0	1	2	3	4

2) $\dfrac{\Delta t}{\Delta x}=1$

4	0	0	0	0	0	0	0	0	1
3	0	0	0	0	0	0	0	1	0
2	0	0	0	0	0	0	1	0	0
1	0	0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0	0	0
n / ζ_i^n / i	-4	-3	-2	-1	0	1	2	3	4

3) $\dfrac{\Delta t}{\Delta x}=2$

4	0	0	0	0	1	-8	24	-32	16
3	0	0	0	0	-1	6	-12	8	0
2	0	0	0	0	1	-4	4	0	0
1	0	0	0	0	-1	2	0	0	0
0	0	0	0	0	1	0	0	0	0
n / ζ_i^n / i	-4	-3	-2	-1	0	1	2	3	4

这个例子一方面显示了该格式的影响区域，另一方面还显示了当$\Delta t/\Delta x$值不同时，计算误差所产生的影响在数值上有很大的不同。当$\Delta t/\Delta x\leqslant 1$时，所产生的影响在数值上不再扩大；当$\Delta t/\Delta x>1$时，所产生的影响在数值上将越来越大。数值上的差别引出了质的不同，因而出现了稳定性问题。

差分格式的数值稳定性，早在1928年就由R. Courant、K. O. Friedrichs和H. Lewy等人发现，并提出了关于双曲型方程差分格式稳定性的必要条件（简称CFL条件）。此后在这方面作了不少研究工作。1950年公开发表了von Neumann提出的稳定性分析法，这是现在比较广泛地用来确定稳定性准则的一种分析方法。

在有限差分法的具体运算中，计算误差总是不可避免的，如舍入误差，以及这种误差的传播、积累。然而人们通过大量的实践和理论分析发现，同一问题的各种差分格式在某一定条件下，对误差的敏感程度不一样。例如某种格式在某一定条件下，若计算中某处产生了误差，则这个误差将对以后的计算产生影响。如果这一误差对以后的影响越来越小，或是这个影响保持在某个限度以内，像上面例子中$\Delta t/\Delta x\leqslant 1$的情况，那么就称这个差分格式在给定的条件下稳定，这个条件就是它的稳定准则。如果误差的影响随着n的增大越来越大，像上面$\Delta t/\Delta x>1$的情况，使计算的结果随着n的增大越来越偏离差分格式的精确解，而毫无实用价值，那么这种情况就是不稳定的。实际表明，有些格式在一定条件下稳定；有些格式在任何情况下都不稳定，称为完全不稳定。有些格式是无条件稳定的，称为完全稳定。

值得强调的是，这里所说的某种格式稳定或不稳定，是一种简略的说法。实际上，不可能孤立地研究某种格式，必然是针对某一微分问题来研究某差分格式，所以差分格式的相容性、收敛性、稳定性都是针对给定的微分问题而言的。

现在以适当的数学式子给出稳定性定义。为此将差分解ζ_i^n表示为连续函数$Z(x,t)$，则稳定性的一种定义为

$$\|Z(x,t)\|\leqslant K\|Z(x,0)\| \tag{3.1-73}$$

这里K是某个有限常数，称为Lipschitz常数，不随$\Delta x\to 0$、$\Delta t\to 0$而变。这就是说，当上述不等式成立时，只要差分问题初始值所含的误差为小量时，此后的解与差分问题的精确解的误差也一定为小量。由于计算误差不仅可以来自初值（包括在某一时刻前的任一时刻），还可以来自边界值，而且可以来自右端项，所以也有将稳定性定义为

$$\|Z\|\leqslant K_1\|D_\Delta(Z)\|+K_2\|B_\Delta(Z)\| \tag{3.1-74}$$

式中，D_Δ和B_Δ分别是对应于微分方程和定解条件的差分算子；K_1、K_2分别是对应于$D_\Delta(Z)$、$B_\Delta(Z)$的Lipschitz常数。若取

$$K=\max(K_1,K_2) \tag{3.1-75}$$

则为

$$\|Z\|\leqslant K(\|D_\Delta(Z)\|+\|B_\Delta(Z)\|) \tag{3.1-76}$$

在建立了稳定性概念之后，可以进一步判定格式是否稳定，或在什么条件下稳定，因篇幅关系，这里不再详述。

4 Lax 等价定理

前面讨论了差分问题的相容性、收敛性和稳定性。已经知道，相容性是收敛性的必要条件；还发现，稳定性与收敛性有一定的联系。Lax 等价定理就是阐述相容性、收敛性和稳定性三者之间关系的。

Lax 等价定理：对一个适定的线性微分问题及一个与其相容的差分格式，如果该格式稳定则必收敛，不稳定必不收敛。换言之，若线性微分问题适定，差分格式相容，则稳定性是收敛性的必要和充分条件。这也可表示为

$$\text{稳定性} \underset{\text{差分格式相容}}{\overset{\text{线性微分问题适定}}{\Leftarrow=\Rightarrow}} \text{收敛性}$$

下面对此定理作一些简略的说明。

由于在定解域内有

$$D(\zeta)=f \text{ 及 } D_{\Delta}(Z)=f \qquad (3.1\text{-}77)$$

其中，D 和 D_{Δ} 分别为微分算子和差分算子，是线性的；f 为已知函数；ζ 和 Z 分别为微分解和差分解。两式相减得

$$D_{\Delta}(Z)-D(\zeta)=0 \qquad (3.1\text{-}78)$$

改写成

$$[D_{\Delta}(Z)-D_{\Delta}(\zeta)]+[D_{\Delta}(\zeta)-D(\zeta)]=0 \quad (3.1\text{-}79)$$

因为算子是线性的，故式中第一个 [ ] 内相当于 $D_{\Delta}(Z-\zeta)$；而第二个 [ ] 内就是截断误差 R，所以有

$$D_{\Delta}(Z-\zeta)=-R \qquad (3.1\text{-}80)$$

若定解条件为

$$B(\zeta)=g \text{ 及 } B_{\Delta}(Z)=g \qquad (3.1\text{-}81)$$

其中，B 和 B_{Δ} 分别为微分算子和差分算子，且是线性的；g 为已知函数。按照以上对方程的同样推导法，可导得

$$B_{\Delta}(Z-\zeta)=-r \qquad (3.1\text{-}82)$$

其中 $r=B_{\Delta}(\zeta)-B(\zeta)$ 是截断误差。若差分格式是稳定的，按稳定性的定义，应该有

$$\|Z-\zeta\| \leqslant K[\|D_{\Delta}(Z-\zeta)\|+\|B_{\Delta}(Z-\zeta)\|] \qquad (3.1\text{-}83)$$

将式 (3.1-80)、式 (3.1-82) 代入得

$$\|Z-\zeta\| \leqslant K(\|R\|+\|r\|) \qquad (3.1\text{-}84)$$

当差分格式相容时，可得

$$\lim_{\substack{\Delta x\to 0\\ \Delta t\to 0}} \|Z-\zeta\|=0 \qquad (3.1\text{-}85)$$

从而保证了收敛性。

根据此定理，在线性适定和格式相容的条件下，只要证明了格式是稳定的，则一定收敛；若不稳定，则不收敛。由于收敛性的证明往往比稳定性更难，故人们就可以把注意力集中在稳定性的研究上。

编写：周建新（华中科技大学）

第 2 章　弹性问题有限元基本方法

1　弹性问题有限元方法的一般过程

弹性问题有限元法是有限元应用的基础，是实际工程中应用最广泛、最成功的数值方法之一。它主要有基于最小势能原理或虚功原理的位移法；基于最小余能原理的力法；基于修正余能原理的杂交法；基于 Reissner 变分原理的混合法。在这些方法中位移法最成熟，已广泛应用于机械、航空等各领域的线弹性、几何非线性、材料非线性等问题的分析。下面以线弹性位移法为例说明有限元方法的一般过程。

1.1　离散化过程

有限元法的基本思想是：

1）将连续的求解系统离散为由一组节点相互联在一起的单元组合体；

2）在每个单元内假设近似函数来分片表示系统的求解场函数。

其中，1）反映了所分析系统的有限元网格划分过程。2）通过在单元内假设不同的插值场函数，建立不同的单元模型，适应各种各样的变形模式和受力模式。例如分析汽车车身在一定工况下的应力和位移响应时，可以根据车身的实际结构和受力情况，将车身模型进行适当简化，划分成由杆、梁、板、壳、三维实体等单元的组合体，这是网格划分过程。杆、梁、板、壳、三维实体等单元的变形模式和受力模式各不相同，为了能分别描述它们所具有的特性，单元的插值场函数也是各不相同的。实际上单元内的插值场函数与单元模型是一一对应的。

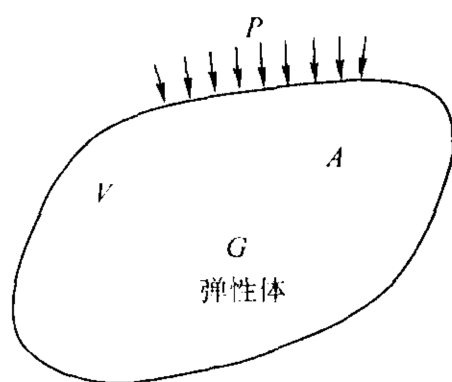


图 3.2-1　弹性体受力情况

如图 3.2-1 所示，所要分析的系统为一个受力弹性体。其中 A 为弹性体的表面积，V 为弹性体的体积，G 为弹性体的体力，P 为作用于弹性体表面的分布力。则弹性体的势能 Π_p 为

$$\Pi_p = W_i - W_e \tag{3.2-1}$$

式中，W_i 为弹性体变形后所具有的内能；W_e 为弹性体所受的外力功。它们的具体表达式分别为

$$W_i = \frac{1}{2}\int_V \varepsilon^T \sigma \mathrm{d}V \tag{3.2-2}$$

$$W_e = \int_A u^T P \mathrm{d}A + \int_V u^T G \mathrm{d}V \tag{3.2-3}$$

式中，ε 为弹性体的应变；σ 为弹性体的应力；u 为弹性体的可容位移。

弹性体处于平衡状态时，其势能最小，即

$$\delta\Pi_p = 0 \tag{3.2-4}$$

$$\delta\Pi_p = \int_V \delta\varepsilon^T \sigma \mathrm{d}V - \int_A \delta u^T P \mathrm{d}A - \int_V \delta u^T G \mathrm{d}V = 0 \tag{3.2-5}$$

式（3.2-5）是最小势能原理，它是位移法有限元方法的数学基础。

按照有限元法的基本思想，对所要分析的系统选择适当的单元模型进行网格划分，对于系统中的每个单元式（3.2-5）同样成立。在此基础上推导单元平衡方程，一般还要知道 3 个关系式，即单元模型的插值关系、单元运动与变形的几何关系以及单元变形与受力的本构关系，这 3 个关系式是位移法推导单元平衡方程时必不可少的。

1）单元插值关系（单元内任意点的位移 u 与单元节点位移 u^e 之间的关系）

$$u = \boldsymbol{N}u^e \tag{3.2-6}$$

式中，$\boldsymbol{N}$ 为单元形函数矩阵，后面将详细介绍各种常见单元的形函数矩阵和具体求法。

2）单元几何关系（单元内任意点的位移 u 与应变 ε 之间的关系）

$$\varepsilon = \boldsymbol{L}u \tag{3.2-7}$$

式中，$\boldsymbol{L}$ 为单元几何微分算子，后面将详细介绍常见运动关系的几何微分算子。

3）单元本构关系（单元内任意点的应变 ε 与应力 σ 之间的关系）

$$\sigma = \boldsymbol{D}^e \varepsilon \tag{3.2-8}$$

式中，$\boldsymbol{D}^e$ 为单元弹性矩阵，后面将详细介绍常见弹性本构矩阵。

将式（3.2-6）、式（3.2-7）和式（3.2-8）代入最小势能原理式（3.2-5）可得

$$\int_v \delta(u^e)^T \boldsymbol{B}^T \boldsymbol{D}^e \boldsymbol{B} u^e \mathrm{d}v - \int_a \delta(u^e)^T \boldsymbol{N}^T P \mathrm{d}a - \int_v \delta(u^e)^T \boldsymbol{N}^T G \mathrm{d}v = 0 \tag{3.2-9}$$

式中，v 为单元体积；a 为单元表面积；$\boldsymbol{B}$ 为应变矩阵。

$$\boldsymbol{B} = \boldsymbol{L}\boldsymbol{N} \tag{3.2-10}$$

由于单元可容位移 u^e 的任意性，得

$$\int_v \boldsymbol{B}^T \boldsymbol{D}^e \boldsymbol{B} u^e \mathrm{d}v - \int_a \boldsymbol{N}^T P \mathrm{d}a - \int_v \boldsymbol{N}^T G \mathrm{d}v = 0 \tag{3.2-11}$$

式（3.2-11）是单元平衡方程，或称为单元刚度方程

$$\boldsymbol{k}u^e = \boldsymbol{f} \tag{3.2-12}$$

式中，$\boldsymbol{k}$ 称为单元刚度矩阵

$$\boldsymbol{k} = \int_v \boldsymbol{B}^T \boldsymbol{D}^e \boldsymbol{B} \mathrm{d}v \tag{3.2-13}$$

$\boldsymbol{f}$ 称为单元载荷向量

$$\boldsymbol{f} = \int_a \boldsymbol{N}^T P \mathrm{d}a + \int_v \boldsymbol{N}^T G \mathrm{d}v \tag{3.2-14}$$

单元刚度矩阵的特性如下。

1）对称性。如果单元弹性矩阵 $\boldsymbol{D}^e$ 是对称的，则 $\boldsymbol{k}$ 也是对称的。

2）奇异性。单元处于平衡状态时，单元平衡方程组的各方程间是线性相关的。

3）主元恒正且对角占优。$\boldsymbol{k}$ 的对角线元素 k_{ii} 始终满足

$$k_{ii} > 0 \tag{3.2-15}$$

并且 k_{ii} 是所在行列中比较大的一个数。

1.2　单元平衡方程组装过程

每个单元的平衡方程求得以后无法分别单独求解，主要原因是，对于单元平衡方程来说，单元间的节点内力也是边界条件，它们在单元应力确定前是未知的。因此，要把所有

的单元平衡方程进行组装，消除单元间的节点内力。而组装后总体平衡方程的边界条件都是实际存在的、已知的。

单元平衡方程组装成总体平衡方程是按有限元节点自由度对应关系进行的，将单元平衡方程左右端的元素累加到总体平衡方程的左右端。下面以二维桁架结构说明组装过程。

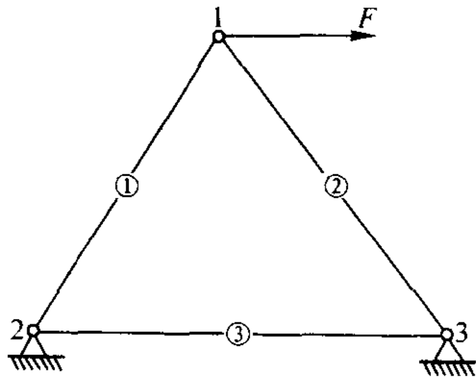


图 3.2-2　简单二维桁架结构的有限元单元划分

如图3.2-2所示平面桁架结构，经过有限元单元划分，1、2、3为节点号，①、②、③为单元号。假设1节点对应的自由度为1和2，2节点对应的自由度为3和4，3节点对应的自由度为5和6。按式（3.2-14）和式（3.2-15）可以求得单元①和②的单元平衡方程

$$\begin{array}{c} \\ 1\\2\\3\\4 \end{array}\begin{array}{c} \begin{matrix}1 & 2 & 3 & 4\end{matrix} \\ \begin{bmatrix} k_{11}^1 & k_{12}^1 & k_{13}^1 & k_{14}^1 \\ k_{21}^1 & k_{22}^1 & k_{23}^1 & k_{24}^1 \\ k_{31}^1 & k_{32}^1 & k_{33}^1 & k_{34}^1 \\ k_{41}^1 & k_{42}^1 & k_{43}^1 & k_{44}^1 \end{bmatrix}\end{array}\begin{Bmatrix} u_1^1 \\ u_2^2 \\ u_3^3 \\ u_4^4 \end{Bmatrix} = \begin{Bmatrix} f_1^1 \\ f_2^1 \\ f_3^1 \\ f_4^1 \end{Bmatrix} \tag{3.2-16}$$

$$\begin{array}{c} \\ 1\\2\\5\\6 \end{array}\begin{array}{c} \begin{matrix}1 & 2 & 5 & 6\end{matrix} \\ \begin{bmatrix} k_{11}^2 & k_{12}^2 & k_{13}^2 & k_{14}^2 \\ k_{21}^2 & k_{22}^2 & k_{23}^2 & k_{24}^2 \\ k_{31}^2 & k_{32}^2 & k_{33}^2 & k_{34}^2 \\ k_{41}^2 & k_{42}^2 & k_{43}^2 & k_{44}^2 \end{bmatrix}\end{array}\begin{Bmatrix} u_1^1 \\ u_2^2 \\ u_3^5 \\ u_4^6 \end{Bmatrix} = \begin{Bmatrix} f_1^2 \\ f_2^2 \\ f_3^2 \\ f_4^2 \end{Bmatrix} \tag{3.2-17}$$

式中，k_{ij}^n表示单刚中每个元素，i 和 j 表示它在单刚中的位置，n 表示单元号，区分不同单元。u_i^m 表示单元的位移自由度，i 表示单元自由度位置，m 表示它在整个系统中的自由度位置。相应地，在 k_{ij}^n矩阵的左侧和上方也标出了它的行列在整个系统中的自由度位置。

组装过程就是按照所给的自由度对应位置直接累加到系统的总体平衡方程中，即

$$\begin{array}{c} \\ 1\\2\\3\\4\\5\\6 \end{array}\begin{array}{c} \begin{matrix}1 & 2 & 3 & 4 & 5 & 6\end{matrix} \\ \begin{bmatrix} k_{11}^1+k_{11}^2 & k_{12}^1+k_{12}^2 & k_{13}^1 & k_{14}^1 & k_{13}^2 & k_{14}^2 \\ k_{21}^1+k_{21}^2 & k_{22}^1+k_{22}^2 & k_{23}^1 & k_{24}^1 & k_{23}^2 & k_{24}^2 \\ k_{31}^1 & k_{32}^1 & k_{33}^1 & k_{34}^1 & 0 & 0 \\ k_{41}^1 & k_{42}^1 & k_{43}^1 & k_{44}^1 & 0 & 0 \\ k_{31}^2 & k_{32}^2 & 0 & 0 & k_{33}^2 & k_{34}^2 \\ k_{41}^2 & k_{42}^2 & 0 & 0 & k_{43}^2 & k_{44}^2 \end{bmatrix}\end{array}\begin{Bmatrix} u^1 \\ u^2 \\ u^3 \\ u^4 \\ u^5 \\ u^6 \end{Bmatrix} = \begin{Bmatrix} f_1^1+f_1^2 \\ f_2^1+f_2^2 \\ f_3^1 \\ f_4^1 \\ f_3^2 \\ f_4^2 \end{Bmatrix} \tag{3.2-18}$$

式（3.2-18）就是经过单元①和②组装后的总体平衡方程，所有单元平衡方程都进行类似的组装后就得到最终的总体平衡方程，或称总体刚度方程

$$\boldsymbol{KU}=\boldsymbol{F} \tag{3.2-19}$$

式中，$\boldsymbol{K}$ 称为总体刚度矩阵，$\boldsymbol{U}$ 称为位移向量，$\boldsymbol{F}$ 称为载荷向量。

总体刚度矩阵 $\boldsymbol{K}$ 具有以下特性。

1）对称性。如果单元刚度矩阵 $\boldsymbol{k}$ 是对称的，组装后 $\boldsymbol{K}$ 也是对称的。

2）奇异性。在没有经过边界条件约束处理前，$\boldsymbol{K}$ 是奇异的。

3）稀疏性。对于大规模求解问题来说，$\boldsymbol{K}$ 中的零元素是非常多的，一般大规模问题，有时会含有99%以上的零元素。

4）非零元素带状分布。如果有限元节点号排列的比较合理，$\boldsymbol{K}$ 的带宽会比较小。

1.3　约束处理过程

组装后的总体平衡方程还不能马上求解。总体平衡方程式（3.2-19）没有经过边界条件约束处理，它还是一个奇异的方程组，也就是说它有无穷多组解。对于固体力学问题的物理意义是，受载系统如果不受任何外部约束，将进行无限制的刚体运动。而对于实际问题来说，系统处于平衡状态的边界条件将确定方程组式（3.2-19）的惟一一组解。因此，有限元分析过程中，边界条件的约束处理是必不可少的过程。

1.3.1　边界条件分类

线弹性问题的边界条件都是已知的，主要分力和位移两类边界条件。

力边界条件常见的主要有集中载荷力、作用于系统表面的分布力、系统重量引起的自重力、热交换和热流引起的温度载荷等等。这些载荷项的分配可以通过式（3.2-14）求得，将载荷分配到相应的有限元节点上。

位移边界条件会制约系统的刚体位移，消除 $\boldsymbol{K}$ 的奇异性。常见约束类型主要如下。

1）固定位移。例如，$U_2=0.0$；$U_5=0.0$；$U_6=0.0$ 等约束位移大小等于零的边界条件。

2）强制位移。例如，$U_1=1.0$；$U_2=0.5$；$U_4=1.5$ 等约束位移大小是常数且不等于零的边界条件。

3）关联位移。例如，$U_8=kU_7+C$，式中 k 和 C 是一个确定的常数。

如果 $k=0$，则关联位移约束条件就退化成强制位移约束条件。如果 $k=0$，$C=0$，则关联位移约束条件就退化成固定位移约束条件。因此，下面主要给出关联位移边界条件的约束处理方法。

1.3.2　边界条件约束处理方法

为了公式推导方便，我们以6阶方程组为例说明约束处理过程，然后总结约束处理方法。假设要约束处理的原始有限元平衡方程为

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23} & K_{24} & K_{25} & K_{26} \\ K_{31} & K_{32} & K_{33} & K_{34} & K_{35} & K_{36} \\ K_{41} & K_{42} & K_{43} & K_{44} & K_{45} & K_{46} \\ K_{51} & K_{52} & K_{53} & K_{54} & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63} & K_{64} & K_{65} & K_{66} \end{bmatrix}\begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{Bmatrix} \tag{3.2-20}$$

式中，方程组系数矩阵 K_{ij}是对称的。

假设关联约束方程为

$$U_4=kU_3+C \tag{3.2-21}$$

常见的约束处理方法有两种：赋0赋1法和乘大数法。下面将分别说明它们的具体实施方法。

（1）赋0赋1法

将关联约束方程式（3.2-21）代入到有限元平衡方程式（3.2-20）中，并将常数项移到方程组等式的右端，则式（3.2-20）变换成

$$\begin{bmatrix} K_{11} & K_{12} & K_{13}+kK_{14} & 0 & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23}+kK_{24} & 0 & K_{25} & K_{26} \\ K_{31} & K_{32} & K_{33}+kK_{34} & 0 & K_{35} & K_{36} \\ K_{41} & K_{42} & K_{43}+kK_{44} & 0 & K_{45} & K_{46} \\ K_{51} & K_{52} & K_{53}+kK_{54} & 0 & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63}+kK_{64} & 0 & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1-CK_{14} \\ F_2-CK_{24} \\ F_3-CK_{34} \\ F_4-CK_{44} \\ F_5-CK_{54} \\ F_6-CK_{64} \end{Bmatrix} \tag{3.2-22}$$

式（3.2-22）是约束后的有限元平衡方程，它有6个方程，5个未知数（其中 U_4 已经被消去），如果约束方程式（3.2-21）可以消除有限元平衡方程式（3.2-20）的奇异性，则取式（3.2-22）中任意5个方程联立求解，都会得到方程组的惟一一组解。

但是式（3.2-22）的系数矩阵由原来的对称的变成了非对称的，这对于大规模有限元方程组求解是十分不利的，采用相同的求解方法，在求解时间和矩阵存贮容量方面都增加了一倍。

比较式（3.2-20）和式（3.2-22）可以发现，式（3.2-22）的系数矩阵是在式（3.2-20）的基础上做了一次列初等变换，因此，为了保持有限元平衡方程组系数矩阵的对称性，对式（3.2-22）再做一次相同的行初等变换，即式（3.2-22）的第4行乘以系数 k 加到第3行，并去掉第4行，为了保持系数矩阵的阶数，将第4行的所有元素赋0，在其对角线元素位置赋1，即所谓赋0赋1法，可得

$$\begin{bmatrix} K_{11} & K_{12} & K_{13}+kK_{14} & 0 & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23}+kK_{24} & 0 & K_{25} & K_{26} \\ K_{31}+kK_{41} & K_{32}+kK_{42} & K_{33}+kK_{34}+k(K_{43}+kK_{44}) & 0 & K_{35}+kK_{45} & K_{36}+kK_{46} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ K_{51} & K_{52} & K_{53}+kK_{54} & 0 & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63}+kK_{64} & 0 & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1-CK_{14} \\ F_2-CK_{24} \\ F_3-CK_{34}+k(F_4-CK_{44}) \\ 0 \\ F_5-CK_{54} \\ F_6-CK_{64} \end{Bmatrix} \tag{3.2-23}$$

可以发现经过初等变换，方程组式（3.2-23）的系数矩阵仍然保持对称性，而且所做的初等变换也不会改变方程组的解，也就是说式（3.2-23）与式（3.2-22）的解是相同的。通过方程组式（3.2-23）可以求得除 U_4 以外的5个未知数，然后再通过关联约束方程式（3.2-21）回代求解 U_4。

式（3.2-23）的第4行实际上不参与方程组求解，因此我们对式（3.2-23）可以进一步改进，将关联约束方程式（3.2-21）直接引进方程组式（3.2-23）中，使自由度 U_4 也直接参与方程组求解。具体做法是：用关联约束方程式（3.2-21）取代方程组式（3.2-23）的赋0赋1行（第4行），然后再做对称化处理。也就是将取代后的一行方程（第4行）乘以 $-k$ 加到第3行，经过这样处理方程组式（3.2-23）的系数矩阵仍然保持对称性

$$\begin{bmatrix} K_{11} & K_{12} & K_{13}+kK_{14} & 0 & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23}+kK_{24} & 0 & K_{25} & K_{26} \\ K_{31}+kK_{41} & K_{32}+kK_{42} & K_{33}+kK_{34}+k(K_{43}+kK_{44})+k^2 & -k & K_{35}+kK_{45} & K_{36}+kK_{46} \\ 0 & 0 & -k & 1 & 0 & 0 \\ K_{51} & K_{52} & K_{53}+kK_{54} & 0 & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63}+kK_{64} & 0 & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1-CK_{14} \\ F_2-CK_{24} \\ F_3-CK_{34}+k(F_4-CK_{44})-kC \\ C \\ F_5-CK_{54} \\ F_6-CK_{64} \end{Bmatrix} \tag{3.2-24}$$

这样通过求解方程组式（3.2-24）就可以一次解出所有未知数。上面所述就是关联位移边界条件的约束处理过程，其基本原理就是利用初等变换对求解方程组进行相同的行列变换，既保证方程组解不会改变，又可以保持方程组系数矩阵的对称性。在进行初等变换时，只要保证对方程组系数矩阵做相同的行列变换，就可以保持方程组系数矩阵的对称性。

固定位移和强制位移边界条件约束处理可以通过式（3.2-24）进行简化。关联约束方程式（3.2-21）中如果 $k=0$，则退化成强制位移边界条件，即

$$U_4 = C \tag{3.2-25}$$

则约束后的方程组式（3.2-24）就简化为

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & 0 & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23} & 0 & K_{25} & K_{26} \\ K_{31} & K_{32} & K_{33} & 0 & K_{35} & K_{36} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ K_{51} & K_{52} & K_{53} & 0 & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63} & 0 & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1-CK_{14} \\ F_2-CK_{24} \\ F_3-CK_{34} \\ C \\ F_5-CK_{54} \\ F_6-CK_{64} \end{Bmatrix} \tag{3.2-26}$$

如果关联约束方程式（3.2-21）中，$k=0$，$C=0$，则退化成固定位移边界条件，即

$$U_4 = 0 \tag{3.2-27}$$

则约束后的方程组式（3.2-26）就进一步简化为

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & 0 & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23} & 0 & K_{25} & K_{26} \\ K_{31} & K_{32} & K_{33} & 0 & K_{35} & K_{36} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ K_{51} & K_{52} & K_{53} & 0 & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63} & 0 & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \\ F_3 \\ 0 \\ F_5 \\ F_6 \end{Bmatrix} \tag{3.2-28}$$

从式（3.2-26）和式（3.2-28）可以看出，固定位移和强制位移边界条件的赋0赋1约束处理相对比较简单，而且它们的系数矩阵约束后是相同的，只是简单地将方程组系数矩阵中要约束自由度的行列分别赋0，对角线元素赋1，这也是赋0赋1法的由来。在方程组载荷右端项的处理方法上两者是不同，处理固定位移边界条件时，只要将对应自由度的载荷赋0即可。处理强制位移边界条件时，要在方程组系数矩阵未赋0赋1法前，先将对应自由度的列乘以系数 C 减到载荷右端项，再将对应自由度的载荷位置赋 C。

（2）乘大数法

乘大数法约束位移边界条件原理还是利用矩阵的初等变换不改变方程组解的思想。

仍然以有限元平衡方程式（3.2-20）和关联位移约束方程式（3.2-21）为例说明乘大数法约束处理方法的具体实施过程。

将关联约束方程式（3.2-21）乘以一个大数 A。大数 A 一般是式（3.2-20）系数矩阵中对角线元素 K_{44} 的 10^{10} 倍量级

左右。乘大数 A 后式（3.2-21）整理成如下形式

$$-AkU_3 + AU_4 = AC \tag{3.2-29}$$

将式（3.2-29）加到有限元平衡方程式（3.2-20）约束处理对应自由度行，即第 3 行或第 4 行，这里我们加到第 4 行

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23} & K_{24} & K_{25} & K_{26} \\ K_{31} & K_{32} & K_{33} & K_{34} & K_{35} & K_{36} \\ K_{41} & K_{42} & K_{43} - Ak & K_{44} + A & K_{45} & K_{46} \\ K_{51} & K_{52} & K_{53} & K_{54} & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63} & K_{64} & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 + AC \\ F_5 \\ F_6 \end{Bmatrix} \tag{3.2-30}$$

如果约束方程式（3.2-21）可以消除有限元平衡方程式（3.2-20）的奇异性，方程组式（3.2-30）的解就是方程组式（3.2-20）惟一的一组解。但是经过约束处理后，方程组式（3.2-30）的系数矩阵是非对称的，下面将利用初等变换方法将变换成对称的。从式（3.2-30）可以发现，系数矩阵中只（3，4）和（4，3）两个元素是不对称的，因此，将已乘完大数的关联约束方程式（3.2-29）再乘以系数 $-k$，得

$$Ak^2 U_3 - AkU_4 = -AkC \tag{3.2-31}$$

将经过处理的关联约束方程加到方程组式（3.2-30）的第 3 行，得

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23} & K_{24} & K_{25} & K_{26} \\ K_{31} & K_{32} & K_{33} + Ak^2 & K_{34} - Ak & K_{35} & K_{36} \\ K_{41} & K_{42} & K_{43} - Ak & K_{44} + A & K_{45} & K_{46} \\ K_{51} & K_{52} & K_{53} & K_{54} & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63} & K_{64} & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \\ F_3 - AkC \\ F_4 + AC \\ F_5 \\ F_6 \end{Bmatrix} \tag{3.2-32}$$

经过这些初等变换以后，约束后的有限元平衡方程式（3.2-32）仍然保持对称性。以上就是乘大数法约束处理关联位移约束边界条件的方法和过程。

固定位移和强制位移边界条件约束处理可以通过式（3.2-32）进行简化。关联约束方程式（3.2-21）中如果 $k=0$，则退化成强制位移边界条件，约束后的方程组式（3.2-32）就简化为

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23} & K_{24} & K_{25} & K_{26} \\ K_{31} & K_{32} & K_{33} & K_{34} & K_{35} & K_{36} \\ K_{41} & K_{42} & K_{43} & K_{44} + A & K_{45} & K_{46} \\ K_{51} & K_{52} & K_{53} & K_{54} & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63} & K_{64} & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 + AC \\ F_5 \\ F_6 \end{Bmatrix} \tag{3.2-33}$$

如果关联约束方程式（3.2-21）中，$k=0$，$C=0$，则退化成固定位移边界条件，约束后的方程组式（3.2-33）就进一步简化为

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} & K_{15} & K_{16} \\ K_{21} & K_{22} & K_{23} & K_{24} & K_{25} & K_{26} \\ K_{31} & K_{32} & K_{33} & K_{34} & K_{35} & K_{36} \\ K_{41} & K_{42} & K_{43} & K_{44} + A & K_{45} & K_{46} \\ K_{51} & K_{52} & K_{53} & K_{54} & K_{55} & K_{56} \\ K_{61} & K_{62} & K_{63} & K_{64} & K_{65} & K_{66} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{Bmatrix} \tag{3.2-34}$$

从式（3.2-33）和式（3.2-34）可以看出，固定位移和强制位移边界条件的乘大数约束处理相对比较简单，而且它们的系数矩阵约束后是相同的，只是简单地将方程组系数矩阵中要约束自由度的对角线元素加上一个相对大数 A 即可。从这点上说，乘大数法的叫法并不十分准确，应该叫加大数法更贴切。当然对于固定位移和强制位移边界条件约束处理来说，乘大数和加大数对方程组式（3.2-33）和式（3.2-34）的解影响非常小，但是对关联位移边界条件约束处理来说，只乘大数就无法实施整个约束处理过程了。

（3）赋 0 赋 1 法和乘大数法的比较

两种方法都可以消除有限元平衡方程的奇异性，得到符合实际边界条件的惟一一组解。但两种方法还是有很大的区别。

1）赋 0 赋 1 法在约束处理过程中是严格精确的，而乘大数法是一种近似约束处理方法，它的精度取决于所乘大数 A 值。

2）采用乘大数法约束处理后的有限元平衡方程在求解时可能造成解的失真，大数 A 值越大可能解的偏差会越大。而赋 0 赋 1 法就不会出现类似的问题，它在约束过程和求解过程都是精确的。

3）乘大数法相对于赋 0 赋 1 法在约束处理过程上简单一些。

4）赋 0 赋 1 法实际上是将关联位移约束方程代入到有限元平衡方程中的，是代入法。而乘大数是将占绝对优势的关联位移约束方程合并到有限元平衡方程中的，是罚方法，计算误差来自于合并过程，计算精度取决于关联位移约束方程的优势大小。

在现有商业软件中，位移边界条件的约束处理都采用赋 0 赋 1 法，乘大数很少被采用，主要原因是它是一种近似方法，而且大数的大小也不好确定，有时还会造成求解失败。

1.4　方程组求解过程

位移法最终要求解约束处理后的式（3.2-19）线性方程组，它也是有限元计算过程中很重要的一部分，直接影响有限元法的实际应用。在有限元法的发展过程中，有限元方程的求解效率一直是其应用的最大瓶颈之一，这也曾是有限元法的重点研究内容。

前面已经介绍过大规模有限元方程组的系数矩阵具有对称、稀疏、带状分布以及正定、主元占优等特点，因此在有限元方程组求解时要有效地利用这些特点，以减少系数矩阵的存贮量，提高方程组求解效率。

线性方程组的解法主要分两大类：直接解法和迭代解法。

直接解法以高斯消去法为基础，以等带宽或变带宽方式存贮系数矩阵内元素。对于求解规模比较大的问题，要存贮的元素非常巨大，但是这种方法的求解效率比较高。因此，在方程组阶数不是非常高时，一般采用直接解法。迭代解法只需要存贮系数矩阵中非零元素，存贮量很小，一般是变带宽存贮量的 20% 或更少，有些算法的求解效率也非常高，适合求解大规模线性方程组。但是这种解法对接近病态的方程组很难保证收敛性。

1.4.1　带宽

有限元方程组系数矩阵是稀疏的、非零元素呈带状分布，带宽就是它的宽度。带宽直接关系到系数矩阵元素的存贮量和方程组的求解效率。带宽的大小是由系统有限元网格的节点号排序决定的，具体求法是

$$\text{带宽} = (\text{单元最大节点号之差} + 1) \times \text{节点自由度数} \tag{3.2-35}$$

带宽是网格节点标注方法直接决定的，不同标注方法带宽可能相关很大。例如图 3.2-3a、b、c 所示四边形网格的三种节点号标注方法，每个节点是 2 个自由度。则按图 3.2-3a 方法标注节点号，结构的带宽是 12，按图 3.2-3b 方法标注节点号，带宽是 18，按图 3.2-3c 方法标注节点号，带宽是 56。三者之间有比较大的差距，其中图 3.2-3a 和图 3.2-3c 之

间相差近5倍，这就意味着系数矩阵的存贮量也是相差5倍。因此，对于大规模复杂系统的节点号优化是十分必要的。

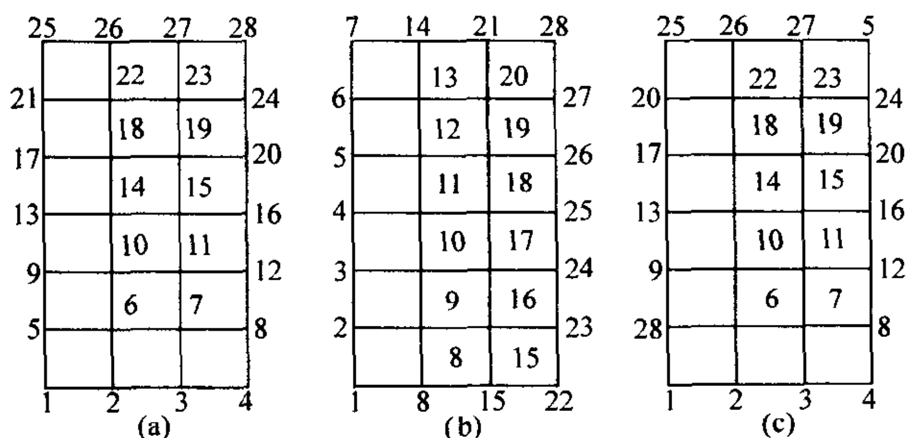


图 3.2-3　二维四边形网格的三种节点标注方法

1.4.2　系数矩阵的存贮

如果节点号排序优化的比较好，系数矩阵的存贮量就会减少很多。根据系数矩阵的对称性，一般都是按半带宽存贮。主要有两种方法：二维等带宽存贮和一维变带宽存贮。

1）二维等带宽存贮　对于 n 阶的系数矩阵，若取最大的半带宽 D 为带宽，则上三角阵中的全部非零元素都将包括在这条以主对角元素为一边的一条等带宽中，如图 3.2-4 所示。二维等带宽存贮是将这条带中的元素以二维数组的形式存贮在计算机的内存中，如图 3.2-4b 所示，二维数组的大小为 $n\times D$。

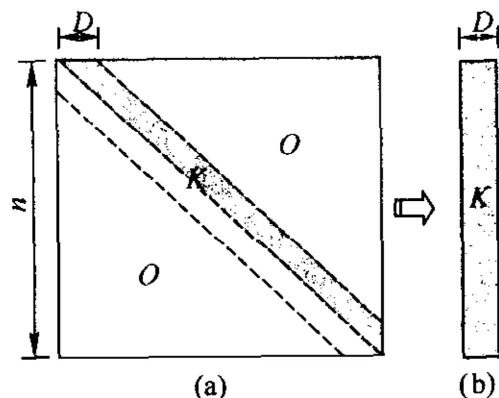


图 3.2-4　二维等带宽存贮

二维等带宽存贮消除了最大带宽以外的全部零元素，节省了系数矩阵元素的存贮量。但是由于取最大带宽为存贮范围，因此不能排除在带宽内的大量零元素。当系数矩阵的各行带宽变化不大时，适合采用二维等带宽存贮，方程组求解过程中系数矩阵元素的寻址也比较方便，求解效率较高。例如，图 3.2-3a 和 b 所示系统的节点号标注方法中，每个单元的带宽是相同，因此适合用二维等带宽存贮系数矩阵。但是当出现局部带宽特别大的情况时，采用二维等带宽存贮时，将由于局部带宽过大而使整体系数矩阵的存贮大大增加。例如，图 3.2-3c 所示结构的节点号标注方法中，多数单元的带宽是 12，但有一个单元的带宽是 56，如果按二维等带宽存贮，要存贮的系数矩阵的大小为 56×56，与系数矩阵的大小是相同的。因此，这时应采用另外一种存贮方式：一维变带宽存贮方法。

2）一维变带宽存贮　一维变带宽存贮方法就是把变化的带宽内的元素按一定的顺序存贮在一个一维数组中。由于它不按最大带宽存贮，因此比二维等带宽存贮更节省内存。按照解法可分为按行一维变带宽存贮和按列一维变带宽存贮。下面以按行一维变带宽存贮为例说明存贮过程。

如图 3.2-5a 所示，K_{ij} 为 8×8 的对称系数矩阵，实线内为各行带宽内元素，这里仍然包含了大量的零元素，它们都要保存到一个一维数组 A 中。按行一维变带宽存贮是按行依次存贮系数矩阵元素，每行应从第一个非零元素开始直到主对角元素，然后再从下一行的第一个非零元素开始到主对角元素，以此类推，直到带宽内所有元素都被存贮，如图 3.2-5b 所示，按照图（b）所示规律将图（a）实线内的各行元素依次保存到 A 数组中。为了区别 A 数组中的元素在系数矩阵 K_{ij} 中的位置，还需要一个辅助的寻址数组 M，记录 K_{ij} 中主对角元素在 A 数组中的位置

$$M=(1\quad 3\quad 4\quad 6\quad 9\quad 12\quad 18\quad 20)\tag{3.2-36}$$

一维变带宽存贮是最节省内存的一种方法，但是由于要借助于寻址数组寻找系数矩阵元素的位置，相对二维等带宽存贮方法来说要复杂一些，而且在程序实现时也要复杂得多，方程组求解过程中也要消耗一些数组寻址时间。因此，在选用存贮方法时要权衡两者的利弊，统盘考虑。一般当带宽变化不大，计算机内存允许时，采用二维等带宽存贮方法是比较合适的。

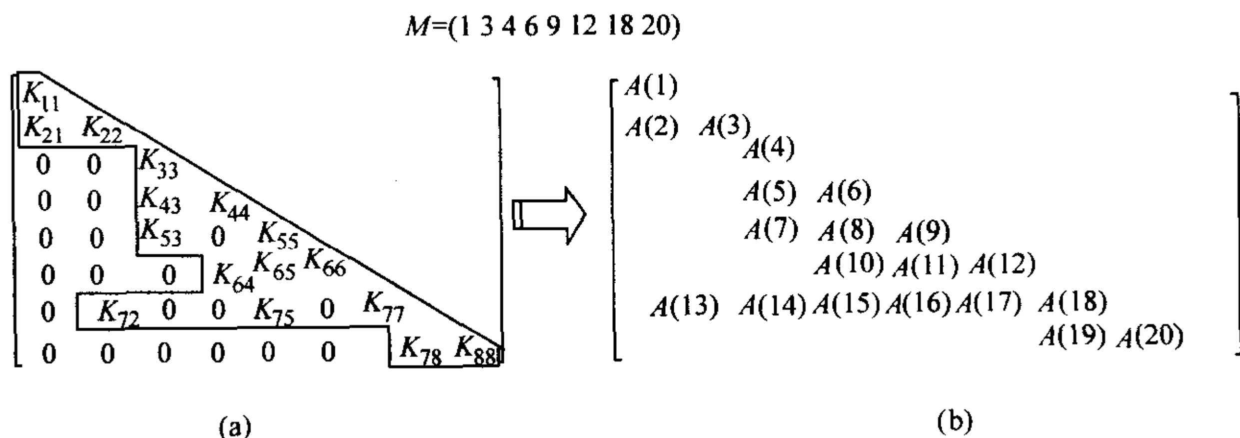


图 3.2-5　一维变带宽存贮对应关系

1.4.3　高斯消去法

对于约束处理后的 n 阶线性有限元平衡方程 $\boldsymbol{KU}=\boldsymbol{F}$，需要进行 $n-1$ 次消元。采用循序消去时，第 m 次消元以 $m-1$次消元后的第 m 行元素作为主元行，$K_{mm}^{(m-1)}$ 为主元，对第 i 行元素（$i>m$）的消元

$$K_{ij}^{(m)}=K_{ij}^{(m-1)}-\frac{K_{im}^{(m-1)}}{K_{mm}^{(m-1)}}K_{mj}^{(m-1)}\tag{3.2-37}$$

$$F_{i}^{(m)}=F_{i}^{(m-1)}-\frac{K_{im}^{(m-1)}}{K_{mm}^{(m-1)}}F_{m}^{(m-1)}\tag{3.2-38}$$

式中，$m=1,\ 2,\ \cdots,\ n-1$；$i,\ j=m+1,\ m+2,\ \cdots,\ n$；$K_{ij}^{(m)}$，$F_{i}^{(m)}$ 的上角标表示该元素是经过 m 次消元后得到的结果。式（3.2-37）和式（3.2-38）表明，第 m 次消元是在第 $m-1$ 次消元基础上进行的。可以把经过 m 次消元后的系数矩阵和载荷向量分别表示为 $\boldsymbol{K}^{(m)}$ 和 $\boldsymbol{F}^{(m)}$。

消元完成后，就可进行回代求解。把消元过程的最后结果记为

$$\left.\begin{aligned}\boldsymbol{K}^{(n-1)}&=\boldsymbol{S}\\ \boldsymbol{F}^{(n-1)}&=\boldsymbol{V}\end{aligned}\right.\tag{3.2-39}$$

式中，S 为上三角阵，回代公式可表示为

$$a_n = \frac{F_n^{(n-1)}}{K_{nn}^{(n-1)}} = \frac{V_n}{S_{nn}}$$

$$a_i = (F_i^{(n-1)} - \sum_{j=i+1}^{n} K_{ij}^{(n-1)} a_j)/K_{ii}^{(n-1)} = (V_i - \sum_{j=i+1}^{n} S_{ij} a_j)/S_{ii} \tag{3.2-40}$$

式中，$i = n-1, n-2, \cdots, 3, 2, 1$。

回代过程是从最后一行开始的，当回代求解 a_i 时，a_{i+1}，a_{i+2}，…，a_n 已经解出。在程序实现过程中，为节省计算机内存，一般将求得的解 a 放在 F 中，因此回代公式可改写为

$$\begin{aligned} F_n &= \frac{F_n}{K_{nn}} \\ F_i &= (F_i - \sum_{j=i+1}^{n} K_{ij} F_j)/K_{ii} \end{aligned} \tag{3.2-41}$$

式中，$i = n-1, n-2, \cdots, 3, 2, 1$。

1.4.4 三角分解法

三角分解定理：设 n 阶矩阵 $\boldsymbol{K}$ 所有主子式

$$K_{11}, \begin{vmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{vmatrix}, \cdots, \begin{vmatrix} K_{11} & K_{12} & \cdots & K_{1n} \\ K_{21} & K_{22} & \cdots & K_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ K_{n1} & K_{n2} & \cdots & K_{nn} \end{vmatrix} \tag{3.2-42}$$

均不等于零，则矩阵 $\boldsymbol{K}$ 必然可以分解为下三角阵 $\boldsymbol{L}$ 和上三角阵 $\boldsymbol{S}$ 的乘积，即

$$\boldsymbol{K} = \boldsymbol{LS} \tag{3.2-43}$$

当 $\boldsymbol{L}$ 或 $\boldsymbol{S}$ 为单位三角阵时，这种分解是唯一的，称为 Crout 分解。

有限元平衡方程的系数矩阵 $\boldsymbol{K}$ 符合三角分解定理的条件，可以进行三角分解。由于 $\boldsymbol{K}$ 是对称矩阵，$\boldsymbol{K}$ 还可以表示为

$$\boldsymbol{K} = \boldsymbol{LDL}^{\mathrm{T}} \tag{3.2-44}$$

式中，$\boldsymbol{D}$ 是对角矩阵，$\boldsymbol{DL}^{\mathrm{T}} = \boldsymbol{S}$。

根据矩阵乘法规则

$$K_{ij} = \sum_{r=1}^{n} l_{ir} d_{rr} l_{rj}^{T} = \sum_{r=1}^{n} l_{ri}^{\mathrm{T}} \mathrm{d}_{rr} l_{rj}^{\mathrm{T}} \tag{3.2-45}$$

式中，K_{ij}，l_{ir}，d_{rr}，l_{rj}^{T} 分别为矩阵 $\boldsymbol{K}$，$\boldsymbol{L}$，$\boldsymbol{D}$，$\boldsymbol{L}^{\mathrm{T}}$ 中的元素。由于 $\boldsymbol{L}$ 和 $\boldsymbol{L}^{\mathrm{T}}$ 互为转置，因此

$$l_{ir} = l_{ri}^{\mathrm{T}} \tag{3.2-46}$$

考虑到 $\boldsymbol{L}^{\mathrm{T}}$ 是上三角阵，因此有

$$\begin{aligned} &\text{当 } r > i \text{ 时}\quad l_{ri}^{\mathrm{T}} = 0 \\ &\text{当 } r > j \text{ 时}\quad l_{rj}^{\mathrm{T}} = 0 \end{aligned} \tag{3.2-47}$$

当 K_{ij} 为 $\boldsymbol{K}$ 中上三角矩阵中的元素时，则 $j \geqslant i$，式（3.2-47）中求和 $r = 1 \sim n$ 可改成 $r = 1 \sim i$。因此式（3.2-45）可写成

$$K_{ij} = \sum_{r=1}^{i} l_{ri}^{\mathrm{T}} d_{rr} l_{rj}^{\mathrm{T}} = \sum_{r=1}^{i-1} l_{ri}^{\mathrm{T}} d_{rr} l_{rj}^{\mathrm{T}} + l_{ii}^{\mathrm{T}} d_{ii} l_{ij}^{\mathrm{T}} \tag{3.2-48}$$

$\boldsymbol{L}^{\mathrm{T}}$ 是单位上三角阵，因此对角元素 $l_{ii}^{\mathrm{T}} = 1$。由式（3.2-48）可得递推公式

$$\begin{aligned} d_{ii} l_{ij}^{\mathrm{T}} &= S_{ij} = K_{ij} - \sum_{r=1}^{i-1} l_{ri}^{\mathrm{T}} d_{rr} l_{rj}^{\mathrm{T}} \\ l_{ij}^{\mathrm{T}} &= S_{ij}/d_{ii} \end{aligned} \tag{3.2-49}$$

式中

$$d_{ii} = S_{ii} \tag{3.2-50}$$

进行三角分解时，可以对原系数矩阵按行或按列进行分解。由于系数矩阵的对称性，只需要分解系数矩阵的上三角部分。

三角分解后的求解步骤是：

1）由 $\boldsymbol{LV} = \boldsymbol{F}$ 向前回代求解 $\boldsymbol{V}$；
2）求对角阵 $\boldsymbol{D}$ 的逆矩阵 $\boldsymbol{D}^{-1}$；
3）由 $\tilde{\boldsymbol{V}} = \boldsymbol{D}^{-1}\boldsymbol{V}$ 求修正的分解后载荷矩阵 $\tilde{\boldsymbol{V}}$；
4）由 $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{a} = \tilde{\boldsymbol{V}}$ 向后回代求解基本未知量 $\boldsymbol{a}$。

1.4.5 高斯消去法与三角分解法的比较

高斯消去法分为消去过程和回代求解过程两部分。三角分解法从形式上看似乎没有消去过程，只有回代求解过程，但是三角分解法在回代求解前必须要进行系数矩阵的三角分解过程，这实际上就相当于高斯消去法的消去过程。

由于这两种解法都是用的高斯消去原理，因此它们的运算量基本相同。但在程序实现时，对相同的系数矩阵，三角分解消耗的机时却比高斯消去法要少。这是由于系数矩阵都以一维或二维数组的形式存贮在计算机的内存中，对一个元素进行运算时，需要经过由数组中取出该元素，以及运算后将该元素再送回数组中去的读取和保存元素的过程，这些元素在计算机运行过程中需要在内存和 CPU 的一级缓存、二级缓存间堆栈交换数据，因此也要消耗一定的机时，特别是当数据堆栈次数很多或堆栈量很大时，要消耗更多的计算时间。三角分解法运算过程中读取和保存元素的次数少，每个元素的分解都是一次完成。而高斯消去法的一个元素的消元要读取和保存 $1 \sim \boldsymbol{D}^{-1}$ 次。因此，采用三角分解法比高斯消去法节省计算时间。

1.4.6 雅可比（Jacobi）迭代法

假设约束处理后的有限元平衡方程为

$$\boldsymbol{KU} = \boldsymbol{F} \tag{3.2-51}$$

由于系数矩阵 $\boldsymbol{K}$ 是非奇异的，且主对角元素 $K_{ii} \neq 0$，则 $\boldsymbol{K}$ 可分解为

$$\boldsymbol{K} = \boldsymbol{D} + \boldsymbol{K}_0 \tag{3.2-52}$$

式中，$\boldsymbol{D}$ 为对角矩阵

$$\boldsymbol{D} = \begin{bmatrix} K_{11} & 0 & \cdots & 0 \\ 0 & K_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & K_{nn} \end{bmatrix} \tag{3.2-53}$$

$$\boldsymbol{K}_0 = \begin{bmatrix} 0 & K_{12} & K_{13} & \cdots & K_{1n} \\ K_{21} & 0 & K_{23} & \cdots & K_{2n} \\ K_{31} & K_{32} & 0 & \cdots & K_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ K_{n1} & K_{n2} & K_{n3} & \cdots & 0 \end{bmatrix} \tag{3.2-54}$$

进一步将式（3.2-51）变换为

$$\boldsymbol{U} = \overline{\boldsymbol{K}}\boldsymbol{U} + \overline{\boldsymbol{F}} \tag{3.2-55}$$

式中

$$\overline{\boldsymbol{K}} = \begin{bmatrix} 0 & -\frac{K_{12}}{K_{11}} & -\frac{K_{13}}{K_{11}} & \cdots & -\frac{K_{1n}}{K_{11}} \\ -\frac{K_{21}}{K_{22}} & 0 & -\frac{K_{23}}{K_{22}} & \cdots & -\frac{K_{2n}}{K_{22}} \\ -\frac{K_{31}}{K_{33}} & -\frac{K_{32}}{K_{33}} & 0 & \cdots & -\frac{K_{3n}}{K_{33}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\frac{K_{n1}}{K_{nn}} & -\frac{K_{n2}}{K_{nn}} & -\frac{K_{n3}}{K_{nn}} & \cdots & 0 \end{bmatrix} \tag{3.2-56}$$

$$\overline{\boldsymbol{F}} = \left[\frac{F_1}{K_{11}} \quad \frac{F_2}{K_{22}} \quad \frac{F_3}{K_{33}} \quad \cdots \quad \frac{F_n}{K_{nn}} \right]^{\mathrm{T}} \tag{3.2-57}$$

系数矩阵 $\boldsymbol{K}$ 与 $\overline{\boldsymbol{K}}$ 及载荷向量 $\boldsymbol{F}$ 与 $\overline{\boldsymbol{F}}$ 之间存在如下关系

$$\overline{\boldsymbol{K}} = \boldsymbol{I} - \boldsymbol{D}^{-1}\boldsymbol{K} \tag{3.2-58}$$

$$\overline{\boldsymbol{F}} = \boldsymbol{D}^{-1}\boldsymbol{F} \tag{3.2-59}$$

雅可比迭代法的计算公式为

$$\boldsymbol{U}^{k+1} = \overline{\boldsymbol{K}}\boldsymbol{U}^{k} + \overline{\boldsymbol{F}} \tag{3.2-60}$$

式中，$\overline{\boldsymbol{K}}$ 为雅可比迭代法的迭代矩阵，k 为迭代次数。

雅可比迭代法公式简单，迭代思路明确，每迭代一次只需要计算 n 个方程的向量乘法。

1.4.7　高斯－赛德尔（Gauss－Seidel）迭代法

由雅可比迭代法可知，在迭代计算 x^{k+1} 过程中，都是采用上一迭代步的结果 x^k。实际上在计算新分量 x_i^{k+1} 时，x_1^{k+1}、x_2^{k+1}、…、x_{i-1}^{k+1} 分量已经被求得，可以认为新计算出来的这些分量比上一迭代步的分量有所改善。这样利用这些分量可以提高迭代过程的收敛速度，这就是高斯－赛德尔迭代法，简称G－S迭代法。

有限元平衡方程组的系数矩阵 $\boldsymbol{K}$ 在式（3.2-52）的基础上还可进一步分解，若将 $\boldsymbol{K}_0$ 再分解为一个下三角矩阵 $\boldsymbol{K}_0^L$ 和一个上三角矩阵 $\boldsymbol{K}_0^U$，则系数矩阵的分解可表示为

$$\boldsymbol{K}=\boldsymbol{D}+\boldsymbol{K}_0^L+\boldsymbol{K}_0^U \tag{3.2-61}$$

式中

$$\boldsymbol{K}_0^L=\begin{bmatrix}0&0&0&\cdots&0\\K_{21}&0&0&\cdots&0\\K_{31}&K_{32}&0&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\K_{n1}&K_{n2}&K_{n3}&\cdots&0\end{bmatrix} \tag{3.2-62}$$

$$\boldsymbol{K}_0^U=\begin{bmatrix}0&K_{12}&K_{13}&\cdots&K_{1n}\\0&0&K_{23}&\cdots&K_{2n}\\0&0&0&\cdots&K_{3n}\\\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&0&\cdots&0\end{bmatrix} \tag{3.2-63}$$

G－S迭代法的计算公式为

$$\boldsymbol{D}\boldsymbol{U}^{k+1}=\boldsymbol{F}-\boldsymbol{K}_0^L\boldsymbol{U}^{k+1}-\boldsymbol{K}_0^U\boldsymbol{U}^k \tag{3.2-64}$$

进一步整理可得

$$(\boldsymbol{D}+\boldsymbol{K}_0^L)\ \boldsymbol{U}^{k+1}=\boldsymbol{F}-\boldsymbol{K}_0^U\boldsymbol{U}^k \tag{3.2-65}$$

若 $\boldsymbol{D}+\boldsymbol{K}_0^L$ 的逆存在，则

$$\boldsymbol{U}^{k+1}=(\boldsymbol{D}+\boldsymbol{K}_0^L)^{-1}\ (\boldsymbol{F}-\boldsymbol{K}_0^U\boldsymbol{U}^k) \tag{3.2-66}$$

这样，G－S迭代法的迭代公式可表示为

$$\boldsymbol{U}^{k+1}=\boldsymbol{E}\boldsymbol{U}^k+\boldsymbol{P} \tag{3.2-67}$$

式中，$\boldsymbol{E}$ 为G－S迭代法的迭代矩阵

$$\boldsymbol{E}=-\ (\boldsymbol{D}+\boldsymbol{K}_0^L)^{-1}\boldsymbol{K}_0^U \tag{3.2-68}$$

$$\boldsymbol{P}=(\boldsymbol{D}+\boldsymbol{K}_0^L)^{-1}\boldsymbol{F} \tag{3.2-69}$$

G－S迭代法的每步迭代的计算量与雅可比迭代法相当，但在程序实现过程中只需存放 $\boldsymbol{U}$ 一个数组，比雅可比迭代法少一个。两种方法都存在解的收敛性问题，当方程组性态不好时，两种解法在迭代过程中都可能发散。要保证迭代过程的收敛性，系数矩阵 $\boldsymbol{K}$ 必须是严格对角占优。这对于临界失稳状态的有限元平衡方程来说不能得到保证，这点在实际应用过程中必须要注意。因此，迭代法与直接法比较的话，收敛性是最大的缺点，但是相对于直接法，它在求解大规模问题方面有很大的优势。直接法由于受计算内存和计算量的限制，只适合求解中等规模的问题。而迭代法在存贮系数矩阵时，只存系数矩阵中非零元素，不是按二维等带宽和一维变带宽方法存贮。这对于大规模稀疏矩阵来说具有很大的优势，一般大规模问题的有限元系数矩阵即使按一维变带宽方法存贮，其中仍然含有80%以上的零元素，这80%以上的零元素在迭代法的系数矩阵存贮时都可以不保存，如果在迭代收敛速度上得到保证，迭代法就可以求解更大规模的实际问题。

1.5　应变、应力回代过程

1.5.1　单元应变和应力回代求解

通过求解有限元平衡方程得到有限元节点位移后，就可以进行系统的刚度校核。如果所分析问题要进行强度校核，就要回代求解单元的应变和应力。

由单元插值关系式（3.2-6）和单元几何关系式（3.2-7）可得单元应变

$$\boldsymbol{\varepsilon}=\boldsymbol{L}\boldsymbol{N}\mathrm{u}^e=\boldsymbol{B}\mathrm{u}^e \tag{3.2-70}$$

得到单元应变后，再由单元本构关系式（3.2-8）求得单元应力。

采用位移法求得单元应力和应变的精度相对比较低。由式（3.2-10）可知，应变矩阵 $\boldsymbol{B}$ 是单元形状矩阵 $\boldsymbol{N}$ 对坐标求导后得到的。求导一次也就意味着插值多项式的阶数就要降低一次，因此通过式（3.2-70）求得的单元应变相对于节点位移误差较大，再由式（3.2-8）求得的单元应力精度也不高。

一般单元应力误差主要表现于：

1）单元内部不满足平衡方程；

2）单元与单元的交界面上应力不连续；

3）在力的边界上也不满足力的边界条件。

1.5.2　单元应力处理

最简单的应力处理方法是单元平均或节点平均，即取相邻单元或围绕节点各单元应力的平均值。

1）相邻单元应力平均值　这种方法常用于3节点三角形单元中。由于这种单元是常应变单元，在各单元间应变和应力是不连续的，呈梯田状，而且单元的应力也总是上下振荡。为了减少这种不合理的振荡，提高单元应力的计算精度，可以取相邻单元应力的平均值作为此二个单元合成的较大四边形单元形心处的应力。

平均应力可以采用算术平均法计算，即

$$\text{平均应力}=(\text{单元1应力}+\text{单元2应力})/2 \tag{3.2-71}$$

也可以采用精确一些的面积加权平均法计算，即

$$\text{平均应力}=\frac{\text{单元1应力}\times\text{单元1面积}+\text{单元2应力}\times\text{单元2面积}}{\text{单元1面积}+\text{单元2面积}} \tag{3.2-72}$$

当相邻单元面积相差不大时，二者的计算结果基本相同。在单元网格划分时应避免相邻单元的面积相差太大，从而使求解的误差相近。

2）节点应力平均值　取围绕该节点周围的相关单元计算得到该节点的应力平均值 σ_i

$$\sigma_i=\frac{1}{m}\sum_{e=1}^{m}\sigma_i^e \tag{3.2-73}$$

式中，m 是围绕在 i 节点周围的单元个数；σ_i^e 是节点 i 周围的单元应力。

2　常用单元模型

2.1　单元模型分类

2.1.1　一维单元

一维常用单元见表3.2-1。主要有：2节点线单元、3节点线单元、梁单元。

表3.2-1　一维常用单元

阶次	单元名称	图示
一次	2节点线单元	1●———●2
	梁单元	1▭▭▭2
二次	3节点线单元	1●——3——●2

2.1.2　二维单元

二维常用单元见表3.2-2。主要有：3节点三角形线性单元、6节点三角形二次单元、10节点三角形三次单元、4节点四边形双线性单元、8节点四边形二次单元、12节点四边形三次单元。

表 3.2-2 二维常用单元

阶次	单元名称	图示
一次	3节点三角形单元	
	4节点四边形单元	
二次	6节点三角形单元	
	8节点四边形单元	
三次	10节点三角形单元	
	12节点四边形单元	

2.1.3 三维单元

三维常用单元见表3.2-3。主要有：4节点四面体线性单元、10节点四面体二次单元、8节点六面体线性单元、20节点六面体二次单元。

表 3.2-3 三维常用单元

阶次	单元名称	图示
一次	4节点四面体单元	
	8节点六面体单元	
二次	10节点四面体单元	
	20节点六面体单元	

2.1.4 空间单元

常用的空间单元主要有：桁架单元、框架单元、板单元、壳单元。

这些空间单元在实际应用时都要根据系统的受力情况引入单元局部随体坐标系，以简化单元的受力状态，提高有限元的计算效率。

1）桁架单元 一维2节点线单元＋单元局部随体坐标系。

由于每根杆的两端都是角接，因此空间桁架单元在受力过程中只能承受拉压，即只能反抗轴力，这与一维2节点线单元相同。因此我们采用相同的标准和度量在每个空间杆上建立一个单元局部随体坐标系。在局部坐标系中，空间桁架的每根杆每变成了一维2节点线单元。

2）框架单元 三维梁单元＋一维2节点线单元＋单元局部随体坐标系。

在空间框架结构中每根的两端都是刚性联结的，因此它在变形过程中，可以承受拉压、弯曲、扭转3种变形模式。为了能全面反映它们的作用，选用三维梁单元＋一维2节点线单元＋单元局部随体坐标系作为空间框架单元。

3）板单元 板单元主要分薄板单元和中厚板单元。

平板在承受法向载荷作用时，主要以弯曲和横向剪切2种变形模式抵抗板的变形，即中厚板单元。如果板很薄、则可以忽略横向剪切抗力，认为抵抗载荷的主要因素是弯矩，即薄板单元。

4）壳单元 壳单元从几何和变形方面区分，与板单元类似，也分为薄壳单元和中厚壳单元。从单元构造方面区分，主要有3种方法。

① 组合单元 抵抗拉压变形的二维单元＋板单元＋单元局部随体坐标系。适合于薄壳单元和中厚壳单元。

② 壳理论单元 由空间壳理论严格构造的壳单元。适合于薄壳单元和中厚壳单元。

③ 退化单元 由三维实体单元退化成的壳单元。只适合于中厚壳单元。

2.2 单元模型构造

有限元法的基本思想是通过单元分片近似，在每个单元内假设近似函数来分片表示单元的场函数。因此，有限元法的一个重要内容是在每个单元内选择一个简单、实用的近似函数。这种表示单元内部解的性态的近似函数称为插值函数。

插值函数一般都采用多项式函数，主要原因是：

1）采用多项式插值函数比较容易推导单元平衡方程，特别是易于进行微分和积分运算；

2）随着多项式函数阶次的增加，可以提高有限元法的计算精度。从理论上说，无限提高多项式的阶数，可以求得系统的精确解。

2.2.1 单元模型构造方法

单元模型与单元插值函数是一一对应的，单元插值函数确定下来，就建立了相应的单元模型。

单元插值函数构造方法主要有两种：整体坐标系法和局部坐标系法。

从插值方法区分主要有两种：Lagrange插值方法和Hermite插值方法。

1）Lagrange插值方法 对于有 n 个节点的单元，如果它的节点参数中只含有场函数的节点值 u_i，$i=1,2,\cdots,n$，则单元内某一确定自由度方向的场函数 u 可插值为

$$u = \sum_{i=1}^{n} N_i u_i \tag{3.2-74}$$

式中，N_i 是插值函数，它有以下性质

$$N_i(x_j, y_j, z_j, \cdots) = \delta_{ij}$$
$$\sum_{i=1}^{n} N_i = 1 \tag{3.2-75}$$

式中，δ_{ij} 是Kroneckerδ；x_j，y_j，$z_j\cdots$是单元节点坐标值。

式（3.2-75）也反映了插值函数的一个共同特性：插值函数 N_i 在 i 节点处的值等于1，在其他节点处的值等于0，也称为插值函数的正交性。

插值函数 N_i 的和等于1，也称为插值函数的正规性。

2）Hermite插值方法 如果希望在单元间的公共节点上还保持场函数导数的连续性，则节点参数中还应包含场函数导数的节点值。这时可以采用Hermite多项式作为单元的插值函数。例如，对于只有两个节点的一维单元，形函数 N_i

采用 Hermite 多项式插值为

$$u(\xi)=\sum_{i=1}^{2}H_i^{(0)}(\xi)u_i+\sum_{i=1}^{2}H_i^{(1)}(\xi)\left(\frac{\mathrm{d}u}{\mathrm{d}\xi}\right)_i \tag{3.2-76}$$

或者

$$u(\xi)=\sum_{i=1}^{4}N_i(\xi)Q_i \tag{3.2-77}$$

式中

$$N_1=H_1^{(0)}(\xi),N_2=H_2^{(0)}(\xi)$$
$$N_3=H_1^{(1)}(\xi),N_4=H_2^{(1)}(\xi) \tag{3.2-78}$$

$$Q_1=u_1,\quad Q_2=u_2$$
$$Q_3=\left(\frac{\mathrm{d}u}{\mathrm{d}\xi}\right)_1,\quad Q_4=\left(\frac{\mathrm{d}u}{\mathrm{d}\xi}\right)_2 \tag{3.2-79}$$

Hermite 多项式具有以下性质

$$H_i^{(0)}(\xi_j)=\delta_{ij},\quad \left.\frac{\mathrm{d}H_i^{(0)}(\xi)}{\mathrm{d}\xi}\right|_{\xi_j}=0$$
$$H_i^{(1)}(\xi_j)=0,\quad \left.\frac{\mathrm{d}H_i^{(0)}(\xi)}{\mathrm{d}\xi}\right|_{\xi_j}=\delta_{ij} \tag{3.2-80}$$

式（3.2-76）在两个节点处最高保持场函数的一阶导数连续性，这种多项式称为一阶 Hermite 多项式。零阶 Hermite 多项式就退化为 Lagrange 多项式。进一步推广，在节点处保持至场函数的 n 阶导数连续性，就称为 n 阶 Hermite 多项式。

2.2.2　插值多项式收敛性条件与插值多项式选择条件

插值多项式对有限元数值解有很大的影响，当单元逐渐缩小时，如果插值多项式满足以下条件，则数值解将收敛于精确解。

1）插值多项式收敛性条件

① 在单元内，场函数必须是连续的。

② 当单元无限缩小时，插值多项式应该能够描述场函数及其各阶导数（直至最高阶导数）的均匀状态。

③ 场函数及其偏导数（直至比最高阶导数低一阶的各阶导数）在各单元边界必须连续。

采用多项式形式本身就是连续的，因此①是自动满足的。

均匀状态是场函数最基本的变化状态，插值多项式应该能够描述这种状态。类似地，当单元逐渐缩小时，场函数的各阶导数在单元内将趋于常数，因此也要求插值多项式具备这种表达能力。从固体力学角度来看，插值多项式零阶导数所描述的场函数均匀状态的物理意义就是刚体位移，一阶导数的均匀状态对应的是常应变状态。

插值多项式满足条件①和③的单元，称为协调单元。满足条件②的单元，称为完备单元。如果场函数的第 r 阶导数是连续的，则称此单元具有 c^r 连续性。

从收敛性条件来看，完备单元具有 c^r 连续性。协调单元在单元边界具有 c^{r-1} 连续性。

对于一般固体力学问题来说，协调性要求单元在变形时，相邻单元之间不应引起开裂、重叠或其他不连续现象。例如，梁、板、壳等单元，在单元边界不但要求位移是连续的，而且其一阶导数也必须是连续的。板、壳单元位移函数沿单元边界的法向导数（转角）的连续性一般比较难实现，因此出现了许多不完全满足协调性要求的“非协调单元”或“部分协调单元”，有时它们的精度也很好。

2）插值多项式选择条件　在选择插值多项式函数时，应该尽量满足以下3点：

① 插值多项式应该尽可能满足其收敛性条件；

② 由插值多项式所确定的场函数变化应该与局部坐标系的选择无关；

③ 假设的插值多项式系数的数量应该等于单元的节点数。

由收敛性条件②可知，插值多项式中必须含有常数项（刚体位移项），高阶项的次数必须依次增加，不允许有跳跃。例如，$u(x)=\alpha_0+\alpha_1x+\alpha_2x^2+\alpha_3x^3$ 就是一个正确的插值多项式，它满足收敛性条件②的要求。而 $u(x)=\alpha_0+\alpha_1x+\alpha_3x^3$ 就是错误的，$u(x)$ 的二阶导数不满足均匀状态。

由选择条件②可知，插值多项式函数在所有自由度方向上要满足各向同性性，这样就不会随局部坐标系变化而改变了。例如，$u(x,y)=\alpha_0+\alpha_1x+\alpha_2y+\alpha_3xy$ 是正确的，满足插值多项式的各向同性性条件。而 $u(x,y)=\alpha_0+\alpha_1x+\alpha_2y+\alpha_3x^2$ 就是错误的，当 x 和 y 坐标互换时，插值多项式的形式就发生了改变，是不满足各向同性性条件的。

选择条件③是为了能由单元节点值惟一确定插值多项式。例如，4节点四边形的插值多项式应该是 $u(x,y)=\alpha_0+\alpha_1x+\alpha_2y+\alpha_3xy$，其中插值多项式系数 α_i（$i=0,1,2,3$）也是4个。这样由4个节点值就可以唯一确定 α_i。

2.2.3　常用单元模型构造

下面采用整体坐标系法和局部坐标系构造常用的单元模型。构造的基本原则是依据2.2.2小节所讨论的插值多项式选择条件。

（1）整体坐标系法

整体坐标系法的基本思想是针对弹性体有限元网格建立一个统一的坐标系，每个单元的插值多项式都在这个坐标系上建立。例如，图3.2-6所示二维系统，针对弹性体有限元网格建立了一个整体坐标系 oxy，单元的插值多项式都在 oxy 坐标系上建立。

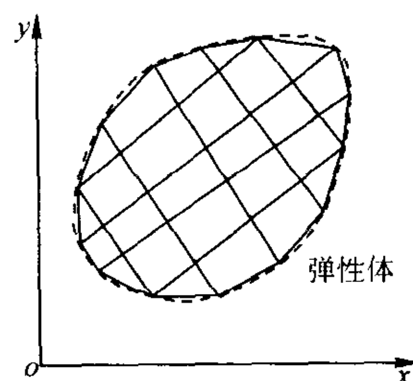


图3.2-6　弹性体网格划分与整体坐标系

1）一维2节点线单元　如图3.2-7所示，单元处于一维整体坐标系 ox 的任意位置。

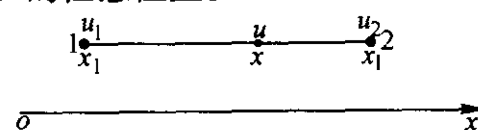


图3.2-7　一维2节点线单元

根据插值多项式选择条件，一维2节点线单元内的位移插值多项式 $u(x)$ 为

$$u(x)=\alpha_0+\alpha_1x \tag{3.2-81}$$

代入节点1和节点2的位移和坐标值，联立成方程组

$$\begin{cases}u_1=\alpha_0+\alpha_1x_1\\u_2=\alpha_0+\alpha_1x_2\end{cases} \tag{3.2-82}$$

求解方程组式（3.2-82）可求得插值系数 α_0 和 α_1

$$\alpha_0=\frac{u_1x_2-u_2x_1}{x_2-x_1}=\frac{u_1x_2-u_2x_1}{l}$$
$$\alpha_1=\frac{u_2-u_1}{x_2-x_1}=\frac{u_2-u_1}{l} \tag{3.2-83}$$

式中，l 为单元长度。

将 α_0 和 α_1 代入式（3.2-81）并按 u_1 和 u_2 整理为

$$u(x)=\frac{x_2-x}{l}u_1+\frac{x-x_1}{l}u_2 \quad (3.2\text{-}84)$$

进一步记为

$$u(x)=N_1u_1+N_2u_2=[N_1 \quad N_2]\begin{Bmatrix}u_1\\u_2\end{Bmatrix}=\boldsymbol{N}\boldsymbol{u} \quad (3.2\text{-}85)$$

式中，N_1 和 N_2 称为单元的形函数；$\boldsymbol{N}$ 称为单元的形函数矩阵；$\boldsymbol{u}$ 为单元节点位移向量。

式 (3.2-85) 就是一维 2 节点线单元的插值关系。

2）二维 3 节点三角形单元　如图 3.2-8 所示，在整体坐标系 oxy 下的二维 3 节点三角形单元。

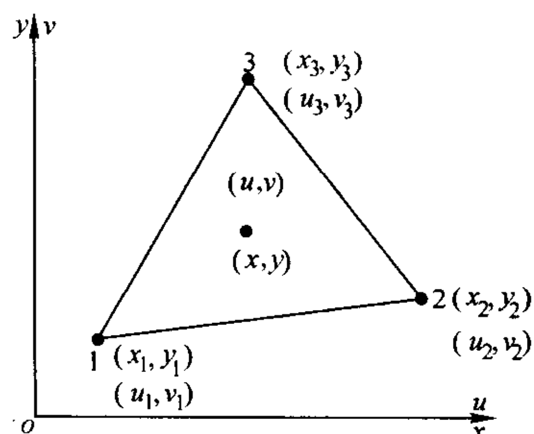


图 3.2-8 二维 3 节点三角形单元

根据插值多项式选择条件，单元内 x 自由度方向的位移插值多项式$u(x,y)$为

$$u(x,y)=\alpha_0+\alpha_1x+\alpha_2y \quad (3.2\text{-}86)$$

类似于一维 2 节点线单元的作法，可求得 x 自由度方向的插值关系

$$u(x,y)=N_1u_1+N_2u_2+N_3u_3 \quad (3.2\text{-}87)$$

式中，N_1、N_2、N_3 为单元形函数，分别为

$$\begin{aligned}N_1&=\frac{1}{2A}\ (a_1+b_1x+c_1y)\\N_2&=\frac{1}{2A}\ (a_2+b_2x+c_2y)\\N_3&=\frac{1}{2A}\ (a_3+b_3x+c_3y)\end{aligned} \quad (3.2\text{-}88)$$

$$\begin{aligned}&a_1=x_2y_3-x_3y_3;\ a_2=x_3y_1-x_1y_3;\ a_3=x_1y_2-x_2y_1\\&b_1=y_2-y_3;\ b_2=y_3-y_1;\ b_3=y_1-y_2\\&c_1=x_3-x_2;\ c_2=x_1-x_3;\ c_3=x_2-x_1\end{aligned} \quad (3.2\text{-}89)$$

A 为三角形单元面积

$$A=\frac{1}{2}(x_1y_2+x_2y_3+x_3y_1-x_1y_3-x_2y_1-x_3y_2) \quad (3.2\text{-}90)$$

同理也可得到 y 自由度方向的插值关系

$$v(x,y)=N_1v_1+N_2v_2+N_3v_3 \quad (3.2\text{-}91)$$

将式 (3.2-87) 和式 (3.2-91) 联立起来，就得到二维 3 节点三角形单元的插值关系

$$\begin{Bmatrix}u\\v\end{Bmatrix}=\begin{bmatrix}N_1&0&N_2&0&N_3&0\\0&N_1&0&N_2&0&N_3\end{bmatrix}\begin{Bmatrix}u_1\\v_1\\u_2\\v_2\\u_3\\v_3\end{Bmatrix} \quad (3.2\text{-}92)$$

或记为

$$\boldsymbol{u}=\boldsymbol{N}\boldsymbol{u}^e \quad (3.2\text{-}93)$$

3）三维 4 节点四面体单元　如图 3.2-9 所示，在整体坐标系 $oxyz$ 下的三维 4 节点四面体单元。

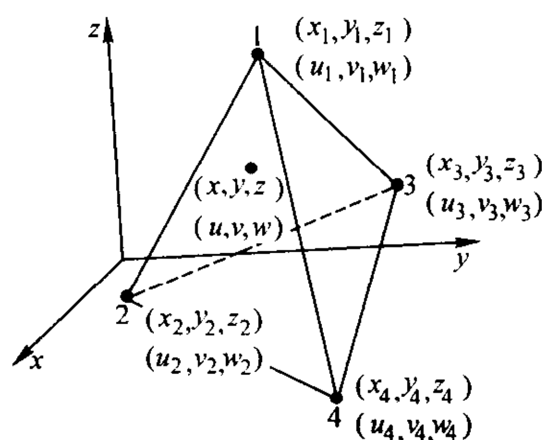


图 3.2-9 三维 4 节点四面体单元

根据插值多项式选择条件，单元内 x 自由度方向的位移插值多项式$u(x,y,z)$为

$$u(x,y,z)=\alpha_0+\alpha_1x+\alpha_2y+\alpha_4z \quad (3.2\text{-}94)$$

根据 4 个节点的位移量和坐标值，可求得 x 自由度方向的插值关系

$$u(x,y,z)=N_1u_1+N_2u_2+N_3u_3+N_4u_4 \quad (3.2\text{-}95)$$

同理也可得到 y 自由度和 z 自由度方向的插值关系分别为

$$v(x,y,z)=N_1v_1+N_2v_2+N_3v_3+N_4v_4 \quad (3.2\text{-}96)$$

$$w(x,y,z)=N_1w_1+N_2w_2+N_3w_3+N_4w_4 \quad (3.2\text{-}97)$$

式中，N_1、N_2、N_3、N_4 为单元形函数，分别为

$$N_i=\frac{1}{6V}\ (a_i+b_ix+c_iy+d_iz)\ (i=1,\ 2,\ 3,\ 4) \quad (3.2\text{-}98)$$

$$a_1=\begin{vmatrix}x_2&y_2&z_2\\x_3&y_3&z_3\\x_4&y_4&z_4\end{vmatrix};\ b_1=-\begin{vmatrix}1&y_2&z_2\\1&y_3&z_3\\1&y_4&z_4\end{vmatrix};$$

$$c_1=-\begin{vmatrix}x_2&1&z_2\\x_3&1&z_3\\x_4&1&z_4\end{vmatrix};\ d_1=\begin{vmatrix}x_2&y_2&1\\x_3&y_3&1\\x_4&y_4&1\end{vmatrix} \quad (3.2\text{-}99)$$

循环轮换脚标 1、2、3、4，相应可以得到 a_2，b_2，c_2，d_2、a_2，b_3，c_3，d_3、a_4，b_4，c_4，d_4。V 为单元体积

$$V=\frac{1}{6}\begin{vmatrix}1&x_1&y_1&z_1\\1&x_2&y_2&z_2\\1&x_3&y_3&z_3\\1&x_4&y_4&z_4\end{vmatrix} \quad (3.2\text{-}100)$$

联立式 (3.2-95)、式 (3.2-96) 和式 (3.2-97)，得到三维 4 节点四面体单元的插值关系

$$\begin{Bmatrix} u \\ v \\ w \end{Bmatrix} = \begin{bmatrix} N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 & 0 & 0 \\ 0 & N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 & 0 \\ 0 & 0 & N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 \end{bmatrix} \begin{Bmatrix} u_1 \\ v_1 \\ w_1 \\ u_2 \\ v_2 \\ w_2 \\ u_3 \\ v_3 \\ w_3 \\ u_4 \\ v_4 \\ w_4 \end{Bmatrix} \tag{3.2-101}$$

或记为

$$\boldsymbol{u} = \boldsymbol{N}\boldsymbol{u}^e \tag{3.2-102}$$

以上利用整体坐标系法求得了一维2节点线单元、二维3节点三角形单元和三维4节点四面体单元的插值关系。经过验证不难发现，3种单元的形函数都满足形函数的特性，即正交性和正规性。后面采用局部坐标系求单元的形函数时，就是利用了单元形函数的这些特性。

由插值关系式（3.2-85）、式（3.2-92）和式（3.2-101）可以发现，位移场 u 实际上是形函数 N_i 的线性组合，因此插值多项式的特性对于形函数 N_i 来说也应该同样具有。这一点在利用局部坐标系法求单元形函数时要用到。

从理论上说，整体坐标系方法可以求任意单元的插值形函数，但实际操作起来相当困难。因为前面3个单元都是线性单元，插值多项式都是线性的，非常简单，因此要求解的方程组式（3.2-82）是一个线性方程组，可以用解析法求得方程组的解。如果插值多项式中含有二次或更高阶次项，则方程组式（3.2-82）就是一个非线性方程组，一般很难得到解析解，也就不能给出形函数的解析表达式了。因此，对于高阶单元一般都采用局部坐标系法求单元的形函数。

(2) 局部坐标系法

采用局部坐标系法求单元的形函数一般需要如下过程：

1) 在单元上假设一种局部坐标系，确定局部坐标系的度量，并在单元节点上标出局部坐标值；

2) 根据插值多项式选择条件假设形函数多项式；

3) 利用单元形函数的特性（正交性）求单元的形函数。

单元局部坐标系是一种与单元形状相关的无因次坐标，也叫自然坐标系。

1) 一维单元

① 长度坐标　如图3.2-10所示，在一维单元上假设的一种自然坐标系。ox 为一维整体坐标系，x_1 和 x_2 是单元节点1和节点2在整体坐标系下的坐标值，l 是单元长度。单元上任意一点 P 到1节点的距离是 l_2，到2节点的距离是 l_1，假设单元的自然坐标为 L_1 和 L_2。

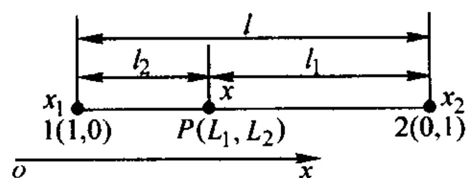


图3.2-10　一维单元长度自然坐标系

定义一维单元的自然坐标 L_1 和 L_2 分别为

$$\begin{aligned} L_1 &= \frac{l_1}{l} = \frac{x_2 - x}{l} \\ L_2 &= \frac{l_2}{l} = \frac{x - x_1}{l} \end{aligned} \tag{3.2-103}$$

自然坐标 L_1 和 L_2 是用单元长度 l、l_1 和 l_2 定义的，因此它们也叫长度坐标。

长度 L_1 和 L_2 不是相互独立的，它们存在如下关系

$$L_1 + L_2 = \frac{l_1}{l} + \frac{l_2}{l} = 1 \tag{3.2-104}$$

比较式（3.2-84）和式（3.2-103）很容易发现，一维2节点线单元的形函数可以表示为

$$\begin{aligned} N_1 &= L_1 \\ N_2 &= L_2 \end{aligned} \tag{3.2-105}$$

单元内任意一点 P 的坐标 x 可以用 x_1 和 x_2 表示为

$$x = x_1 L_1 + x_2 L_2 \tag{3.2-106}$$

式（3.2-104）和式（3.2-106）联立，得到自然坐标系与整体坐标系之间的变换关系

$$\begin{Bmatrix} 1 \\ x \end{Bmatrix} = \begin{bmatrix} 1 & 1 \\ x_1 & x_2 \end{bmatrix} \begin{Bmatrix} L_1 \\ L_2 \end{Bmatrix} \tag{3.2-107}$$

或者

$$\begin{Bmatrix} L_1 \\ L_2 \end{Bmatrix} = \frac{1}{l} \begin{bmatrix} x_2 & -1 \\ -x_1 & 1 \end{bmatrix} \begin{Bmatrix} 1 \\ x \end{Bmatrix} \tag{3.2-108}$$

② 一维二次单元　建立长度坐标系的目的是为了求一维高阶单元的形函数。下面利用长度求一维3节点单元的形函数。如图3.2-11所示，一维3节点单元，节点3位于1、2节点的中间，图中已经标出了每个节点的长度坐标。

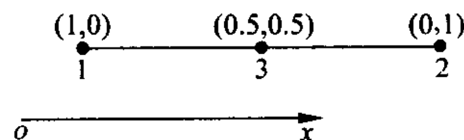


图3.2-11　一维3节点单元

根据插值多项式与形函数之间的线性组合关系，一维3节点单元的形函数多项式可以假设为

$$N_i = \alpha_1^i L_1 + \alpha_2^i L_2 + \alpha_3^i L_1 L_2 \quad (i = 1,\ 2,\ 3) \tag{3.2-109}$$

根据 N_i 的正交性，N_1 在1节点处的值等于1，在2、3节点处的值等于0，即

$$N_1 = \begin{cases} 1 & 当\ L_1 = 1,\ L_2 = 0\ 时 \\ 0 & 当\ L_1 = 0,\ L_2 = 1\ 时 \\ 0 & 当\ L_1 = 0.5,\ L_2 = 0.5\ 时 \end{cases} \tag{3.2-110}$$

由这3个条件得到方程组

$$\begin{cases} \alpha_1^1 = 1 \\ \alpha_2^1 = 0 \\ 0.5\alpha_1^1 + 0.5\alpha_2^1 + 0.25\alpha_3^1 = 0 \end{cases} \tag{3.2-111}$$

由方程组式（3.2-111）很容易确定

$$\alpha_1^1 = 1,\ \alpha_2^1 = 0,\ \alpha_3^1 = -2 \tag{3.2-112}$$

于是可以得到形函数 N_1

$$N_1 = L_1 - 2L_1 L_2 \tag{3.2-113}$$

同理可得

$$N_2 = L_2 - 2L_1 L_2 \tag{3.2-114}$$

$$N_3 = 4L_1 L_2 \tag{3.2-115}$$

这样，就得到了一维3节点单元的形函数。计算过程表明，采用长度坐标系求2次单元形函数比整体坐标系法要简单得多，而且采用这种方法还可以求更高阶单元的形函数。

由于长度坐标 L_i 本身就含有常数项和一次项，因此式(3.2-109)完全满足插值多项式选择条件要求。又由于长度坐标 L_1 和 L_2 不是相互独立的，形函数多项式的假设就会出现多种形式，只要它们满足插值多项式选择条件要求即可。例如

$$N_i = \alpha_1^i L_1 + \alpha_2^i L_2 + \alpha_3^i L_1^2 \quad (i=1,\ 2,\ 3) \tag{3.2-116}$$

$$N_i = \alpha_1^i + \alpha_2^i L_2 + \alpha_3^i L_1^2 \quad (i=1,\ 2,\ 3) \tag{3.2-117}$$

等形式都可以。实际计算过程中要考虑求解的繁易程度。

③ 一维三次单元　如图3.2-12所示，一维4节点单元，各节点长度坐标如图所示。

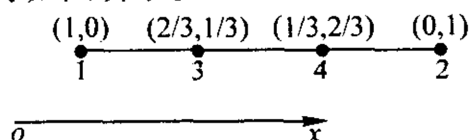


图3.2-12　一维4节点单元

一维4节点单元的形函数形式假设为

$$N_i = \alpha_1^i L_1 + \alpha_2^i L_2 + \alpha_3^i L_1 L_2 + \alpha_4^i L_1 L_2^2 \tag{3.2-118}$$

根据形函数 N_i 的正交性，可分别求得单元的形函数

$$\begin{aligned} N_1 &= L_1\left(1-\frac{9}{2}L_1L_2\right) \\ N_2 &= -\frac{9}{2}L_1L_2(1-L_1) \\ N_3 &= -\frac{9}{2}L_1L_2(1-3L_1) \\ N_4 &= 9L_1L_2\left(1-\frac{2}{3}L_1\right) \end{aligned} \tag{3.2-119}$$

④ 正规自然坐标　实际上对于一维单元来说，求单元的形函数时，最常用的是采用正规自然坐标。正规自然坐标系是一种正规化的曲线坐标系，如图3.2-13所示。坐标系原点位于单元形心 P 点处，坐标轴 r 与单元重合并指向2节点，坐标系的度量假设为 P 点为0，2节点为1，1节点为 -1。

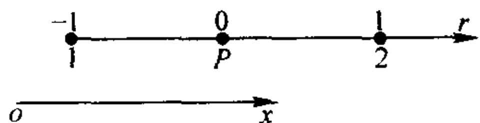


图3.2-13　一维2节点单元正规自然坐标系

⑤ 一维2节点单元　根据插值多项式与形函数之间的线性组合关系，一维2节点单元的形函数多项式可以假设为

$$N_i = \alpha_0 + \alpha_1 r \tag{3.2-120}$$

根据 N_i 的正交性，N_1 在1节点处的值等于1，在2节点处的值等于0，即

$$N_1 = \begin{cases} 1 & 当\ r=-1\ 时 \\ 0 & 当\ r=1\ 时 \end{cases} \tag{3.2-121}$$

由式(3.2-121)可求得

$$\alpha_0 = \frac{1}{2},\ \alpha_1 = -\frac{1}{2} \tag{3.2-122}$$

代入式(3.2-120)得形函数 N_1 为

$$N_1 = \frac{1}{2}(1-r) \tag{3.2-123}$$

同理可得形函数 N_2 为

$$N_2 = \frac{1}{2}(1+r) \tag{3.2-124}$$

⑥ 一维3节点单元　如图3.2-14所示，一维3节点单元及其正规自然坐标系，上面已经标出了每个节点正规自然坐标值。

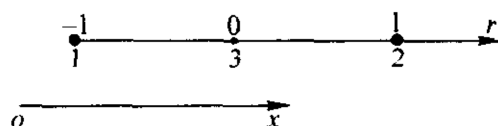


图3.2-14　一维3节点单元正规自然坐标系

一维3节点单元的形函数 N_i 的形式假设为

$$N_i = \alpha_0 + \alpha_1 r + \alpha_2 r^2 \tag{3.2-125}$$

根据 N_i 的正交性可求得单元的形函数分别为

$$\begin{aligned} N_1 &= \frac{1}{2}r(r-1) \\ N_2 &= 1-r^2 \\ N_3 &= \frac{1}{2}r(r+1) \end{aligned} \tag{3.2-126}$$

采用这种方法同样可求更高阶单元的形函数。

2) 二维单元

① 面积坐标　如图3.2-15所示，在二维三角形单元上采用单元面积建立了一种自然坐标系，也叫面积坐标系。oxy 为二维整体坐标系，$(x_i,\ y_i)$ 为节点 i 的整体坐标系的坐标值，单元内任意一点 P 将三角形单元分为3部分，面积分别为 A_1、A_2、A_3，假设三角形单元的面积坐标分别为 L_1、L_2 和 L_3。

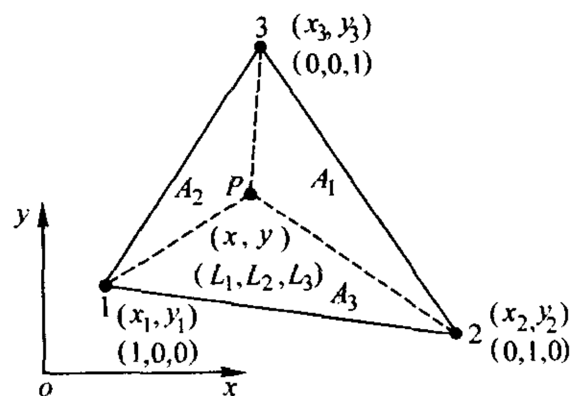


图3.2-15　二维三角形单元的面积坐标系

定义面积坐标 L_1、L_2 和 L_3 分别为

$$L_i = \frac{A_i}{A},\quad i=1,\ 2,\ 3 \tag{3.2-127}$$

式中，A 为三角形单元面积。

面积坐标 L_1、L_2 和 L_3 之间不是相互独立的，它们存在如下关系

$$L_1 + L_2 + L_3 = \frac{A_1}{A} + \frac{A_2}{A} + \frac{A_3}{A} = 1 \tag{3.2-128}$$

式(3.2-88)和式(3.2-127)经过验证，面积坐标 L_1、L_2 和 L_3 恰好是二维3节点三角形单元的形函数，即

$$N_i = L_i,\quad i=1,\ 2,\ 3 \tag{3.2-129}$$

面积坐标系与整体坐标系之间的转换关系为

$$\begin{Bmatrix} 1 \\ x \\ y \end{Bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} \begin{Bmatrix} L_1 \\ L_2 \\ L_3 \end{Bmatrix} \tag{3.2-130}$$

或者

$$\begin{Bmatrix} L_1 \\ L_2 \\ L_3 \end{Bmatrix} = \frac{1}{2A} \begin{bmatrix} x_2y_3 - x_3y_2 & y_2 - y_3 & x_3 - x_2 \\ x_3y_1 - x_1y_3 & y_3 - y_1 & x_1 - x_3 \\ x_1y_2 - x_2y_1 & y_1 - y_2 & x_2 - x_1 \end{bmatrix} \begin{Bmatrix} 1 \\ x \\ y \end{Bmatrix} \tag{3.2-131}$$

② 二维6节点三角形单元　采用整体坐标系方法很难求得二维6节点三角形单元的形函数。定义了面积坐标后，这个问题就变得很容易了。图3.2-16所示为二维6节点三角形单元，图中已经标出了每个节点的面积坐标。

二维6节点三角形单元的形函数多项式可以假设为

$$\begin{aligned} N_i = {} & \alpha_1^i L_1 + \alpha_2^i L_2 + \alpha_3^i L_3 + \\ & \alpha_4^i L_1 L_2 + \alpha_5^i L_2 L_3 + \alpha_6^i L_1 L_3 \quad (i=1,\ 2,\ 3,\ 4,\ 5,\ 6) \end{aligned} \tag{3.2-132}$$

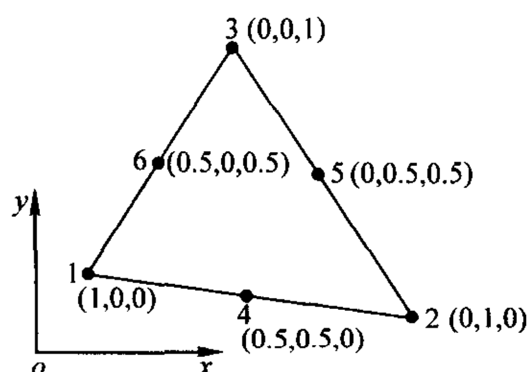


图 3.2-16 二维 6 节点三角形单元

根据 N_i 的正交性，N_1 在 1 节点处的值等于 1，在其他节点处的值等于 0，即

$$N_1=\begin{cases}1 & 当\ L_1=1,\ L_2=0,\ L_3=0\ 时\\ 0 & 当\ L_1=0,\ L_2=1,\ L_3=0\ 时\\ 0 & 当\ L_1=0,\ L_2=0,\ L_3=1\ 时\\ 0 & 当\ L_1=0.5,\ L_2=0.5,\ L_3=0\ 时\\ 0 & 当\ L_1=0,\ L_2=0.5,\ L_3=0.5\ 时\\ 0 & 当\ L_1=0.5,\ L_2=0,\ L_3=0.5\ 时\end{cases}$$

可求得二维 6 节点三角形单元的形函数分别为

$$N_i=L_i(2L_i-1),i=1,2,3$$
$$N_4=4L_1L_2$$
$$N_5=4L_2L_3$$
$$N_6=4L_3L_1$$

③ 二维 10 节点三角形单元 图 3.2-17 所示为二维 10 节点三角形单元，图中已经标出了每个节点的面积坐标。

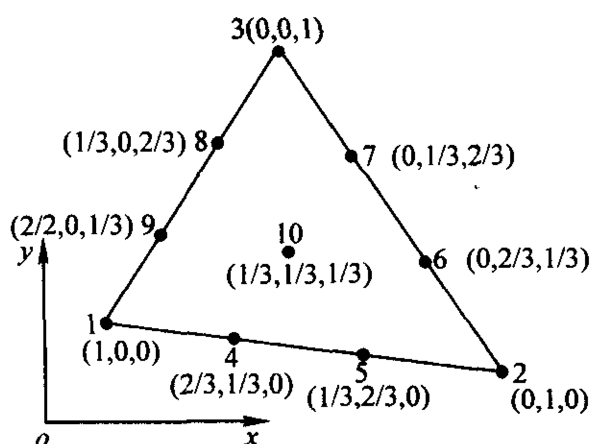


图 3.2-17 二维 10 节点三角形单元

二维 10 节点单元的形函数多项式可以假设为

$$N_i=\alpha_1^iL_1+\alpha_2^iL_2+\alpha_3^iL_3+\alpha_4^iL_1L_2+\alpha_5^iL_2L_3+\alpha_6^iL_1L_3+\alpha_7^iL_1^2L_2+\alpha_8^iL_2^2L_3+\alpha_9^iL_1L_3^2+\alpha_{10}^iL_1L_2L_3 \quad (i=1,\ 2,\ \cdots,\ 10) \qquad (3.2\text{-}133)$$

根据 N_i 的正交性，可求得二维 10 节点三角形单元的形函数为

$$N_i=\frac{1}{2}L_i(3L_i-1)(3L_i-2),i=1,2,3$$
$$N_4=\frac{9}{2}L_1L_2(3L_1-1)$$
$$N_5=\frac{9}{2}L_1L_2(3L_2-1)$$
$$N_6=\frac{9}{2}L_2L_3(3L_2-1)$$
$$N_7=\frac{9}{2}L_2L_3(3L_3-1)$$
$$N_8=\frac{9}{2}L_1L_3(3L_3-1)$$
$$N_9=\frac{9}{2}L_1L_3(3L_1-1)$$
$$N_{10}=27L_1L_2L_3$$

④ 二维 4 节点四边形单元 面积坐标是针对三角形单元假设的，它不适合四边形单元。四边形单元一般采用正规自然坐标系，类似于一维单元的正规自然坐标系。如图 3.2-18 所示，取任意四边形的对边中点连线，分别作为 r 轴和 s 轴，两轴的交点 P 作为坐标系的原点。坐标系度量定义为：原点 P 点的坐标是（0，0），四边形两对对边分别取 $r=\pm1$ 和 $s=\pm1$，这样就建立了四边形单元的一种正规自然坐标系。四边形的 4 个节点的自然坐标依次为（1，1）、（−1，1）、（−1，−1）、（1，−1）。

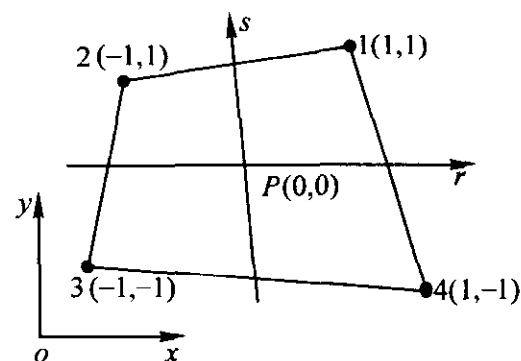


图 3.2-18 二维 4 节点四边形单元

二维 4 节点四边形单元的形函数多项式假设为

$$N_i=\alpha_0+\alpha_1r+\alpha_2s+\alpha_2rs \qquad (3.2\text{-}134)$$

根据 N_i 的正交性，可以求得 4 节点四边形单元的形函数分别为

$$N_i=\frac{1}{4}\ (1+rr_i)\ (1+ss_i),\ i=1,\ 2,\ 3,\ 4$$

⑤ 二维 8 节点四边形单元 图 3.2-19 所示为二维 8 节点四边形单元，正规自然坐标系的建立与 4 节点单元类似。由于单元的四个边可以是曲线，所以正规自然坐标系是一个曲线坐标系。单元节点的自然坐标已经在图中标出。

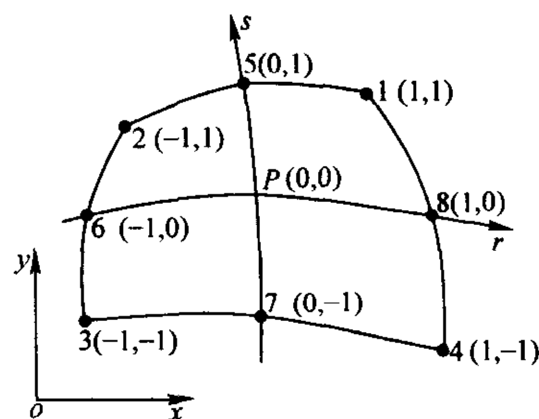


图 3.2-19 二维 8 节点四边形单元

二维 8 节点四边形单元的形函数多项式假设为

$$N_i=\alpha_0+\alpha_1r+\alpha_2s+\alpha_3rs+\alpha_4r^2+\alpha_5s^2+\alpha_6r^2s+\alpha_7rs^2 \qquad (3.2\text{-}135)$$

根据 N_i 的正交性，可以求得 8 节点四边形单元的形函数分别为

$$N_i=\frac{1}{4}(1+rr_i)(1+ss_i)(rr_i+ss_i-1),i=1,2,3,4$$
$$N_5=\frac{1}{2}(1-r^2)(1+ss_5)$$
$$N_6=\frac{1}{2}(1+rr_6)(1-s^2)$$
$$N_7=\frac{1}{2}(1-r^2)(1+ss_7)$$
$$N_8=\frac{1}{2}(1+rr_8)(1-s^2)$$

⑥ 二维 12 节点四边形单元 图 3.2-20 所示为二维 12 节点四边形单元，正规自然坐标系的建立与 8 节点单元相同。单元节点的自然坐标已经在图中标出。

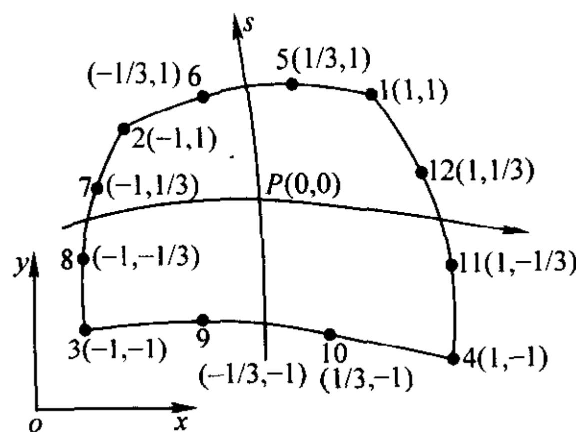


图 3.2-20　二维 12 节点四边形单元

二维 12 节点四边形单元的形函数多项式假设为

$$N_i = \alpha_0 + \alpha_1 r + \alpha_2 s + \alpha_3 rs + \alpha_4 r^2 + \alpha_5 s^2 + \alpha_6 r^2 s + \alpha_7 rs^2 + \alpha_8 r^3 + \alpha_9 s^3 + \alpha_{10} r^3 s + \alpha_{11} rs^3 \tag{3.2-136}$$

根据 N_i 的正交性，可以求得 12 节点四边形单元的形函数分别为

$$\begin{aligned} N_i &= \frac{1}{32}(1+rr_i)(1+ss_i)(9r^2+9s^2-10),\ i=1,2,3,4 \\ N_i &= \frac{9}{32}(1+ss_i)(1-r^2)(1+9rr_i),\ i=5,6,9,10 \\ N_i &= \frac{9}{32}(1+rr_i)(1-s^2)(1+9ss_i),\ i=7,8,11,12 \end{aligned} \tag{3.2-137}$$

3）三维单元

① 体积坐标　如图 3.2-21 所示，三维四面体单元内任意一点 P，则点 $243P$ 围成的体积为 V_1，点 $134P$ 围成的体积为 V_2，$142P$ 围成的体积为 V_3，$123P$ 围成的体积为 V_4。

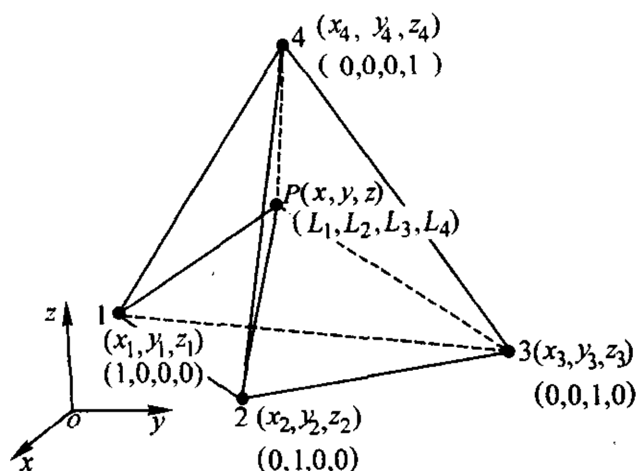


图 3.2-21　三维四面体单元的体积坐标系

定义单元的体积坐标 L_1、L_2、L_3 和 L_4 分别为

$$L_i = \frac{V_i}{V},\ i=1,\ 2,\ 3,\ 4 \tag{3.2-138}$$

式中，V 为四面体单元的体积。

体积坐标 L_1、L_2、L_3 和 L_4 之间不是相互独立的，它们存在如下关系

$$L_1 + L_2 + L_3 + L_4 = \frac{V_1}{V} + \frac{V_2}{V} + \frac{V_3}{V} + \frac{V_4}{V} = 1 \tag{3.2-139}$$

经过验证，面积坐标 L_1、L_2、L_3 和 L_4 恰好是三维 4 节点四面体单元的形函数，即

$$N_i = L_i,\ i=1,\ 2,\ 3,\ 4 \tag{3.2-140}$$

面积坐标系与整体坐标系之间的转换关系为

$$\begin{Bmatrix} 1 \\ x \\ y \\ z \end{Bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \end{bmatrix} \begin{Bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \end{Bmatrix} \tag{3.2-141}$$

或者

$$\begin{Bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \end{Bmatrix} = \frac{1}{6V} \begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{bmatrix} \begin{Bmatrix} 1 \\ x \\ y \\ z \end{Bmatrix} \tag{3.2-142}$$

式中，a_i、b_i、c_i、d_i（$i=1,\ 2,\ 3,\ 4$）按式（3.2-99）计算。

② 三维 10 节点四面体单元　定义了体积坐标后，可以利用它求四面体高阶单元的形函数。图 3.2-22 所示为三维 10 节点四面体单元，图中标出了每个节点的体积坐标。

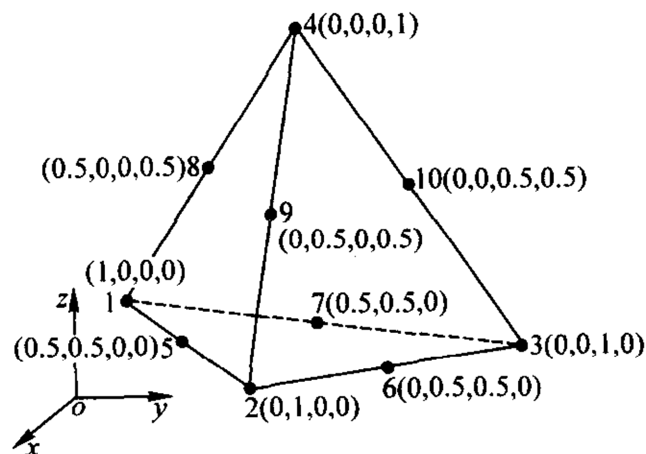


图 3.2-22　三维 10 节点四面体单元

三维 10 节点四面体单元的形函数多项式假设为

$$N_i = \alpha_1^i L_1 + \alpha_2^i L_2 + \alpha_3^i L_3 + \alpha_4^i L_4 + \alpha_5^i L_1 L_2 + \alpha_6^i L_2 L_3 + \alpha_7^i L_3 L_4 + \alpha_8^i L_4 L_1 + \alpha_9^i L_1 L_3 + \alpha_{10}^i L_2 L_4 \tag{3.2-143}$$

利用 N_i 的正交性，可以求得 10 节点四面体单元的形函数分别为

$$\begin{aligned} N_i &= L_i(2L_i - 1),\ i=1,2,3,4 \\ N_5 &= 4L_1 L_2 \\ N_6 &= 4L_2 L_3 \\ N_7 &= 4L_3 L_1 \\ N_8 &= 4L_1 L_4 \\ N_9 &= 4L_2 L_4 \\ N_{10} &= 4L_3 L_4 \end{aligned} \tag{3.2-144}$$

③ 三维 8 节点六面体单元　求三维 8 节点六面体单元的形函数要采用正规自然坐标系。如图 3.2-23 所示，每个面的对边中点连线得到 6 个交叉点，取相对面的交叉点连线分别作为 r 轴、s 轴和 t 轴，三个轴的交叉点 P 作为坐标系的原点。坐标系度量定义为：原点 P 点的坐标是（0，0，0），六面体的 3 个对面分别取 $r=\pm1$、$s=\pm1$ 和 $t=\pm1$，这样就建立了六面体单元的一种正规自然坐标系。图中已经标出每个节点的自然坐标值。

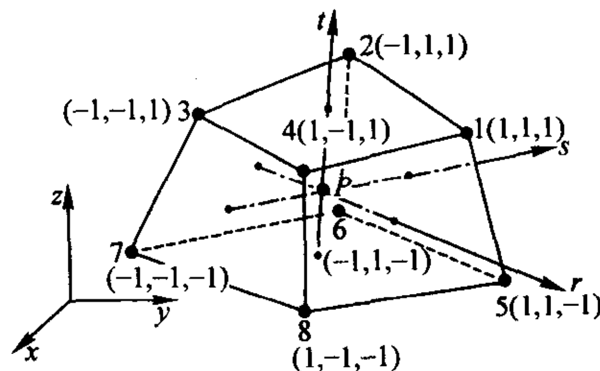


图 3.2-23　三维 8 节点六面体单元

三维 8 节点六面体单元的形函数多项式假设为

$$N_i = \alpha_0 + \alpha_1 r + \alpha_2 s + \alpha_3 t + \alpha_4 rs + \alpha_5 st + \alpha_6 tr + \alpha_7 rst \tag{3.2-145}$$

根据 N_i 的正交性，可以求得 8 节点六面体单元的形函数分别为

$$N_i=\frac{1}{8}(1+rr_i)(1+ss_i)(1+tt_i),\ i=1,2,\cdots,8 \tag{3.2-146}$$

④ 三维 20 节点六面体单元　三维 20 节点六面体单元是一个二次实体单元，它的每个面可以是曲面，如图 3.2-24 所示。正规自然坐标系的定义与 8 节点六面体单元类似，而这种单元的自然坐标系是一个 3 维曲线坐标系。节点自然坐标依次为1(1,1,1)、2(-1,1,1)、3(-1,-1,1)、4(1,-1,1)、5(1,1,-1)、6(-1,1,-1)、7(-1,-1,-1)、8(1,-1,-1)、9(0,1,1)、10(-1,0,1)、11(0,-1,1)、12(1,0,1)、13(1,1,0)、14(-1,1,0)、15(-1,-1,0)、16(1,-1,0)、17(0,1,-1)、18(-1,0,-1)、19(0,-1,-1)、20(1,0,-1)。

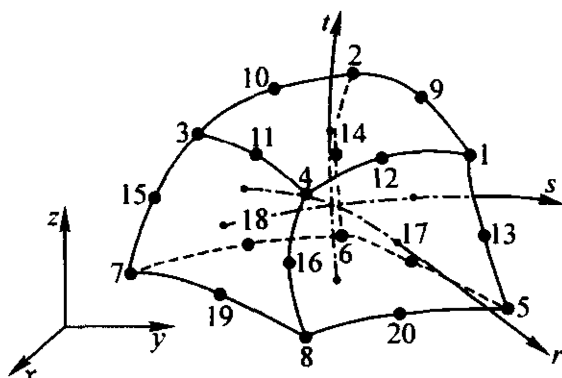


图 3.2-24　三维 20 节点六面体单元

三维 20 节点六面体单元的形函数多项式假设为

$$\begin{aligned}N_i=&\alpha_0+\alpha_1 r+\alpha_2 s+\alpha_3 t+\alpha_4 rs+\alpha_5 st+\alpha_6 tr+\\&\alpha_7 rst+\alpha_8 r^2+\alpha_9 s^2+\alpha_{10}t^2+\alpha_{11}r^2 s+\alpha_{12}s^2 t+\\&\alpha_{13}t^2 r+\alpha_{14}r^2 t+\alpha_{15}s^2 r+\alpha_{16}t^2 s+\alpha_{17}r^2 st+\\&\alpha_{18}rs^2 t+\alpha_{19}rst^2\end{aligned} \tag{3.2-147}$$

根据 N_i 的正交性，可以求得 20 节点六面体单元的形函数分别为

$$\begin{aligned}&N_i=\frac{1}{8}(1+rr_i)(1+ss_i)(1+tt_i),i=1,2,\cdots,8\\&N_i=\frac{1}{4}(1-r^2)(1+ss_i)(1+tt_i),i=9,11,17,19\\&N_i=\frac{1}{4}(1-s^2)(1+rr_i)(1+tt_i),i=10,12,18,20\\&N_i=\frac{1}{4}(1-t^2)(1+rr_i)(1+ss_i),i=13,14,15,16\end{aligned} \tag{3.2-148}$$

2.3 等参单元

一种单元模型的建立，就是确定它的形函数，单元位移的插值关系就建立起来了，即单元内任意一点的位移 u 与单元节点位移 u^e 之间的关系为

$$\boldsymbol{u}=\boldsymbol{N}\boldsymbol{u}^e \tag{3.2-149}$$

在有限元方法中，一般单元坐标的插值关系也采用与位移插值关系相同的变换关系，即单元内任意一点的坐标 x 与单元节点坐标 x^e 之间的关系为

$$\boldsymbol{x}=\boldsymbol{N}\boldsymbol{x}^e \tag{3.2-150}$$

凡是几何形状和位移场采用同阶同参数插值关系来描述的单元，称为等参单元。例如，二维 3 节点三角形单元的位移插值关系和坐标插值关系分别为

$$\begin{Bmatrix}u\\v\end{Bmatrix}=\begin{bmatrix}N_1&0&N_2&0&N_3&0\\0&N_1&0&N_2&0&N_3\end{bmatrix}\begin{Bmatrix}u_1\\v_1\\u_2\\v_2\\u_3\\v_3\end{Bmatrix} \tag{3.2-151}$$

$$\begin{Bmatrix}x\\y\end{Bmatrix}=\begin{bmatrix}N_1&0&N_2&0&N_3&0\\0&N_1&0&N_2&0&N_3\end{bmatrix}\begin{Bmatrix}x_1\\y_1\\x_2\\y_2\\x_3\\y_3\end{Bmatrix} \tag{3.2-152}$$

前面介绍的所有单元都属于等参单元。但是在描述单元的几何形状和位移场时，并不一定非采用同阶插值关系。如果几何形状插值函数的阶数低于位移场插值函数的阶数，称为次参单元。反之，如果几何形状插值函数的阶数高于位移场插值函数的阶数，称为超参单元。

3 线弹性问题几何方程

3.1 三维问题

对于三维问题，单元的几何关系为

$$\boldsymbol{\varepsilon}=\begin{Bmatrix}\varepsilon_{xx}\\\varepsilon_{yy}\\\varepsilon_{zz}\\\varepsilon_{xy}\\\varepsilon_{yz}\\\varepsilon_{zx}\end{Bmatrix}=\begin{Bmatrix}\frac{\partial u}{\partial x}\\\frac{\partial v}{\partial y}\\\frac{\partial w}{\partial z}\\\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\\\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\\\frac{\partial w}{\partial x}+\frac{\partial u}{\partial z}\end{Bmatrix}=\begin{bmatrix}\frac{\partial}{\partial x}&0&0\\0&\frac{\partial}{\partial y}&0\\0&0&\frac{\partial}{\partial z}\\\frac{\partial}{\partial y}&\frac{\partial}{\partial x}&0\\0&\frac{\partial}{\partial z}&\frac{\partial}{\partial y}\\\frac{\partial}{\partial z}&0&\frac{\partial}{\partial x}\end{bmatrix}\begin{Bmatrix}u\\v\\w\end{Bmatrix}=\boldsymbol{L}\boldsymbol{u} \tag{3.2-153}$$

3.2 二维问题

二维问题的单元几何关系可由三维情况式（3.2-153）退化而来。

3.2.1 平面应力和平面应变

$$\boldsymbol{\varepsilon}=\begin{Bmatrix}\varepsilon_{xx}\\\varepsilon_{yy}\\\varepsilon_{xy}\end{Bmatrix}=\begin{Bmatrix}\frac{\partial u}{\partial x}\\\frac{\partial v}{\partial y}\\\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\end{Bmatrix}=\begin{bmatrix}\frac{\partial}{\partial x}&0\\0&\frac{\partial}{\partial y}\\\frac{\partial}{\partial y}&\frac{\partial}{\partial x}\end{bmatrix}\begin{Bmatrix}u\\v\end{Bmatrix}=\boldsymbol{L}\boldsymbol{u} \tag{3.2-154}$$

3.2.2 轴对称

$$\boldsymbol{\varepsilon}=\begin{Bmatrix}\varepsilon_{rr}\\\varepsilon_{\theta\theta}\\\varepsilon_{zz}\\\varepsilon_{rz}\end{Bmatrix}=\begin{Bmatrix}\frac{\partial u}{\partial r}\\\frac{u}{r}\\\frac{\partial w}{\partial z}\\\frac{\partial u}{\partial z}+\frac{\partial w}{\partial r}\end{Bmatrix}=\begin{bmatrix}\frac{\partial}{\partial r}&0\\\frac{1}{r}&0\\0&\frac{\partial}{\partial z}\\\frac{\partial}{\partial z}&\frac{\partial}{\partial r}\end{bmatrix}\begin{Bmatrix}u\\w\end{Bmatrix}=\boldsymbol{L}\boldsymbol{u} \tag{3.2-155}$$

4 线弹性问题本构方程

4.1 三维问题

对于三维问题，单元的本构关系为

$$\sigma=\boldsymbol{D}^e\varepsilon \tag{3.2-156}$$

式中

$$D^e=\frac{E}{(1+\nu)(1-2\nu)}\begin{bmatrix}1-\nu & \nu & \nu & 0 & 0 & 0\\ \nu & 1-\nu & \nu & 0 & 0 & 0\\ \nu & \nu & 1-\nu & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 & 0\\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} & 0\\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2}\end{bmatrix} \quad (3.2\text{-}157)$$

式中，E 为弹性模量；ν 为泊松比。

4.2 二维问题

二维问题可由三维情况退化而来。

4.2.1 平面应力

对于平面应力情况来说，应力分量 $\sigma_z=\sigma_{xz}=\sigma_{yz}=0$，则

$$D^e=\frac{E}{1-\nu^2}\begin{bmatrix}1 & \nu & 0\\ \nu & 1 & 0\\ 0 & 0 & \frac{1-\nu}{2}\end{bmatrix} \quad (3.2\text{-}158)$$

4.2.2 平面应变

对于平面应变情况来说，应力分量 $\sigma_{xz}=\sigma_{yz}=0$，应变分量 $\varepsilon_z=0$，则

$$D^e=\frac{E}{(1+\nu)(1-2\nu)}\begin{bmatrix}1-\nu & \nu & 0\\ \nu & 1-\nu & 0\\ 0 & 0 & \frac{1-2\nu}{2}\end{bmatrix} \quad (3.2\text{-}159)$$

4.2.3 轴对称

对于轴对称情况来说，应力分量 $\sigma_{r\theta}=\sigma_{\theta z}=0$，则

$$D^e=\frac{E}{(1+\nu)(1-2\nu)}\begin{bmatrix}1-\nu & \nu & \nu & 0\\ \nu & 1-\nu & \nu & 0\\ \nu & \nu & 1-\nu & 0\\ 0 & 0 & 0 & \frac{1-2\nu}{2}\end{bmatrix} \quad (3.2\text{-}160)$$

5 单元平衡方程列式

5.1 三角形单元

三角形单元中，3 节点单元是实际工程应用最广泛的单元之一。这种单元的计算精度虽然不高，但是单元列式简单，尤其适合复杂边界的网格划分。如果单元划分密度达到一定程度的话，计算精度也能满足工程分析需要。

5.1.1 单元刚度矩阵

前面已经推导出了单元平衡方程的一般列式和 3 节点三角形单元的形函数，利用它们可以进一步推导 3 节点三角形单元的刚度矩阵。

假设 3 节点三角形单元的节点位移向量 u^e 为

$$u^e=(u_1\quad v_1\quad u_2\quad v_2\quad u_3\quad v_3)^T \quad (3.2\text{-}161)$$

则对应的单元形函数矩阵为

$$N=\begin{bmatrix}N_1 & 0 & N_2 & 0 & N_3 & 0\\ 0 & N_1 & 0 & N_2 & 0 & N_3\end{bmatrix} \quad (3.2\text{-}162)$$

式中，N_i，$i=1,2,3$ 是 3 节点三角形单元的形函数，具体形式参见式（3.2-88）~式（3.2-90）。

单元刚度矩阵的一般表达式为

$$k=\int_V B^T D^e B\mathrm{d}V \quad (3.2\text{-}163)$$

式中，B 为单元的应变矩阵；D^e 为弹性矩阵，具体表达式参见式（3.2-158）~式（3.2-160）。

1）平面应力和平面应变情况的应变矩阵 B 为

$$B=\begin{bmatrix}\frac{\partial}{\partial x} & 0\\ 0 & \frac{\partial}{\partial y}\\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x}\end{bmatrix}\begin{bmatrix}N_1 & 0 & N_2 & 0 & N_3 & 0\\ 0 & N_1 & 0 & N_2 & 0 & N_3\end{bmatrix}=\frac{1}{2a}\begin{bmatrix}b_1 & 0 & b_2 & 0 & b_3 & 0\\ 0 & c_1 & 0 & c_2 & 0 & c_3\\ c_1 & b_1 & c_2 & b_2 & c_3 & b_3\end{bmatrix} \quad (3.2\text{-}164)$$

假设平面应力情况板厚或平面应变情况的厚向长度为 t，单元面积为 a，则单元刚度矩阵 k 为

$$k=\int_V B^T D^e B\mathrm{d}v=B^T D^e B\int_V \mathrm{d}v=B^T D^e Bt\int_a \mathrm{d}a=atB^T D^e B \quad (3.2\text{-}165)$$

2）轴对称情况的应变矩阵 B 为（坐标系依然是 xy）

$$B=\begin{bmatrix}\frac{\partial}{\partial x} & 0\\ \frac{1}{x} & 0\\ 0 & \frac{\partial}{\partial y}\\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x}\end{bmatrix}\begin{bmatrix}N_1 & 0 & N_2 & 0 & N_3 & 0\\ 0 & N_1 & 0 & N_2 & 0 & N_3\end{bmatrix}=\begin{bmatrix}b_1 & 0 & b_2 & 0 & b_3 & 0\\ \frac{N_1}{x} & 0 & \frac{N_2}{x} & 0 & \frac{N_3}{x} & 0\\ 0 & c_1 & 0 & c_2 & 0 & c_3\\ c_1 & b_1 & c_2 & b_2 & c_3 & b_3\end{bmatrix} \quad (3.2\text{-}166)$$

轴对称问题的应变矩阵 B 与平面应变/应力情况不同，它不是一个常数矩阵。矩阵中的坐标 x 和形函数 N_i 都与坐标系直接相关，而且在对称轴上 $x=0$ 会造成数值奇异。为了简化计算和消除这些麻烦，实际计算过程中取单元的形心处的坐标 $\bar{x}$ 和 $\bar{y}$ 来近似，即

$$x\approx\bar{x}=\frac{1}{3}(x_1+x_2+x_3),\quad y\approx\bar{y}=\frac{1}{3}(y_1+y_2+y_3) \quad (3.2\text{-}167)$$

式中，x_i、y_i，$i=1,2,3$ 是单元的节点坐标。

这样轴对称问题的应变矩阵 B 就是与坐标无关的常数矩阵了。

假设对于轴对称情况，单元面积为 a，则

$$k=\int_v B^T D^e B\mathrm{d}v=B^T D^e B\int_v \mathrm{d}v=2\pi\bar{x}aB^T D^e B \quad (3.2\text{-}168)$$

5.1.2 载荷向量

载荷向量按式（3.2-15）计算，即

$$f=\int_a N^T P\mathrm{d}a+\int_v N^T G\mathrm{d}v \quad (3.2\text{-}169)$$

（1）平面应变和应力问题

3 节点三角形单元在单元面内和单元边界上都是线性插值，如果分布力 P 在单元内是常数，则 P 就可以按平均方法分配，在单元边界上每个节点承受 1/2 的外力，在单元面内每个节点承受 1/3 的外力。

1）分布力　如果分布力作用于单元的某一条边上或单元面内。

① 作用于 1、2 边　假设单元 1、2 边的长度为 l_{12}，则通过式（3.2-169）计算，可以得到分布力作用于 1、2 边时

单元载荷向量 f_P 为

$$f_P = \frac{l_{12}t}{2}\begin{Bmatrix} P_x \\ P_y \\ P_x \\ P_y \\ 0 \\ 0 \end{Bmatrix} \tag{3.2-170}$$

式中，P_x 和 P_y 为分布力 P 在坐标 x 和 y 方向的分量；t 为单元厚度。

② 作用于2、3边　假设单元2、3边的长度为 l_{23}，则单元载荷向量 f_P 为

$$f_P = \frac{l_{23}t}{2}\begin{Bmatrix} 0 \\ 0 \\ P_x \\ P_y \\ P_x \\ P_y \end{Bmatrix} \tag{3.2-171}$$

③ 作用于3、1边　假设单元3、1边的长度为 l_{31}，则单元载荷向量 f_P 为

$$f_P = \frac{l_{31}t}{2}\begin{Bmatrix} P_x \\ P_y \\ 0 \\ 0 \\ P_x \\ P_y \end{Bmatrix} \tag{3.2-172}$$

④ 作用于单元面内　假设单元的面积为 A，则单元载荷向量 f_P 为

$$f_P = \frac{A}{3}\begin{Bmatrix} P_x \\ P_y \\ P_x \\ P_y \\ P_x \\ P_y \end{Bmatrix} \tag{3.2-173}$$

2) 自重力　通过式（3.2-169）可以得到自重力作用的单元载荷向量 f_G 为

$$f_P = \frac{At}{3}\begin{Bmatrix} G_x \\ G_y \\ G_x \\ G_y \\ G_x \\ G_y \end{Bmatrix} \tag{3.2-174}$$

式中，G_x 和 G_y 为自重力 G 在坐标 x 和 y 方向的分量。

(2) 轴对称问题

对于轴对称问题来说，分布力和自重力的作用也是轴对称的，作用力的分配一般不是平均分配。

1) 分布力　如果分布力作用于单元的某一条边上。

① 作用于1、2边　假设单元1、2边的长度为 l_{12}，则通过式（3.2-169）计算，可以得到分布力作用于1、2边时单元载荷向量 f_P 为

$$f_P = \frac{\pi l_{12}}{3}\begin{Bmatrix} (2x_1+x_2)P_x \\ (2x_1+x_2)P_y \\ (x_1+2x_2)P_x \\ (x_1+2x_2)P_y \\ 0 \\ 0 \end{Bmatrix} \tag{3.2-175}$$

式中，P_x 和 P_y 为分布力 P 在坐标 x 和 y 方向的分量。

② 作用于2、3边　假设单元2、3边的长度为 l_{23}，则单元载荷向量 f_P 为

$$f_P = \frac{\pi l_{23}}{3}\begin{Bmatrix} 0 \\ 0 \\ 2(x_2+x_3)P_x \\ (2x_2+x_3)P_y \\ (x_2+2x_3)P_x \\ (x_2+2x_3)P_y \end{Bmatrix} \tag{3.2-176}$$

③ 作用于3、1边　假设单元3、1边的长度为 l_{31}，则单元载荷向量 f_P 为

$$f_P = \frac{\pi l_{31}}{3}\begin{Bmatrix} (2x_1+x_3)P_x \\ (2x_1+x_3)P_y \\ 0 \\ 0 \\ (x_1+2x_3)P_x \\ (x_1+2x_3)P_y \end{Bmatrix} \tag{3.2-177}$$

2) 自重力　通过式（3.2-169）可以得到自重力作用的单元载荷向量 f_G 为

$$f_G = \frac{\pi A}{6}\begin{Bmatrix} (2x_1+x_2+x_3)G_x \\ (2x_1+x_2+x_3)G_y \\ (x_1+2x_2+x_3)G_x \\ (x_1+2x_2+x_3)G_y \\ (x_1+x_2+2x_3)G_x \\ (x_1+x_2+2x_3)G_y \end{Bmatrix} \tag{3.2-178}$$

式中，G_x 和 G_y 为自重力 G 在坐标 x 和 y 方向的分量。

5.2　四边形单元

4节点四边形单元是常用的一种单元模型，它的单元平衡方程列式具有代表性。采用正规自然坐标系方法求单元形函数的一类单元，对于一维和三维问题来说，只是节点自由度维数上的差别。

5.2.1　单元刚度矩阵

假设4节点四边形单元的节点位移向量 u^e 为

$$u^e = (u_1 \quad v_1 \quad u_2 \quad v_2 \quad u_3 \quad v_3 \quad u_4 \quad v_4)^T \tag{3.2-179}$$

则对应的单元形函数矩阵为

$$N = \begin{bmatrix} N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4 & 0 \\ 0 & N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4 \end{bmatrix} \tag{3.2-180}$$

式中，N_i，$i=1, 2, 3, 4$ 是4节点四边形单元的形函数，具体形式参见式（3.2-135）。

1) 平面应力和平面应变问题

应变矩阵 B 为

$$B = \begin{bmatrix} \frac{\partial N_1}{\partial x} & 0 & \frac{\partial N_2}{\partial x} & 0 & \frac{\partial N_3}{\partial x} & 0 & \frac{\partial N_4}{\partial x} & 0 \\ 0 & \frac{\partial N_1}{\partial y} & 0 & \frac{\partial N_2}{\partial y} & 0 & \frac{\partial N_3}{\partial y} & 0 & \frac{\partial N_4}{\partial y} \\ \frac{\partial N_1}{\partial y} & \frac{\partial N_1}{\partial x} & \frac{\partial N_2}{\partial y} & \frac{\partial N_2}{\partial x} & \frac{\partial N_3}{\partial y} & \frac{\partial N_3}{\partial x} & \frac{\partial N_4}{\partial y} & \frac{\partial N_4}{\partial x} \end{bmatrix} \tag{3.2-181}$$

单元的 B 矩阵计算不像3节点三角形单元那样很容易得到一个常数矩阵。因为它的形函数是通过自然 r 和 s 坐标描述的，而不是整体坐标 x 和 y，因此它们与算子矩阵 L 作用时，不能直接偏导。

B 矩阵的具体求法如下。

把 N_i，$i=1, 2, 3, 4$ 看成是 x、y 的函数对 r 和 s 求导，根据链锁规则得

$$\begin{Bmatrix}\frac{\partial N_i}{\partial r}\\ \frac{\partial N_i}{\partial s}\end{Bmatrix}=\begin{Bmatrix}\frac{\partial N_i}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial N_i}{\partial y}\frac{\partial y}{\partial r}\\ \frac{\partial N_i}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial N_i}{\partial y}\frac{\partial y}{\partial s}\end{Bmatrix}=\begin{bmatrix}\frac{\partial x}{\partial r} & \frac{\partial y}{\partial r}\\ \frac{\partial x}{\partial s} & \frac{\partial y}{\partial s}\end{bmatrix}\begin{Bmatrix}\frac{\partial N_i}{\partial x}\\ \frac{\partial N_i}{\partial y}\end{Bmatrix}=\boldsymbol{J}\begin{Bmatrix}\frac{\partial N_i}{\partial x}\\ \frac{\partial N_i}{\partial y}\end{Bmatrix} \tag{3.2-182}$$

式中，$\boldsymbol{J}$ 称为雅可比矩阵，表示整体坐标系与自然坐标系之间的变换矩阵

$$\boldsymbol{J}=\begin{bmatrix}\frac{\partial x}{\partial r} & \frac{\partial y}{\partial r}\\ \frac{\partial x}{\partial s} & \frac{\partial y}{\partial s}\end{bmatrix} \tag{3.2-183}$$

利用等参单元关系

$$\begin{Bmatrix}x\\ y\end{Bmatrix}=\begin{bmatrix}N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4 & 0\\ 0 & N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4\end{bmatrix}\begin{Bmatrix}x_1\\ y_1\\ x_2\\ y_2\\ x_3\\ y_3\\ x_4\\ y_4\end{Bmatrix} \tag{3.2-184}$$

或者

$$x=\sum_{i=1}^{4}N_ix_i,\quad y=\sum_{i=1}^{4}N_iy_i \tag{3.2-185}$$

则雅可比矩阵 $\boldsymbol{J}$ 中的 4 个元素分别表示为

$$\frac{\partial x}{\partial r}=\sum_{i=1}^{4}\frac{\partial N_i}{\partial r}x_i;\quad \frac{\partial y}{\partial r}=\sum_{i=1}^{4}\frac{\partial N_i}{\partial r}y_i$$
$$\frac{\partial x}{\partial s}=\sum_{i=1}^{4}\frac{\partial N_i}{\partial s}x_i;\quad \frac{\partial y}{\partial s}=\sum_{i=1}^{4}\frac{\partial N_i}{\partial s}y_i \tag{3.2-186}$$

由式（3.2-182）的逆得

$$\begin{Bmatrix}b_i\\ c_i\end{Bmatrix}=\begin{Bmatrix}\frac{\partial N_i}{\partial x}\\ \frac{\partial N_i}{\partial y}\end{Bmatrix}=\boldsymbol{J}^{-1}\begin{Bmatrix}\frac{\partial N_i}{\partial r}\\ \frac{\partial N_i}{\partial s}\end{Bmatrix} \tag{3.2-187}$$

应变矩阵 $\boldsymbol{B}$ 进一步表示为

$$\boldsymbol{B}=\begin{Bmatrix}b_1 & 0 & b_2 & 0 & b_3 & 0 & b_4 & 0\\ 0 & c_1 & 0 & c_2 & 0 & c_3 & 0 & c_4\\ c_1 & b_1 & c_2 & b_2 & c_3 & b_3 & c_4 & b_4\end{Bmatrix} \tag{3.2-188}$$

单元刚度矩阵 $\boldsymbol{k}$ 为

$$\boldsymbol{k}=\int_v \boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{e}}\boldsymbol{B}\mathrm{d}v=t\int_a \boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{e}}\boldsymbol{B}\mathrm{d}a \tag{3.2-189}$$

由于 b_i 和 c_i 是关于自然坐标 r、s 的函数，所以矩阵 $\boldsymbol{B}$ 和 $\boldsymbol{k}$ 也都是关于自然坐标 r、s 的函数，这时一般要采用高斯积分方法计算单刚 $\boldsymbol{k}$。

令 $\mathrm{d}a=\mathrm{d}x\mathrm{d}y=\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s$，并将积分限取为 -1 到 $+1$，即

$$\boldsymbol{k}=t\int_{-1}^{1}\int_{-1}^{1}\boldsymbol{B}^{\mathrm{T}}D^{\mathrm{e}}B\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s \tag{3.2-190}$$

2）轴对称问题　应变矩阵 $\boldsymbol{B}$ 为（坐标系依然是 xy）

$$\boldsymbol{B}=\begin{bmatrix}\frac{\partial N_1}{\partial x} & 0 & \frac{\partial N_2}{\partial x} & 0 & \frac{\partial N_3}{\partial x} & 0 & \frac{\partial N_4}{\partial x} & 0\\ \frac{N_1}{x} & 0 & \frac{N_2}{x} & 0 & \frac{N_3}{x} & 0 & \frac{N_4}{x} & 0\\ 0 & \frac{\partial N_1}{\partial y} & 0 & \frac{\partial N_2}{\partial y} & 0 & \frac{\partial N_3}{\partial y} & 0 & \frac{\partial N_4}{\partial y}\\ \frac{\partial N_1}{\partial y} & \frac{\partial N_1}{\partial x} & \frac{\partial N_2}{\partial y} & \frac{\partial N_2}{\partial x} & \frac{\partial N_3}{\partial y} & \frac{\partial N_3}{\partial x} & \frac{\partial N_4}{\partial y} & \frac{\partial N_4}{\partial x}\end{bmatrix} \tag{3.2-191}$$

具体推导过程与平面应变/应力问题完全相同，$\boldsymbol{B}$ 的最终表达式为

$$\boldsymbol{B}=\begin{bmatrix}b_1 & 0 & b_2 & 0 & b_3 & 0 & b_4 & 0\\ \frac{N_1}{x} & 0 & \frac{N_2}{x} & 0 & \frac{N_3}{x} & 0 & \frac{N_4}{x} & 0\\ 0 & c_1 & 0 & c_2 & 0 & c_3 & 0 & c_4\\ c_1 & b_1 & c_2 & b_2 & c_3 & b_3 & c_4 & b_4\end{bmatrix} \tag{3.2-192}$$

式中，坐标 x 根据等参单元的性质，用式（3.2-185）代入即可。

单元刚度矩阵 $\boldsymbol{k}$ 为

$$\boldsymbol{k}=\int_v \boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{e}}\boldsymbol{B}\mathrm{d}v \tag{3.2-193}$$

令 $\mathrm{d}v=2\pi x\mathrm{d}a=2\pi x\mathrm{d}x\mathrm{d}y=2\pi x\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s$，并将积分限取为 -1 到 $+1$，则式（3.2-193）可以表示为

$$\boldsymbol{k}=2\pi\int_{-1}^{1}\int_{-1}^{1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{e}}\boldsymbol{B}x\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s \tag{3.2-194}$$

5.2.2　载荷向量

（1）平面应变/应力问题

平面应变/应力问题的载荷向量计算按式（3.2-15）进一步简化为

$$\boldsymbol{f}=\int_{-1}^{1}\int_{-1}^{1}\boldsymbol{N}^{\mathrm{T}}P\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s+t\int_{-1}^{1}\int_{-1}^{1}\boldsymbol{N}^{\mathrm{T}}G\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s \tag{3.2-195}$$

代入形函数矩阵 $\boldsymbol{N}$ 直接积分计算。

4 节点四边形单元在单元面内是双线性的，在单元边界上是线性插值，如果分布力 P 在单元内是常数，则 P 就可以按平均方法分配，在单元边界上每个节点承受 1/2 的外力，在单元面内每个节点承受 1/4 的外力。具体计算结果与 3 节点三角形单元类似，这里就不再列出。

（2）轴对称问题

轴对称问题的载荷向量计算按式（3.2-15）进一步简化为

$$\boldsymbol{f}=2\pi\int_{-1}^{1}\int_{-1}^{1}\boldsymbol{N}^{\mathrm{T}}P\det\boldsymbol{J}x\mathrm{d}r+2\pi\int_{-1}^{1}\int_{-1}^{1}\boldsymbol{N}^{\mathrm{T}}G\det\boldsymbol{J}x\mathrm{d}r\mathrm{d}s \tag{3.2-196}$$

代入形函数矩阵 $\boldsymbol{N}$ 直接积分计算。

5.3　空间单元

空间单元一般都要在单元上建立一个随体局部坐标系。假设单元局部坐标系的节点位移向量 $\boldsymbol{u}'^{\mathrm{e}}$ 和整体坐标系的节点位移向量 $\boldsymbol{u}^{\mathrm{e}}$ 存在如下变换关系

$$\boldsymbol{u}'^{\mathrm{e}}=\boldsymbol{\lambda}\boldsymbol{u}^{\mathrm{e}} \tag{3.2-197}$$

式中，$\boldsymbol{\lambda}$ 为坐标变换矩阵。

在随体局部坐标系下，单元的插值关系、几何关系和本构关系分别为

$$\boldsymbol{u}'=\boldsymbol{N}'\boldsymbol{u}'^{\mathrm{e}} \tag{3.2-198}$$
$$\boldsymbol{\varepsilon}'=\boldsymbol{L}'\boldsymbol{u}' \tag{3.2-199}$$
$$\boldsymbol{\sigma}'=\boldsymbol{D}^{\mathrm{e}}\boldsymbol{\varepsilon}' \tag{3.2-200}$$

在随体局部坐标系下，单元的最小势能原理为

$$\delta\Pi_P = \int_v \delta\varepsilon'^T\sigma' dv - \int_a \delta u'^T P' da - \int_v \delta u'^T G' dv = 0$$

(3.2-201)

将单元的坐标变换关系式（3.2-197）以及插值关系、几何关系和本构关系式（3.2-198）~式（3.2-200）分别代入单元的最小势能原理式（3.2-201），整理后得

$$\int_v \delta(u^e)^T \boldsymbol{\lambda}^T \boldsymbol{B}'^T \boldsymbol{D}^e \boldsymbol{B}' \boldsymbol{\lambda} u^e dv -$$

$$\int_a \delta(u^e)^T \boldsymbol{\lambda}^T \boldsymbol{N}'^T P' da - \int_v \delta(u^e)^T \boldsymbol{\lambda}^T \boldsymbol{N}'^T G' dv = 0$$

(3.2-202)

由于 $\boldsymbol{u}^e$ 的任意性，得

$$\boldsymbol{\lambda}^T(\int_v \boldsymbol{B}'^T \boldsymbol{D}^e \boldsymbol{B}' \boldsymbol{u}^e dv)\boldsymbol{\lambda} - \boldsymbol{\lambda}^T(\int_a \boldsymbol{N}'^T P' da - \int_v \boldsymbol{N}'^T G' dv) = 0$$

(3.2-203)

$$\boldsymbol{\lambda}^T \boldsymbol{k}' \boldsymbol{\lambda} \boldsymbol{u}^e = \boldsymbol{\lambda}^T \boldsymbol{f}' \quad (3.2\text{-}204)$$

式中，$\boldsymbol{k}'$ 和 $\boldsymbol{f}'$ 分别为局部坐标系下的单元刚度矩阵和载荷向量

$$\boldsymbol{k}' = \int_v \boldsymbol{B}'^T \boldsymbol{D}^e \boldsymbol{B}' \boldsymbol{u}^e dv \quad (3.2\text{-}205)$$

$$\boldsymbol{f}' = \int_a \boldsymbol{N}'^T P' da + \int_v \boldsymbol{N}'^T G' dv \quad (3.2\text{-}206)$$

因此，整体坐标系下的单元平衡方程可以表示为

$$\boldsymbol{k}\boldsymbol{u}^e = \boldsymbol{f} \quad (3.2\text{-}207)$$

式中，$\boldsymbol{k}$ 和 $\boldsymbol{f}$ 分别为整体坐标系下的单元刚度矩阵和载荷向量

$$\boldsymbol{k} = \boldsymbol{\lambda}^T \boldsymbol{k}' \boldsymbol{\lambda} \quad (3.2\text{-}208)$$

$$\boldsymbol{f} = \boldsymbol{\lambda}^T \boldsymbol{f}' \quad (3.2\text{-}209)$$

6 数值积分

2节点线单元、3节点三角形单元和4节点四面体单元3种单元的单元刚度矩阵是常数矩阵，不需要再进行数值积分运算。除了这3种单元外，一般其他单元的刚度矩阵都是积分变量的函数，要采用数值积分方法进行计算。

常用的单元面内数值积分方法主要有 Hammer 积分和 Gauss 积分。

6.1 Hammer 积分

在三角形单元和四面体单元中，自然坐标是面积坐标和体积坐标。采用这些坐标建立的单元形函数，其单元刚度矩阵的一般形式为

二维

$$I = \int_0^1\int_0^{1-L_1} F(L_1, L_2, L_3) dL_1 dL_2 \quad (3.2\text{-}210)$$

三维

$$I = \int_0^1\int_0^{1-L_1}\int_0^{1-L_1-L_2} F(L_1, L_2, L_3, L_4) dL_1 dL_2 dL_3$$

(3.2-211)

一维单元在实际有限元应用中，一般都采用正规自然坐标系法建立单元的形函数。它采用 Gauss 积分方案。

Hammer 等针对这些积分运算导出了有效的数值积分方案。三角形单元的 Hammer 积分表示为

$$\int_0^1\int_0^{1-L_1} F(L_1, L_2, L_3) dL_1 dL_2 = A_1 F(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) + B_3[F(a,a,b) + F(a,b,a) + F(b,a,a)] + C_3[F(c,c,d) + F(c,d,c) + F(d,c,c)]$$

(3.2-212)

四面体单元的 Hammer 积分表示为

$$\int_0^1\int_0^{1-L_1}\int_0^{1-L_1-L_2} F(L_1, L_2, L_3, L_4) dL_1 dL_2 dL_3 = A_1 F(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}) + B_4[F(a,b,b,b) + F(b,a,b,b) + F(b,b,a,b) + F(b,b,b,a)]$$

(3.2-213)

三角形单元和四面体单元的积分点位置、权函数和误差量级分别列于表3.2-4和表3.2-5。

表 3.2-4 三角形单元的数值积分

阶数	图形	误差	积分点	面积坐标 L_1, L_2, L_3	权系数
线性		$R = O(h^2)$	a	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$	1
二次		$R = O(h^3)$	a	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{3}$
			b	$0, \frac{1}{2}, \frac{1}{2}$	$\frac{1}{3}$
			c	$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{1}{3}$
三次		$R = O(h^4)$	a	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$	$-\frac{27}{48}$
			b	0.6,0.2,0.2	$\frac{25}{48}$
			c	0.2,0.6,0.2	
			d	0.2,0.2,0.6	

表 3.2-5 四面体单元的数值积分

阶数	图形	误差	积分点	体积坐标 L_1, L_2, L_3, L_4	权系数
线性		$R = O(h^2)$	a	$\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$	1
二次		$R = O(h^3)$	a	$\alpha, \beta, \beta, \beta$	1/4
			b	$\beta, \alpha, \beta, \beta$	1/4
			c	$\beta, \beta, \alpha, \beta$	1/4
			d	$\beta, \beta, \beta, \alpha$	1/4
				$\alpha = 0.58541020$ $\beta = 0.13819660$	
三次		$R = O(h^4)$	a	$\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$	$-\frac{4}{5}$
			b	$\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}$	$\frac{9}{20}$
			c	$\frac{1}{6}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}$	$\frac{9}{20}$
			d	$\frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{6}$	$\frac{9}{20}$
			e	$\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{3}$	$\frac{9}{20}$

6.2 Gauss 积分

采用正规自然坐标确定形函数的单元，其单元刚度矩阵的一般形式为

一维单元

$$I = \int_{-1}^1 F(r) dr \quad (3.2\text{-}214)$$

二维单元

$$I = \int_{-1}^1\int_{-1}^1 F(r,s) drds \quad (3.2\text{-}215)$$

三维单元

$$I = \int_{-1}^1\int_{-1}^1\int_{-1}^1 F(r,s,t) drdsdt \quad (3.2\text{-}216)$$

这些积分形式在积分限上与高斯积分完全一致，因此高斯积分方案被广泛应用于此类积分形式中。它们的具体数值形式为

一维单元

$$\int_{-1}^{1} F(r)\mathrm{d}r = \sum_{i=1}^{n} w_i F(r_i) \tag{3.2-217}$$

二维单元

$$\int_{-1}^{1}\int_{-1}^{1} F(r,s)\mathrm{d}r\mathrm{d}s = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j F(r_i, s_j) \tag{3.2-218}$$

三维单元

$$\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} F(r,s,t)\mathrm{d}r\mathrm{d}s\mathrm{d}t = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} w_i w_j w_k F(r_i, s_j t_k) \tag{3.2-219}$$

表 3.2-6 列出了前 6 个高斯积分点的位置和权系数。

表 3.2-6　高斯积分的积分点和权系数

积分点数（n）	积分点坐标（r_i）	权系数（w_i）
1	0.00000 00000 00000	2.00000 00000 00000
2	±0.57735 02691 89626	1.00000 00000 00000
3	±0.77459 66692 41483 0.00000 00000 00000	0.55555 55555 55556 0.88888 88888 88889
4	±0.86113 63115 94053 ±0.33998 10435 84856	0.34785 48451 47454 0.65214 51548 62546
5	±0.90617 98459 38664 ±0.53846 93101 05683 0.00000 00000 00000	0.23692 68850 56189 0.47862 86704 99366 0.56888 88888 88889
6	±0.93246 95142 03152 ±0.66120 93864 66265 ±0.23861 91860 83197	0.17132 44923 79170 0.36076 15730 48139 0.46791 39345 72691

6.3　数值积分的阶次选择

求解单元平衡方程时，绝大多数情况要采用数值积分方法，如何选择数值积分的阶次将直接影响计算精度和计算量。如果积分阶次选择不当，有时甚至会导致计算失败。

选择积分阶次的原则主要依据以下两点。

1）积分精度　积分阶次 n 与被积分多项式的阶次 m 有直接关系。一般来说，有限元积分的经验公式是

$$n = \frac{1}{2}(m+1) \tag{3.2-220}$$

从单元刚度矩阵的一般公式（3.2-13）来看，积分项有两个应变矩阵 $\boldsymbol{B}$ 相乘，因此 m 一定是偶数，因此，由式（3.2-220）计算得到的积分阶数 n 等于 0.5、1.5、2.5、…。下面按照这样的原则，对一些常用单元的积分阶次选择进行说明。

① 一维单元　一维单元的有限元应用过程中，一般都采用正规自然坐标系法得到的形函数。在单元平衡方程中雅可比矩阵中虽然也含有自然坐标，但是它只是单刚的一个系数，只对单刚中的每个元素的大小有相同的影响，不会改变单刚的特性。因此，2 节点线单元的单刚积分项是 0 次，3 节点单元是 2 次，4 节点单元是 4 次，按式（3.2-220）计算，它们的高斯积分阶次应该分别选 0.5、1.5、2.5。因此，

a）2 节点线单元只能取高斯积分点 $n=1$；

b）3 节点单元可以取 $n=1$ 或 $n=2$；

c）4 节点单元可以取 $n=2$ 或 $n=3$。

在有限元法中，把 3 节点单元取 $n=1$ 以及 4 节点单元取 $n=2$ 的积分方案称为减缩积分，而 3 节点单元取 $n=2$ 以及 4 节点单元取 $n=3$ 的积分方案称为正常积分。

实际数值结果表明，有时减缩积分方案会带来很大的计算误差，产生零能模式。而正常积分方案有时计算结果也会偏小，产生闭锁现象。造成这些现象的原因有很多，例如、单元形状、单元相对大小、单元受力状况、分析问题的类型等等。为了避免零能模式和闭锁现象的发生，一般采用减缩积分加阻尼矩阵方法。采用减缩积分方案时，对每个节点施加一个柔性弹簧，通过弹簧的阻尼增加刚度矩阵的稳定性，阻止零能模式的发生。但是弹簧的刚性系数越大，计算误差就越大，因此弹簧系数的选择也有一定的困难。

② 三角形单元　按式（3.2-220）计算，3 节点三角形单元的积分阶次 $n=0.5$，实际计算时只能取 $n=1$。这样就造成计算结果偏硬，有时会产生闭锁现象，实际有限元计算时也证明了这一点。

三角形高阶单元的积分阶次是比较精确的。例如，6 节点三角形单元的积分阶次应该取 $n=1.5$，在单元面内应该是 3 个积分点，这与表 3.2-4 中所给出的积分点数正好相符。但是，这并不意味着单元的精度就比较高，因为单元的精度是由插值多项式本身决定的。

③ 四边形单元　四边形单元与一维单元类似，按式（3.2-220）计算，4 节点、8 节点、12 节点单元的高斯积分阶次应该分别选 1.5、2.5、3.5。因此，

a）4 节点单元可以取减缩积分方案 $n=1$ 或正常积分方案 $n=2$；

b）8 节点单元可以取减缩积分方案 $n=2$ 或正常积分方案 $n=3$；

c）12 节点单元可以取减缩积分方案 $n=3$ 或正常积分方案 $n=4$。

这些单元在数值积分时，同样会象一维单元一样，出现零能模式或闭锁现象。为了避免这些现象发生，同样采用选择高斯积分方案和减缩积分加阻尼矩阵的方法进行刚度矩阵的数值积分。

但是对于 4 节点单元来说，在单元面内减缩积分是 1×1 个积分点，正常积分是 2×2 个积分点，两者相差 4 倍，因此减缩高斯积分方案和减缩积分加阻尼矩阵的方法对于 4 节点单元来改善的效果不大。针对 4 节点减缩积分特点，提出了稳定化矩阵积分方案。这种方法的基本思想是，在自然坐标系 rs 中，单元应变 ε 在点（0，0）泰勒展开，并去掉二阶小项，即

$$\varepsilon(r,s) = [B(0,0) + B_{,r}(0,0)r + B_{,s}(0,0)s]u^e \tag{3.2-221}$$

式中（　）$_{,r}$和（　）$_{,s}$表示对自然坐标的 r 和 s 的导数。将式（3.2-221）代入单元平衡方程式（3.2-12），并考虑到

$$\begin{aligned}\int_v r\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s &= 0\\ \int_v s\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s &= 0\\ \int_v rs\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s &= 0\end{aligned} \tag{3.2-222}$$

则单元刚度矩阵 $\boldsymbol{k}$ 为

$$k = k^0 + k^s \tag{3.2-223}$$

$$\boldsymbol{k}^0 = \int_{-1}^{1}\int_{-1}^{1} \boldsymbol{B}^{\mathrm{T}}(0,0)\boldsymbol{D}^e\boldsymbol{B}(0,0)\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s \tag{3.2-224}$$

$$\boldsymbol{k}^s = \frac{1}{3}\Big[\int_{-1}^{1}\int_{-1}^{1} \boldsymbol{B}_{,r}^{T}(0,0)\boldsymbol{D}^e\boldsymbol{B}_{,r}(0,0)\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s + \int_{-1}^{1}\int_{-1}^{1} \boldsymbol{B}_{,s}^{\mathrm{T}}(0,0)\boldsymbol{D}^e\boldsymbol{B}_{,s}(0,0)\det\boldsymbol{J}\mathrm{d}r\mathrm{d}s\Big] \tag{3.2-225}$$

式中，$\boldsymbol{k}^0$ 是 4 节点四边形的减缩积分单刚；$\boldsymbol{k}^s$ 称为 $\boldsymbol{k}^0$ 的稳定化单元刚度矩阵。

实际计算时，式（3.2-225）中的系数$\frac{1}{3}$可以根据情况

适当改变，例如，改成$\frac{1}{2}$或$\frac{1}{4}$等。经过这样的处理不仅消除了零能模式，还有效地提高有限元计算的精度。

④ 四面体单元 四面体单元与三角形单元相似，它们都没有减缩积分方案。4节点单元取 $n=1$，但计算结果偏硬，有时会产生闭锁现象。其他高阶单元积分阶次是比较精确的，具体情况参见表3.2-5所示。

⑤ 六面体单元 六面体单元与四边形单元相似。

a) 8节点单元可以取减缩积分方案 $n=2$ 或正常积分方案 $n=3$。

b) 20节点单元可以取减缩积分方案 $n=3$ 或正常积分方案 $n=4$。

8节点单元也可以类似4节点四边形单元采用单点积分(0，0)加稳定化矩阵的积分方案。这种方法计算单刚的效率比较高，可以降低有限元计算时间。

2) 刚度矩阵 $\boldsymbol{K}$ 是非奇异的 求解已经约束处理后的有限元平衡方程 $\boldsymbol{KU}=\boldsymbol{F}$ 时，要求方程组存在惟一解，就必须保证系数矩阵的逆 $\boldsymbol{K}^{-1}$ 存在。系数矩阵 $\boldsymbol{K}$ 非奇异的条件是满秩的，即

$$|\boldsymbol{K}|\neq 0 \tag{3.2-226}$$

如果 $\boldsymbol{K}$ 是 N 阶方阵，则要求它的秩为 N。因此，数值积分应该保证 $\boldsymbol{K}$ 是满秩的，否则将使求解失败。

关于矩阵的秩，有以下两个基本规则。

① 矩阵相乘的秩规则

如果几个矩阵相乘

$$\boldsymbol{B}=\boldsymbol{UAV} \tag{3.2-227}$$

则 $\boldsymbol{B}$ 矩阵的秩

$$\text{秩}\,\boldsymbol{B}\leqslant\min\,(\text{秩}\,\boldsymbol{U},\ \text{秩}\,\boldsymbol{A},\ \text{秩}\,\boldsymbol{V}) \tag{3.2-228}$$

即 $\boldsymbol{B}$ 矩阵的秩一定小于等于 $\boldsymbol{U}$、$\boldsymbol{A}$、$\boldsymbol{V}$ 矩阵中秩最小者。

② 矩阵相加的秩规则

如果几个矩阵相加

$$\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B} \tag{3.2-229}$$

则 $\boldsymbol{C}$ 矩阵的秩

$$\text{秩}\,\boldsymbol{C}\leqslant\text{秩}\,\boldsymbol{A}+\text{秩}\,\boldsymbol{B} \tag{3.2-230}$$

即 $\boldsymbol{C}$ 矩阵的秩一定小于等于 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 矩阵秩的和。

单元采用高斯减缩积分方案时，就会造成系数矩阵秩的不足，产生零能模式，即使单元发生变形，应变能仍为零。这种应变能为零不同于刚体运动。此时就要注意检查 $\boldsymbol{K}$ 的非奇异性条件是否得到保证。

$\boldsymbol{K}$ 的非奇异性的必要条件为

$$M\cdot n\cdot d\geqslant N \tag{3.2-231}$$

式中，M 为系统的单元数；N 为系统的自由度度；n 为高斯积分点数；d 为应变分量数，二维问题 $d=3$，轴对称问题 $d=4$，三维问题 $d=6$。

编写：柳玉起（华中科技大学）

第 3 章　板料成形数值模拟方法

板料成形数值模拟是一个非常复杂的问题。它涉及到变分原理和单元模型；板料的面内各向异性、应变强化、随动强化等材料物理模型；摩擦与润滑、拉深筋、压边力分布、坯料形状等工艺条件的模型化及其约束处理；坯料与模具间的界面接触判断与约束处理；以及所有这些模型的正确与有效实施，所有这些因素都直接影响有限元模拟的精度。

1　板壳单元

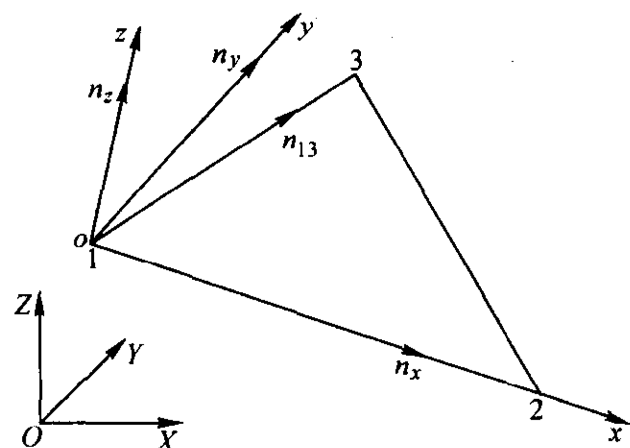
由于板料一般是比较薄的平板或壳体，冲压成形时属于平面应力状态。板很薄时可以不考虑横向剪切变形。根据板料变形的特点和实际工程计算的需要，已经构造出很多实用的单元模型。

1.1　薄膜单元

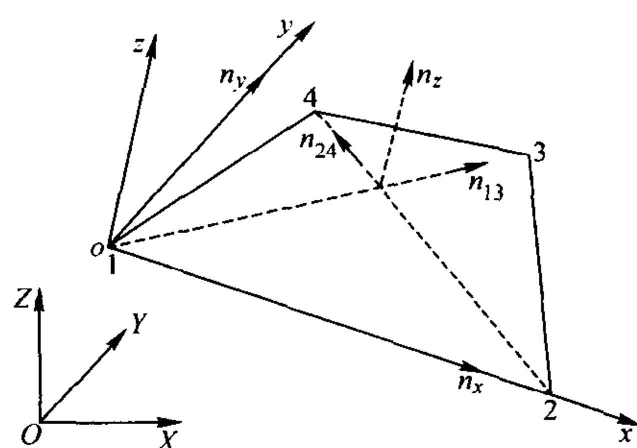
对于液压胀形、半球冲头胀形等一类问题来说，板料在变形过程中主要以拉伸和压缩变形为主，局部弯曲变形对整个成形问题不产生大的影响，这时可以采用薄膜单元。

薄膜单元是由二维三角形单元或四边形单元构造的空间板壳单元。如图 3.3-1a 和图 3.3-1b 所示，*OXYZ* 为空间整体坐标系，在每个单元上建立一个正交的随体局部坐标系 *oxyz*。图 3.3-1a 为三角形单元，局部坐标系以 1 节点为坐标原点，x 轴与单元 1、2 边重合并指向 2 节点，z 轴与单元法向量 n_z 平行。单位法向量 n_z 为

$$n_z = \frac{n_x \times n_{13}}{|n_x \times n_{13}|} \tag{3.3-1}$$



(a) 三角形单元



(b) 四边形单元

图 3.3-1　单元随体局部坐标系

因此，*oxyz* 坐标系 y 轴的单元向量 n_y 为

$$n_y = \frac{n_z \times n_x}{|n_z \times n_x|} \tag{3.3-2}$$

四边形单元的正交随体局部坐标系建立与三角形单元类似。如图 3.3-1b 所示，只是单元单位法向量 n_z 的定义与三角形单元有些差别，考虑到单元的 4 个节点可能不在一个平面内，定义单元法向量 n_z 为

$$n_z = \frac{n_{13} \times n_{24}}{|n_{13} \times n_{24}|} \tag{3.3-3}$$

则单元随体局部坐标系 *oxyz* 与空间整体坐标系 *OXYZ* 之间坐标转换矩阵 $\boldsymbol{\lambda}$ 为

$$\boldsymbol{\lambda} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} \tag{3.3-4}$$

$\boldsymbol{\lambda}$ 是一个 3×3 的正交矩阵。

当然，这里所建立的单元随体局部坐标系还有其他的定义方法，对三角形单元来说没有任何影响。但是对四边形单元来说，如果 4 个节点不在一个平面内，将对计算结果产生一定的影响。

单元的插值关系、几何关系、本构关系都是建立在 *oxyz* 坐标系上的，具体表达式可以参考本章有限元方法基础中的相关章节。

1.2　薄壳单元

薄壳单元主要以 Kirchhoff 直法线假设为理论基础，忽略横向剪切变形的影响，假设板料变形前垂直于中性层的各直线，变形后仍然保持直线并垂直于中性层。

Kirchhoff 理论单元在应用于实际分析中，通常采用协调单元和非协调单元两种列式方法。后者往往计算精度较高，会得到较好的计算结果，但它的收敛性是以通过分片试验为条件的，这就使其应用范围受到一定的限制。因此，协调板壳单元在塑性大变形研究领域仍然受到相当的重视。下面介绍一种离散 Kirchhoff 理论单元模型。

1.2.1　三角形单元

由于薄板成形过程中横向剪切变形很小，这样 Kirchhoff 直法线板壳理论假设是很合理的。而离散 Kirchhoff 理论三角形（DKT）单元以其低阶、简单、有效、位移协调等优点，已被应用于几何非线性、材料非线性的静态和动态壳体结构分析中，以及板料冲压成形过程模拟中，并都获得满意的数值结果。

如图 3.3-2 所示，DKT 单元采用横向位移 w 和中面法线转动 β_1 和 β_2 独立插值，并满足以下约束关系。

1）Kirchhoff 直法线假设作用在

① 角结点

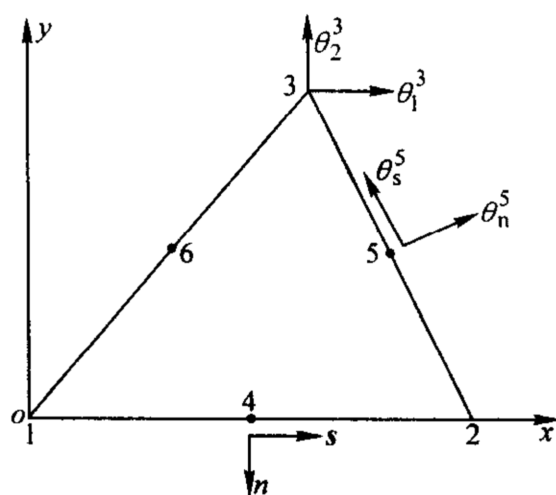
$$\gamma = \begin{bmatrix} \beta_1^i + w^i_{,x_1} \\ \beta_2^i + w^i_{,x_2} \end{bmatrix} = 0 \quad (i = 1,2,3)\text{表示角结点} \tag{3.3-5}$$

② 各边中点

$$\beta_s^k + w^k_{,s} = 0 \quad (k = 4,5,6)\text{表示边中点} \tag{3.3-6}$$

2）沿边界上 w 的变化是三次的

$$w^k_{,s} = -\frac{3}{2l_{ij}}w^i - \frac{1}{4}w^i_{,s} + \frac{3}{2l_{ij}}w^j - \frac{1}{4}w^j_{,s} \tag{3.3-7}$$



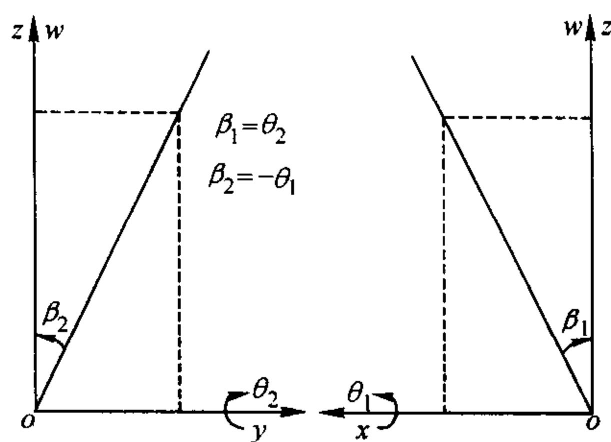


图 3.3-2 DKT 三角形单元

式中，l_{ij} 表示单元 ij 边的长度。

3）β_n 沿边界线性变化

$$\beta_n^k = \frac{1}{2}(\beta_n^i + \beta_n^j) \tag{3.3-8}$$

在单元内部 β_1 和 β_2 是二次变化，插值成

$$\beta_1 = \sum_{i=1}^{6} N_i \beta_{x_1}^i$$
$$\beta_2 = \sum_{i=1}^{6} N_i \beta_{x_2}^i \tag{3.3-9}$$

式中，N_i 为 6 节点三角形单元形函数。

根据式（3.3-5）～式（3.3-9）可将单元内任意点的 β_1 和 β_2 表示成三个角结点参数的插值形式

$$\beta_1 = \boldsymbol{H}_{x_1} \boldsymbol{u}_b^e$$
$$\beta_2 = \boldsymbol{H}_{x_2} \boldsymbol{u}_b^e \tag{3.3-10}$$

式中，$\boldsymbol{u}_b^e$ 表示弯曲结点位移向量。

1.2.2 四边形单元

离散 Kirchhoff 理论四边形（DKQ）单元是在三角形单元基础上开发的一种精度较高单元，如图 3.3-3 所示。

它采用与三角形单元相同的约束关系，即式（3.3-5）～

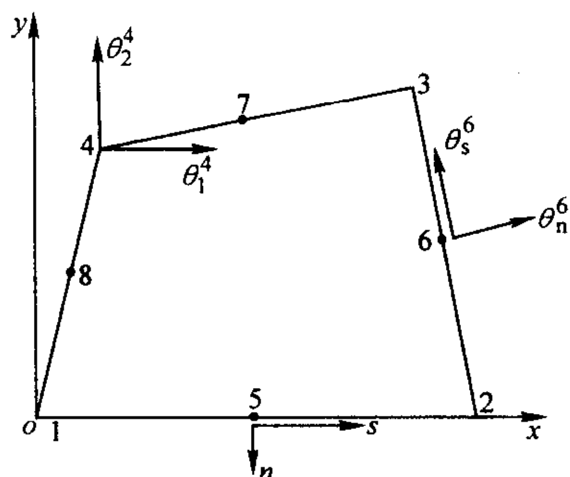


图 3.3-3 DKQ 四边形单元

式（3.3-8），在单元内部 β_1 和 β_2 也是二次变化，插值成

$$\beta_1 = \sum_{i=1}^{8} N_i \beta_{x_1}^i$$
$$\beta_2 = \sum_{i=1}^{8} N_i \beta_{x_2}^i \tag{3.3-11}$$

式中，N_i 为 8 节点四边形单元形函数。

单元内任意一点的 β_1 和 β_2 表示成四个角节点参数的插值形式

$$\beta_1 = \boldsymbol{H}_{x_1} \boldsymbol{u}_b^e$$
$$\beta_2 = \boldsymbol{H}_{x_2} \boldsymbol{u}_b^e \tag{3.3-12}$$

式中，$\boldsymbol{u}_b^e$ 表示弯曲节点位移向量；$\boldsymbol{H}_{x_1}$、$\boldsymbol{H}_{x_2}$ 的具体形式为

$$\dot{\boldsymbol{u}}_b^e = (w^1\ \theta_1^1\ \theta_2^1\ w^2\ \theta_1^2\ \theta_2^2\ w^3\ \theta_1^3\ \theta_2^3\ w^4\ \theta_1^4\ \theta_2^4)^{\mathrm{T}} \tag{3.3-13}$$

$$\boldsymbol{H}_x = [H_{x1}\ H_{x2} \cdots H_{x12}]$$
$$\boldsymbol{H}_y = [H_{y1}\ H_{y2} \cdots H_{y12}] \tag{3.3-14}$$

$$H_{x1} = N_8 \frac{3s}{2l_{23}} - N_5 \frac{3s}{2l_{23}} = 1.5\ (a_5 N_5 - a_8 N_8)$$
$$H_{x2} = -\left(N_5 \frac{3cs}{4} + N_8 \frac{3cs}{4}\right) = N_5 b_5 + N_8 b_8$$
$$H_{x3} = N_1 + \frac{2c^2 - s^2}{4} N_5 + \frac{2c^2 - s^2}{4} N_8 = N_1 - c_5 N_5 - c_8 N_8 \tag{3.3-15}$$

$$H_{y1} = N_5 \frac{3c}{2l_{12}} - N_8 \frac{3c}{2l_{41}} = 1.5\ (d_5 N_5 - d_8 N_8)$$
$$H_{y2} = -N_1 - \frac{2s^2 - c^2}{4} N_5 + \frac{2s^2 - c^2}{4} N_8 = -N_1 + e_2 N_5 + e_8 N_8$$
$$H_{y3} = \frac{3cs}{4} N_5 + \frac{3cs}{4} N_8 = -H_{x2} \tag{3.3-16}$$

函数 H_{x4}，H_{x5}，H_{x6}，H_{y4}，H_{y5}，H_{y6} 可以在上式分别用 6，5，2 来替代 5，8，1 获得，函数 H_{x7}，H_{x8}，H_{x9}，H_{y7}，H_{y8}，H_{y9} 可以在上式分别用 7，6，3 来替代 6，5，2 获得，函数 H_{x10}，H_{x11}，H_{x12}，H_{y10}，H_{y11}，H_{y12} 可以在上式分别用 8，7，4 来替代 7，6，3 获得。

$$a_k = -x_{ij}/l_{ij}^2$$
$$b_k = \frac{3}{4} x_{ij} y_{ij} / l_{ij}^2$$
$$c_k = \left(\frac{1}{4} x_{ij}^2 - \frac{1}{2} y_{ij}^2\right) / l_{ij}^2$$
$$d_k = -y_{ij}/l_{ij}^2$$
$$e_k = \left(\frac{1}{4} y_{ij}^2 - \frac{1}{2} x_{ij}^2\right) / l_{ij}^2$$
$$l_{ij}^2 = (x_{ij}^2 + y_{ij}^2) \tag{3.3-17}$$

式中 $k=5$，6，7，8 分别是 $ij=12$，23，34，41 的边中点。

1.2.3 空间薄壳单元

将薄膜单元与弯曲单元组合建立空间薄壳单元，模拟大变形大应变板材冲压成形问题。如图 3.3-1 所示，在 $oxyz$ 坐

标系下考虑 Kirchhoff 直法线假设，单元内任意点 p（x，y，z）的速度可假定为

$$\begin{aligned} \dot{\vec{u}}(x,y,z) &= \dot{u}(x,y) + z\dot{\beta}(x,y) \\ \dot{\vec{w}}(x,y,z) &= \dot{w}(x,y) \end{aligned} \tag{3.3-18}$$

式中，$\dot{u}$ 和 $\dot{w}$ 分别表示 p 点在中面上相对应点（x，y，0）沿 x、y 和 z 方向的速度；$\dot{\beta}$ 表示单元法线转角速度。由式（3.3-18）可得 p 点处的 Cauchy 应变率 $\dot{\varepsilon}$ 和速度梯度 $\dot{q}$ 分别表示为

$$\begin{aligned} \dot{\varepsilon} &= \dot{\varepsilon}^{m} + \dot{\varepsilon}^{b} \\ \dot{q} &= \dot{q}^{m} + \dot{q}^{b} \end{aligned} \tag{3.3-19}$$

式中，Cauchy 薄膜应变率 $\dot{\varepsilon}^m$ 和速度梯度 $\dot{q}^m$ 分别为

$$\begin{aligned} \dot{\varepsilon}^{m} &= \boldsymbol{L}\dot{u} \\ \dot{q}^{m} &= \boldsymbol{L}_{v}\dot{u} \end{aligned} \tag{3.3-20}$$

算子 $\boldsymbol{L}$ 和 $\boldsymbol{L}_v$ 分别为

$$\boldsymbol{L} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{bmatrix} \tag{3.3-21}$$

$$\boldsymbol{L}_v = \begin{bmatrix} \frac{\partial}{\partial x} & 0 \\ \frac{\partial}{\partial y} & 0 \\ 0 & \frac{\partial}{\partial x} \\ 0 & \frac{\partial}{\partial y} \end{bmatrix} \tag{3.3-22}$$

Cauchy 弯曲应变率 $\dot{\varepsilon}^b$ 和速度梯度 $\dot{q}^b$ 分别为

$$\begin{aligned} \dot{\varepsilon}^{b} &= z\boldsymbol{L}\dot{\beta} \\ \dot{q}^{b} &= z\boldsymbol{L}_{v}\dot{\beta} \end{aligned} \tag{3.3-23}$$

1.3　中厚壳单元

Mindlin 板壳单元考虑了横向剪切变形的影响，可用于分析较厚的壳体。这种单元可以比较好地模拟横向弯曲效应对冲压成形过程中板料起皱的影响，以及更精确地预测回弹问题，充分体现剪切应力对板材成形力学行为的影响。

目前比较实用的 Mindlin 板壳单元主要有：Hughes - Liu 壳单元，Belytschko - Lin - Tsay（BT）壳单元，Belyschko - Wong - Chiang（BWC）壳单元。这 3 种单元在 LS - DYNA 等动力显式软件中是非常重要的，板料成形模拟时主要采用这 3 种壳单元。

1.3.1　Hughes - Liu 壳单元

1）几何关系　图 3.3-4 所示为 4 节点 Hughes - Liu 壳单元。$OXYZ$ 为空间固定直角坐标系，e_1、e_2、e_3 分别为 X、Y、Z 轴的单位向量。

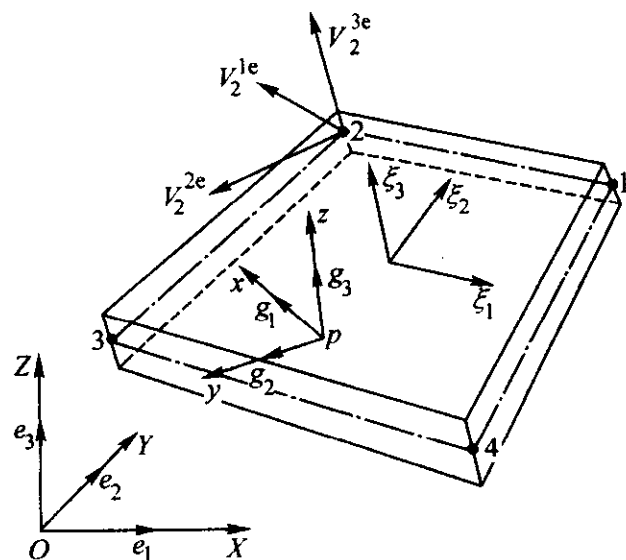


图 3.3-4　Hughes - Liu 壳单元

单元由上、下两个曲面及周边以壳体厚度方向的直线为母线的曲面所围成。ξ_i（$i=1$，2，3）为单元中面上的自然局部坐标系（$-1\leqslant\xi_i\leqslant1$）。于是，壳单元内任一点 p 的空间位置可表示为

$$X = \sum_{i=1}^{4} N_i X_i^{e} + \frac{\xi_3}{2}\sum_{i=1}^{4} N_i h_i^{e} V_i^{3e} \tag{3.3-24}$$

式中，ξ_3 为厚向自然坐标；N_i 为 4 节点四边形单元形函数；h_i^e 为单元节点厚度；V_i^{3e}（$i=1$，2，3，4）为单元节点中面的单位法向量；X_i^e（$i=1$，2，3，4）为单元节点中面坐标，可以由上、下曲面相应的节点坐标 X_i^{et} 和 X_i^{eb} 表示为

$$X_i^{e} = \frac{1}{2}(X_i^{et} + X_i^{eb}) \tag{3.3-25}$$

假设单位向量 V_i^{1e} 和 V_i^{2e} 与 V_i^{3e}（$i=1$，2，3，4）组成节点 i 的正交坐标系，则有

$$V_i^{1e} = \frac{e_1 \times V^{3e}}{|e_1 \times V_i^{3e}|};\quad V_i^{2e} = \frac{V^{3e} \times V^{1e}}{|V_i^{3e} \times V_i^{1e}|} \tag{3.3-26}$$

若 V_i^{3e} 与 e_1 平行，则式（3.3-26）中用 e_2 代替 e_1。

假定壳体的中面法线变形之后仍保持为直线，则 p 点的位移 u 可由中面对应节点沿 X_i^e 方向的 3 个线位移分量 u_i^e 以及绕节点 V_i^{1e} 和 V_i^{2e} 的 2 个角位移 α^e 表示为

$$u = \sum_{i=1}^{4} N_i u_i^{e} + \frac{\xi_3}{2}\sum_{i=1}^{4} N_i h_i^{e}(V_i^{1e}\alpha_{1i}^{e} - V_i^{2e}\alpha_{2i}^{e}) \tag{3.3-27}$$

式中，α_{1i}^e 和 α_{2i}^e 是角位移 α^e 的两个分量。

2）局部坐标系　为描述单元内任一点的应力、应变状态，在 p 点处再引入一个随体正交坐标系 $oxyz$，如图 3.3-4 所示。g_i（$i=1$，2，3）为随体坐标系 $oxyz$ 的单位向量。g_3 设定为 p 处 ξ_3 等于常数的曲面法向，则 g_1 和 g_2 与曲面相切。ξ_3 等于常数的曲面由二组曲线族 ξ_1 和 ξ_2 组成。ξ_i（$i=1$，2）的切向量 S 和 T 分别为

$$S = \frac{\partial X}{\partial \xi_1};\quad T = \frac{\partial X}{\partial \xi_2} \tag{3.3-28}$$

因此

$$g_3 = \frac{S \times T}{|S \times T|};\quad g_1 = \frac{e_1 \times g_3}{|e_1 \times g_3|};\quad g_2 = \frac{g_3 \times g_1}{|g_3 \times g_1|} \tag{3.3-29}$$

若 g_3 与 e_1 平行，则用 e_2 代替 e_1。

因此，$OXYZ$ 与 $oxyz$ 两坐标系间的转换关系为

$$x = \boldsymbol{\lambda} X \tag{3.3-30}$$

式中，$\boldsymbol{\lambda}$ 为坐标转换矩阵，具体表达式为

$$\boldsymbol{\lambda} = \begin{bmatrix} e_1 g_1 & e_1 g_2 & e_1 g_3 \\ e_2 g_1 & e_2 g_2 & e_2 g_3 \\ e_3 g_1 & e_3 g_2 & e_3 g_3 \end{bmatrix} \tag{3.3-31}$$

1.3.2　BT 壳单元

BT 壳单元是最简单的一种组合壳单元，膜变形、弯曲变形和横向剪切变形都是由双线性插值构成的 4 节点四边形壳单元。

1）随体局部坐标系　如图 3.3-5 所示，$OXYZ$ 为空间固定直角坐标系，$oxyz$ 为假设的单元随体局部坐标系，具体定义方法如下。单元法向量 e_3 由单元对角线向量 n_{13} 和 n_{24} 定义

$$e_3 = \frac{n_{13} \times n_{24}}{|n_{13} \times n_{24}|} \tag{3.3-32}$$

局部坐标系的 x 轴建立在单元 1、2 节点边上是最方便的，有利于有限元列式的推导。但是如果单元 4 个节点不在一个平面内就会产生比较大的误差。为了减少这种误差，定义 x 轴与 1、2 节点边近似平行。x 轴的单位向量 e_1 为

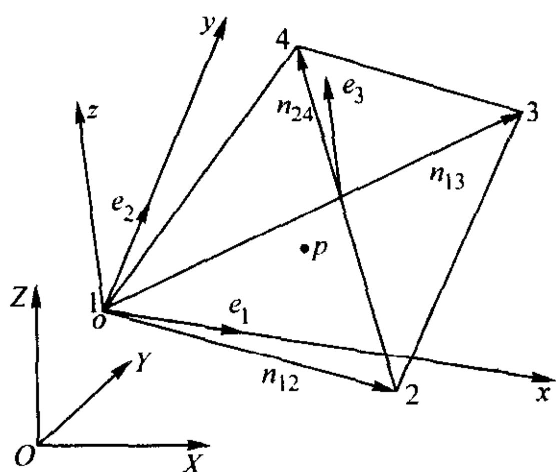


图 3.3-5 随体局部坐标系

$$e_1 = \frac{n_{12} - (n_{12} e_3) e_3}{|n_{12} - (n_{12} e_3) e_3|} \quad (3.3\text{-}33)$$

如果单元4个节点在一个平面内，则 e_1 与 n_{12}平行。y 轴的单位向量 e_2 为

$$e_2 = \frac{e_3 e_1}{|e_3 e_1|} \quad (3.3\text{-}34)$$

因此，$OXYZ$ 与 $oxyz$ 两坐标系间的转换关系为

$$x = \lambda X \quad (3.3\text{-}35)$$

式中，λ 为坐标转换矩阵，具体表达式为

$$\lambda = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \quad (3.3\text{-}36)$$

2）几何关系　在 $oxyz$ 坐标系下考虑 Mindlin 板壳单元直线假设，单元内任意点 p（x，y，z）的速度可假定为

$$\begin{aligned} \dot{\bar{u}}(x,y,z) &= \dot{u}(x,y) + z\dot{\beta}(x,y) \\ \dot{\bar{w}}(x,y,z) &= \dot{w}(x,y) \end{aligned} \quad (3.3\text{-}37)$$

式中，$\dot{u}$ 和 $\dot{w}$ 分别表示 p 点在中面上相对应点（x，y，0）沿 x、y 和 z 方向的速度；$\dot{\beta}$ 表示单元法线转角速度。由式(3.3-37) 可得 p 点处的 Cauchy 应变率 $\dot{\varepsilon}$ 和速度梯度 $\dot{q}$ 分别表示为

$$\begin{aligned} \dot{\varepsilon} &= \dot{\varepsilon}^m + \dot{\varepsilon}^b + \dot{\varepsilon}^r \\ \dot{q} &= \dot{q}^m + \dot{q}^b \end{aligned} \quad (3.3\text{-}38)$$

式中，$\dot{\varepsilon}^m$ 为 Cauchy 面内薄膜应变率；$\dot{\varepsilon}^b$ 为面内弯曲应变率；$\dot{\varepsilon}^r$ 为横向剪切应变率。式 (3.3-38) 的分量表达式为

$$\begin{Bmatrix} \dot{\varepsilon}_{xx} \\ \dot{\varepsilon}_{yy} \\ \dot{\varepsilon}_{xy} \\ \dot{\varepsilon}_{yz} \\ \dot{\varepsilon}_{zx} \end{Bmatrix} = \begin{Bmatrix} \dot{\varepsilon}^m_{xx} \\ \dot{\varepsilon}^m_{yy} \\ \dot{\varepsilon}^m_{xy} \\ 0 \\ 0 \end{Bmatrix} + \begin{Bmatrix} \dot{\varepsilon}^b_{xx} \\ \dot{\varepsilon}^b_{yy} \\ \dot{\varepsilon}^b_{xy} \\ 0 \\ 0 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ 0 \\ \dot{\varepsilon}^y_{yz} \\ \dot{\varepsilon}^r_{zx} \end{Bmatrix} \quad (3.3\text{-}39)$$

$$\begin{aligned} \dot{\varepsilon}^m &= L\dot{u} \\ \dot{\varepsilon}^b &= -zL\dot{\theta} \\ \dot{\varepsilon}^r &= -L_s\dot{r} \end{aligned} \quad (3.3\text{-}40)$$

$$\begin{aligned} \dot{u} &= N\dot{u}^e \\ \dot{\theta} &= N\dot{\theta}^e \\ \dot{r} &= N_b\dot{\beta}^e \end{aligned} \quad (3.3\text{-}41)$$

式中，N 为线位移速度和弯曲角速度形函数矩阵；$\dot{u}^e$ 为单元节点面内线位移速度向量；$\dot{\theta}^e$ 为单元节点角位移速度向量；L_s 为剪切应变算子，具体表达式为

$$L_s = \begin{bmatrix} 1 & 0 & -\dfrac{\partial}{\partial x} \\ 0 & 1 & -\dfrac{\partial}{\partial y} \end{bmatrix} \quad (3.3\text{-}42)$$

$$\begin{aligned} \dot{u}^e &= (\dot{u}_1\ \dot{v}_1\ \dot{u}_2\ \dot{v}_2\ \dot{u}_3\ \dot{v}_3\ \dot{u}_4\ \dot{v}_4)^T \\ \dot{\theta}^e &= (\theta_{1x}\,\theta_{1y}\,\theta_{2x}\,\theta_{2y}\,\theta_{3x}\,\theta_{3y})^T \\ \dot{\beta}^e &= (\dot{w}_1\ \theta_{1x}\,\theta_{1y}\ \dot{w}_2\ \theta_{2x}\,\theta_{2y}\ \dot{w}_3\ \theta_{3x}\,\theta_{3y}\ \dot{w}_4\ \theta_{4x}\,\theta_{4y})^T \end{aligned} \quad (3.3\text{-}43)$$

$$N = \begin{bmatrix} N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4 & 0 \\ 0 & N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4 \end{bmatrix}$$

$$N_b = \begin{bmatrix} N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 & 0 & 0 \\ 0 & N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 & 0 \\ 0 & 0 & N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 \end{bmatrix} \quad (3.3\text{-}44)$$

1.3.3 BWC 壳单元

BT壳单元是双线性插值，当单元挠曲比较明显时（如图3.3-6所示），单元的性态变得比较差。为了解决这个问题，又提出了 Belyschko - Wong - Chiang（BWC）单元，在单元几何关系中增加一个修正项，耦合了曲率项，并在剪切计算中附加了一个节点映射项，改善单元性态的映射因子。这些修正对于单元挠曲比较明显时，能有效改善有限元模拟精度。

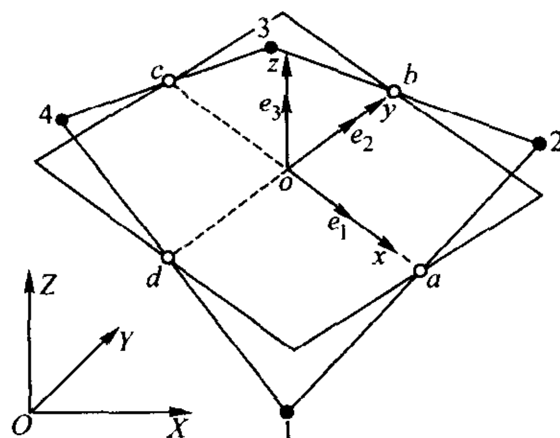


图 3.3-6 挠曲的4节点四边形单元及随体局部坐标系

1）随体局部坐标系　图 3.3-6 所示为一个挠曲的4节点四边形单元。$OXYZ$ 为空间固定直角坐标系，$oxyz$ 为假设的单元随体局部坐标系，具体定义方法如下。a、b、c、d 分别位于单元的各边中点，则单元法向量 e_3 定义为

$$e_3 = \frac{n_{ca} n_{db}}{|n_{ca} n_{db}|} \quad (3.3\text{-}45)$$

$$e_1 = \frac{n_{ca}}{|n_{ca}|} \quad (3.3\text{-}46)$$

$$e_2 = \frac{e_3 e_1}{|e_3 e_1|} \quad (3.3\text{-}47)$$

2）运动方程　如图 3.3-7 所示，考虑单元节点坐标 X_i^e（$i=1$，2，3，4）对应的上下表面坐标 X_i^{top} 和 X_i^{bot}，节点单位法向量 P_i（$i=1$，2，3，4）。

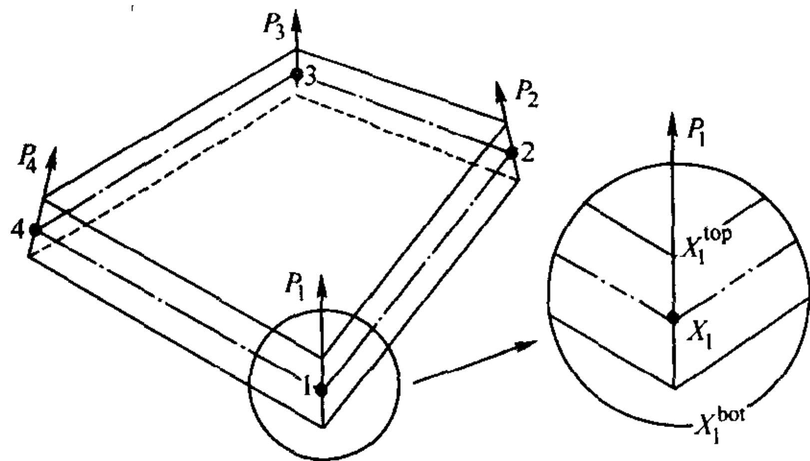


图 3.3-7 单元的几何形状

则单元内任意一点的坐标 X 为

$$X = \sum_{i=1}^{4} N_i \left[\frac{1}{2}(X_i^{top} + X_i^{bot}) + \frac{\xi_3}{2}(X_i^{top} - X_i^{bot}) \right] \quad (3.3\text{-}48)$$

式中，N_i 为4节点四边形单元形函数。

每个节点的法向量 P_i（$i=1$，2，3，4）近似定义为

$$P_i = (X_i^{top} - X_i^{bot})/h \quad (3.3\text{-}49)$$

式中，h 为单元厚度，假设单元4个节点的厚度相同。则中面节点坐标 X_i^e（$i=1, 2, 3, 4$）为

$$X_i^e = \frac{1}{2}(X_i^{top} + X_i^{bot}) \tag{3.3-50}$$

因此

$$X = \sum_{i=1}^{4} N_i\left[X_i^e + \frac{\xi_3 h}{2} P_i\right] \tag{3.3-51}$$

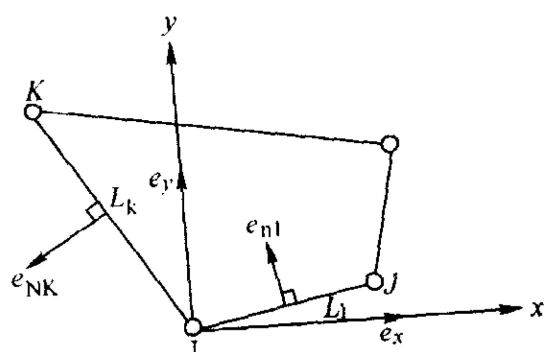
单元中性面内坐标 X^m 的等参插值为

$$X^m = \sum_{i=1}^{4} N_i X_i^e \tag{3.3-52}$$

3）剪切效应　切应变按下式由节点映射来计算

$$\bar{\theta}_n^I = \frac{1}{2}(\theta_{nI}^I + \theta_{nJ}^I) + \frac{1}{L^{IJ}}(\dot{u}_{zJ} - \dot{u}_{zI}) \tag{3.3-53}$$

式中，上角标 I 指第 I 个边，下角标 n 指与第 I 边正交的法向量，如图 3.3-8 所示。



节点和边的号码

I	J	K
1	2	4
2	3	1
3	4	2
4	1	3

图 3.3-8　单元的节点、边和边的法向单位向量

横向切应变 γ_{xz} 和 γ_{yz} 为

$$\gamma_{xz} = -\sum_{i=1}^{4} N_i \bar{\theta}_{yi}$$
$$\gamma_{yz} = -\sum_{i=1}^{4} N_i \bar{\theta}_{xi} \tag{3.3-54}$$

可见，横向剪切并不依赖于 $\dot{u}_z$，因为经过映射 $\dot{u}_z$ 消去了。因此

$$\bar{\theta}_{xi} = (e_n^I e_x)\,\bar{\theta}_n^I + (e_n^K e_x)\,\bar{\theta}_n^K$$
$$\bar{\theta}_{yi} = (e_n^I e_y)\,\bar{\theta}_n^I + (e_n^K e_y)\,\bar{\theta}_n^K \tag{3.3-55}$$

式中，e_x、e_y 为图 3.3-8 中定义的单位向量。

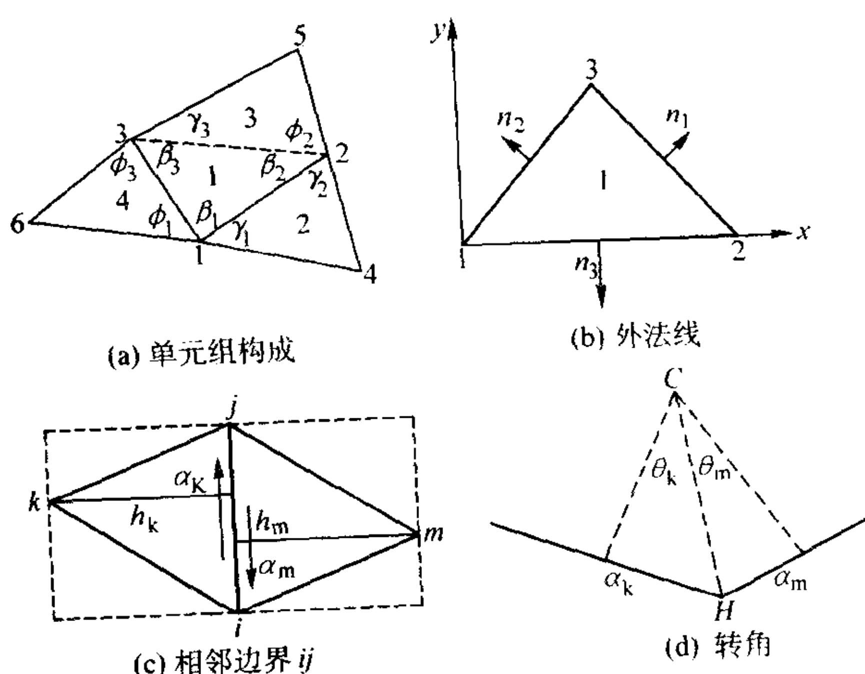
1.4　等效弯曲单元

等效弯曲单元基于具有相邻边界的单元在其相邻边界垂直方向上的曲率保持连续的假设，提出了一种改进的以面外位移来考虑弯曲效应的三角形单元组模型。这种单元能够准确地模拟板材冲压成形过程中弯曲变形引起的起皱现象，计算精度大大提高，同时还具有节点自由度低，计算工作量小等优点。

1.4.1　几何方程

如图 3.3-9 所示，考虑单元组（1，2，3，4）（图 3.3-9a 所示）各节点面外位移引起的单元 1 沿边界外法线 n_1、n_2、n_3 方向的弯曲应变（图 3.3-9b 所示）。首先假设两个有相邻边界的单元初始共面（图 3.3-9c 所示），在垂直公共边界方向上保持常曲率（图 3.3-9d 所示）。令 $CH=R$，则有

$$\theta_k = h_k/2R,\quad \theta_m = h_m/2R \tag{3.3-56}$$



(a) 单元组构成　　(b) 外法线

(c) 相邻边界 ij　　(d) 转角

图 3.3-9　等效弯曲三角形单元

令 α_i 表示单元转动角增量，由几何关系式可得

$$\theta_k + \theta_m = \alpha_k + \alpha_m \tag{3.3-57}$$

由式（3.3-56）、式（3.3-57），可以推出

$$\left.\begin{aligned}\theta_k &= h_k(\alpha_k + \alpha_m)/(h_k + h_m)\\ \theta_m &= h_m(\alpha_k + \alpha_m)/(h_k + h_m)\end{aligned}\right\} \tag{3.3-58}$$

垂直于相邻边界，距离中面厚度为 z 的弯曲应变由下式给出

$$\varepsilon_{nk} = -z/R = -2z\theta_k/h_k = -2z(\alpha_k + \alpha_m)/(h_k + h_m) \tag{3.3-59}$$

则有

$$\boldsymbol{\varepsilon}_n = \begin{Bmatrix}\varepsilon_{n1}\\ \varepsilon_{n2}\\ \varepsilon_{n3}\end{Bmatrix} = zH\boldsymbol{\alpha} \tag{3.3-60}$$

式中

$$\boldsymbol{\alpha} = \{\alpha_1\quad \alpha_2\quad \alpha_3\quad \alpha_4\quad \alpha_5\quad \alpha_6\}^T \tag{3.3-61}$$

矩阵 $\boldsymbol{H}$ 中的非零元素为

$$\begin{aligned}H_{11} &= H_{15} = 2/(h_1 + h_5)\\ H_{22} &= H_{26} = 2/(h_2 + h_6)\\ H_{33} &= H_{34} = 2/(h_3 + h_4)\end{aligned} \tag{3.3-62}$$

单元绕相邻边界的转角增量 α_i 由节点的面外位移增量 w_i 表示为

$$\alpha = Cw \tag{3.3-63}$$

式中

$$\boldsymbol{w} = \{w_1\quad w_2\quad w_3\quad w_4\quad w_5\quad w_6\}^T \tag{3.3-64}$$

矩阵 $\boldsymbol{C}$ 的非零元素为

$$\begin{aligned}&C_{11} = 1/h_1,\ C_{12} = -\cos\beta_3/h_2,\ C_{13} = -\cos\beta_2/h_3\\ &C_{21} = -\cos\beta_3/h_1,\ C_{22} = 1/h_2,\ C_{23} = -\cos\beta_1/h_3\\ &C_{31} = -\cos\beta_2/h_1,\ C_{32} = -\cos\beta_1/h_2,\ C_{33} = 1/h_3\\ &C_{41} = -\cos\gamma_2/q_1,\ C_{42} = -\cos\gamma_1/q_2,\ C_{44} = 1/h_4\\ &C_{52} = -\cos\phi_3/r_2,\ C_{53} = -\cos\phi_2/r_3,\ C_{55} = 1/h_5\\ &C_{61} = -\cos\phi_3/s_1,\ C_{63} = -\cos\phi_1/s_3,\ C_{66} = 1/h_6\end{aligned} \tag{3.3-65}$$

现在考虑图 3.3-9b 局部坐标系 xoy 下，三角形单元（123）的三条边（2，3）、（3，1）、（1，2）的外法线 n_1、n_2、n_3 分别是

$$\begin{aligned}&b_1 = (y_3 - y_2)/l_1\quad b_2 = (y_1 - y_3)/l_2\quad b_3 = (y_2 - y_1)/l_3\\ &c_1 = (x_2 - x_3)/l_1\quad c_2 = (x_3 - x_1)/l_2\quad c_3 = (x_1 - x_2)/l_3\end{aligned} \tag{3.3-66}$$

经过坐标变换和迭加，可得到局部坐标系 xoy 下的应变为

$$\boldsymbol{\varepsilon}=\begin{Bmatrix}\varepsilon_x\\ \varepsilon_y\\ \gamma_{xy}\end{Bmatrix}=\begin{bmatrix}b_1^2 & b_2^2 & b_3^2\\ c_1^2 & c_2^2 & c_3^2\\ 2b_1c_1 & 2b_2c_2 & 2b_3c_3\end{bmatrix}\begin{Bmatrix}\varepsilon_{n1}\\ \varepsilon_{n2}\\ \varepsilon_{n3}\end{Bmatrix}=\boldsymbol{R}\varepsilon_n \tag{3.3-67}$$

1.4.2 边界条件

为描述自由和固定边界条件，可引入一个对称的虚单元（124）和单元（123）边界相邻，如图3.3-10所示。

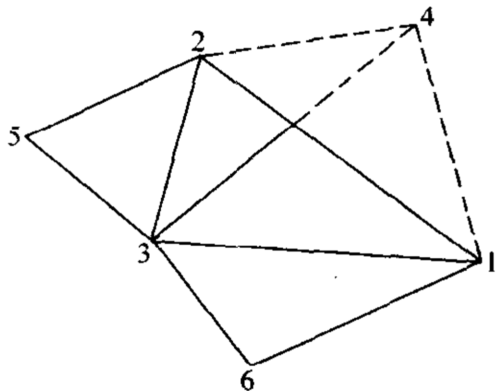


图3.3-10 虚单元

当（1，2）边为自由边界条件时，（1，2）边转动不受限制，沿（1，2）边法线方向的弯曲应变 ε_{n3} 消失。令单元（123）和虚单元（124）上的对称节点3、4的面外位移大小相等、方向相反，即 $w_3=-w_4$，则有 $\alpha_3=-\alpha_4$，则矩阵 $\boldsymbol{C}$ 的第四行改变为

$$[\cos\beta_2/h_1 \quad \cos\beta_1/h_2 \quad -1/h_3 \quad 0 \quad 0 \quad 0] \tag{3.3-68}$$

对于固定边界条件情况，（1，2）边转角为零，可令单元（123）和虚单元（124）上的对称节点3、4的面外位移大小相等、方向相同，即 $w_3=w_4$，则有 $\alpha_3=\alpha_4$，矩阵 $\boldsymbol{C}$ 的第四行改变为

$$[-\cos\beta_2/h_1 \quad -\cos\beta_1/h_2 \quad 1/h_3 \quad 0 \quad 0 \quad 0] \tag{3.3-69}$$

1.4.3 三维情况

考虑三维变形时，相邻单元不在同一平面上的情况。设 δ_{0m}、δ_{0k} 分别是单元 m、k 在 t 时刻的曲率角，δ_m、δ_k 分别是单元 m、k 在 $t+\Delta t$ 时刻的曲率角，θ_m、θ_k 分别是 Δt 时刻内的曲率角的变化，则有

$$\theta_m=\delta_m-\delta_{0m} \quad \theta_k=\delta_k-\delta_{0k} \tag{3.3-70}$$

式中

$$\delta_m=h_m/2R,\ \delta_k=h_k/2R \tag{3.3-71}$$

$$\delta_{0m}=h_{0m}/2R_0,\ \delta_{0k}=h_{0k}/2R_0 \tag{3.3-72}$$

式中，$\delta_{0m}=\delta_{0k}=0$ 对应于单元初始共面情况；h_{0m}、h_{0k}、h_m、h_k 分别是 t 时刻和 $t+\Delta t$ 时刻三角形单元的高；R_0、R 分别是 t 时刻和 $t+\Delta t$ 时刻三角形单元的曲率。

令 α_m 和 α_k 表示单元绕相邻边界转角的增量，仍假设相邻单元沿公共边界为常曲率，则由式（3.3-57）、式（3.3-70）、式（3.3-71）可得

$$\begin{aligned}\theta_k&=[h_k(\alpha_k+\alpha_m+\delta_{0m})-h_m\delta_{0k}]/(h_k+h_m)\\ \theta_m&=[h_m(\alpha_k+\alpha_m+\delta_{0k})-h_k\delta_{0m}]/(h_k+h_m)\end{aligned} \tag{3.3-73}$$

忽略不计转动过程中单元长度的变化，则由式（3.3-72）可得

$$\delta_{0k}=h_{0k}\delta_{0m}/h_{0m}=h_k\delta_{0m}/h_m \tag{3.3-74}$$

由式（3.3-73）、式（3.3-74）可得

$$\left.\begin{aligned}\theta_k&=h_k(\alpha_k+\alpha_m)/(h_k+h_m)\\ \theta_m&=h_m(\alpha_k+\alpha_m)/(h_k+h_m)\end{aligned}\right\} \tag{3.3-75}$$

转角增量 α_m、α_k 引起的沿相邻边界法线方向的弯曲应变为

$$\varepsilon_{nk}=-z\frac{R_0-R}{R_0R}=-2z\left(\frac{\delta_k}{h_k}-\frac{\delta_{0k}}{h_{0k}}\right)=-2z\frac{\delta_k-\delta_{0k}}{h_k}=-2z\frac{\theta_k}{h_k} \tag{3.3-76}$$

由式（3.3-75）、式（3.3-76）可知，转角增量 α_m、α_k 引起的沿相邻边界法线方向的弯曲应变为

$$\varepsilon_{nk}-2z\theta_k/h_k=-2z(\alpha_k+\alpha_m)/(h_k+h_m) \tag{3.3-77}$$

由式（3.3-77）可知，式（3.3-60）中的矩阵 $\boldsymbol{H}$ 保持不变。

由于相邻单元不再共面，面外位移必须投影到每个单元各自的局部坐标系下，则式（3.3-63）变为

$$\boldsymbol{\alpha}=\boldsymbol{C}\boldsymbol{w} \tag{3.3-78}$$

式中，矩阵 $\boldsymbol{C}$ 的非零元素为

$$\begin{aligned}&C_{11}=1/h_1,\ C_{12}=-\cos\beta_3/h_2,\ C_{13}=-\cos\beta_2/h_3\\ &C_{21}=-\cos\beta_3/h_1,\ C_{22}=1/h_2,\ C_{23}=-\cos\beta_1/h_3\\ &C_{31}=-\cos\beta_2/h_1,\ C_{32}=-\cos\beta_1/h_2,\ C_{33}=1/h_3\\ &C_{44}=-\cos\gamma_2/q_1,\ C_{45}=-\cos\gamma_1/q_2,\ C_{46}=1/h_4\\ &C_{57}=-\cos\phi_3/r_2,\ C_{58}=-\cos\phi_2/r_3,\ C_{59}=1/h_5\\ &C_{6,10}=-\cos\phi_3/s_1,\ C_{6,11}=-\cos\phi_1/s_3,\ C_{6,12}=1/h_6\end{aligned} \tag{3.3-79}$$

$$\boldsymbol{w}=[w_1^1\ w_2^1\ w_3^1\ w_1^2\ w_2^2\ w_4^2\ w_2^3\ w_3^3\ w_5^3\ w_3^4\ w_1^4\ w_6^4]^{\mathrm T} \tag{3.3-80}$$

式中 w_i^j 表示节点 i 投影到单元 j 的局部坐标系下的面外位移。

2 本构方程

在金属塑性大变形有限元分析时经常采用流动理论本构方程，其他本构方程很少采用。例如，基于形变理论或非经典的角点理论本构方程虽然可以比较准确反映板料失稳后的局部化变形过程，但是板料成形属于强约束过程，对角点本构方程不敏感，而且板料成形也并不十分关心板料失稳后的局部化变形过程。还有一个重要原因是，目前板料成形有限元模拟的精度也不高，也没有必要十分强调本构方程的影响。

为了推导公式方便，本节将采用张量分量记法。考虑具有光滑屈服面屈服函数的弹塑性体，假设温度对变形速度的影响很小，可以忽略不计。这样全应变率 $\dot\varepsilon_{ij}$ 可以分解为弹性应变率 $\dot\varepsilon_{ij}^{e}$ 和塑性应变率 $\dot\varepsilon_{ij}^{p}$ 之和，即

$$\dot\varepsilon_{ij}=\dot\varepsilon_{ij}^{e}+\dot\varepsilon_{ij}^{p} \tag{3.3-81}$$

采用Hooke定律，弹性应变率 $\dot\varepsilon_{ij}^{e}$ 用第二Piola－Kirchhoff应力的Jaumann导数 $\overset{\nabla}{S}_{ij}$ 表示为

$$\dot\varepsilon_{ij}^{e}=B_{ijkl}^{e}\overset{\nabla}{S}_{kl} \tag{3.3-82}$$

式（3.3-82）的逆关系为

$$\overset{\nabla}{S}_{ij}=D_{ijkl}^{e}\dot\varepsilon_{kl}^{e} \tag{3.3-83}$$

塑性应变率 $\dot\varepsilon_{ij}^{p}$ 用流动法则和屈服函数 f 表示为

$$\dot\varepsilon_{ij}^{p}=\alpha n_{kl}\overset{\nabla}{S}_{kl}n_{ij}/h \tag{3.3-84}$$

式中，n_{ij} 是屈服面（应力空间 $f=0$ 的曲面）的单位法向量；$\alpha=1$ 或 $\alpha=0$，当应力点位于屈服面以内时，应力处于弹性状态 $\alpha=0$，当应力点位于屈服面上，而且应力率 $\overset{\nabla}{S}_{ij}$ 指向屈服面以外时为塑性加载状态 $\alpha=1$ 且 $n_{ij}\overset{\nabla}{S}_{ij}>0$，如果应力率 $\overset{\nabla}{S}_{ij}$ 与屈服面相切时，为卸载状态 $\alpha=0$；h 表示当前状态的加工硬化率，它是应力和应变的函数。

由式（3.3-82）和式（3.3-84）可得全应变率 $\dot\varepsilon_{ij}$ 为

$$\dot\varepsilon_{ij}=\left(B_{ijkl}^{e}+\frac{\alpha}{h}n_{ij}n_{kl}\right)\overset{\nabla}{S}_{kl}=B_{ijkl}^{ep}\overset{\nabla}{S}_{kl} \tag{3.3-85}$$

式（3.3-85）的逆关系就是材料有限变形本构关系

$$\overset{\nabla}{S}_{ij}=\left(D_{ijkl}^{e}+\frac{\alpha}{g}m_{ij}m_{kl}\right)\dot\varepsilon_{kl}=D_{ijkl}^{ep}\dot\varepsilon_{kl} \tag{3.3-86}$$

式中

$$\begin{aligned}m_{ij}&=D_{ijkl}^{ep}n_{kl}\\ g&=h+m_{ij}n_{ij}\end{aligned} \tag{3.3-87}$$

由于 $\overset{\nabla}{S}_{ij}$ 与 Cauchy 应力的 Jaumann 导数 $\overset{\nabla}{\sigma}_{ij}$ 的关系为

$$\overset{\nabla}{S}_{ij} = \overset{\nabla}{\sigma}_{ij} + \sigma_{ij} v_{k,k} \tag{3.3-88}$$

如果材料不可压缩，$v_{k,k}=0$，因此

$$\overset{\nabla}{S}_{ij} = \overset{\nabla}{\sigma}_{ij} \tag{3.3-89}$$

式（3.3-86）可简化为

$$\overset{\nabla}{\sigma}_{ij} = (D^{e}_{ijkl} + D^{p}_{ijkl})\dot{\varepsilon}_{kl} = D^{ep}_{ijkl}\dot{\varepsilon}_{kl} \tag{3.3-90}$$

式（3.3-86）和式（3.3-90）是三维状态下的材料本构方程，对于板料而言一般都属于平面应力状态，因此要把它们简化成平面应力状态下的本构方程。一般平面应力状态假设有

$$\sigma_{33} = 0 \tag{3.3-91}$$

根据 $\overset{\nabla}{\sigma}_{ij}$ 与 σ_{ij} 的关系

$$\overset{\nabla}{\sigma}_{ij} = \dot{\sigma}_{ij} - \sigma_{mj}\omega_{im} - \sigma_{mi}\omega_{jm} \tag{3.3-92}$$

可得平面应力状态下同样有

$$\overset{\nabla}{\sigma}_{33} = 0 \tag{3.3-93}$$

根据式（3.3-90）得

$$\overset{\nabla}{\sigma}_{33} = D^{ep}_{33kl}\dot{\varepsilon}_{kl} = 0 \tag{3.3-94}$$

式（3.3-90）可以表示成

$$\overset{\nabla}{\sigma}_{ij} = D^{ep}_{ijkl}\dot{\varepsilon}_{kl} + D^{ep}_{ij33}\dot{\varepsilon}_{33} \quad (k、l\text{ 不同时等于 }3) \tag{3.3-95}$$

再由式（3.3-93）可得

$$\overset{\nabla}{\sigma}_{33} = D^{ep}_{33kl}\dot{\varepsilon}_{kl} + D^{ep}_{3333}\dot{\varepsilon}_{33} = 0 \quad (k、l\text{ 不同时等于 }3) \tag{3.3-96}$$

$$\dot{\varepsilon}_{33} = -\frac{D^{ep}_{33kl}}{D^{ep}_{3333}}\dot{\varepsilon}_{kl} \quad (k、l\text{ 不同时等于 }3) \tag{3.3-97}$$

将式（3.3-97）代入式（3.3-95）得

$$\overset{\nabla}{\sigma}_{ij} = (D^{ep}_{ijkl} - \frac{D^{ep}_{ij33}D^{ep}_{33kl}}{D^{ep}_{3333}})\dot{\varepsilon}_{kl} = \bar{D}^{ep}_{ijkl}\dot{\varepsilon}_{kl} \quad (k、l\text{ 不同时等于 }3) \tag{3.3-98}$$

式中，$\bar{D}^{ep}_{ijkl}$ 就是平面应力状态下的材料本构矩阵

$$\bar{D}^{ep}_{ijkl} = D^{ep}_{ijkl} - \frac{D^{ep}_{ij33}D^{ep}_{33kl}}{D^{ep}_{3333}} \quad (k、l\text{ 不同时等于 }3) \tag{3.3-99}$$

2.1 J_2 流动理论

弹性问题本构关系为

$$\begin{aligned} \overset{\nabla}{S}_{ij} &= D^{e}_{ijkl}\varepsilon^{e}_{kl} \\ D^{e}_{ijkl} &= 2G[\frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) - \frac{\nu}{1-2\nu}\delta_{ij}\delta_{kl}] \end{aligned} \tag{3.3-100}$$

式中，G 为切变模量；ν 为泊松比。

Mises 屈服函数为

$$f = \frac{1}{2}\sigma'_{ij}\sigma'_{ij} - \frac{1}{3}\bar{\sigma}^2 \tag{3.3-101}$$

式中，σ'_{ij} 表示应力 σ_{ij} 的偏量。在 Mises 屈服函数中，式（3.3-85）中的 n_{ij} 为

$$n_{ij} = \frac{\sigma'_{ij}}{\sqrt{2/3}\bar{\sigma}} \tag{3.3-102}$$

因此式（3.3-85）可以表示成

$$\dot{\varepsilon}_{ij} = (B^{e}_{ijkl} + \frac{3\alpha}{2h\bar{\sigma}^2}\sigma'_{ij}\sigma'_{kl})\overset{\nabla}{S}_{kl} = B^{ep}_{ijkl}\overset{\nabla}{S}_{kl} \tag{3.3-103}$$

式（3.3-103）的逆就是材料的 J_2 流动本构关系

$$\overset{\nabla}{S}_{ij} = (D^{e}_{ijkl} - \frac{2G\alpha}{g}\sigma'_{ij}\sigma'_{kl})\dot{\varepsilon}_{kl} = D^{ep}_{ijkl}\dot{\varepsilon}_{kl} \tag{3.3-104}$$

式中

$$\begin{aligned} g &= \frac{2}{3}\bar{\sigma}^2(1 + \frac{h}{2G}) \\ \bar{\sigma}^2 &= \frac{3}{2}\sigma'_{ij}\sigma'_{ij} \end{aligned} \tag{3.3-105}$$

h 可以由单向拉伸实验确定

$$\frac{1}{h} = \frac{3}{2}[(1 - \frac{1-2\nu}{E}\sigma)\frac{1}{E_t} - \frac{1}{E}] \tag{3.3-106}$$

式中，E 为弹性模量；E_t 为单向拉伸正应力－对数应变曲线的切线模量

$$E_t = \frac{d\sigma}{d\varepsilon} \tag{3.3-107}$$

对于不可压缩材料式（3.3-106）可以简化为

$$\frac{1}{h} = \frac{3}{2}(\frac{1}{E_t} - \frac{1}{E}) \tag{3.3-108}$$

2.2 J_2 随动强化理论

材料的初始屈服面可以表示为

$$f = \frac{1}{2}\sigma'_{ij}\sigma'_{ij} - \frac{1}{3}\bar{\sigma}_0^2 \tag{3.3-109}$$

在随动强化理论中假设材料在塑性变形时，式（3.3-109）所描述的屈服面保持形状和大小不变，只是在应力空间中伴随刚体回转而移动。因此，若以 σ_{ij} 表示当前变形的屈服面中心位置的话，当前状态的屈服面为

$$\begin{aligned} f &= \frac{1}{2}\bar{\sigma}'_{ij}\bar{\sigma}'_{ij} - \frac{1}{3}\bar{\sigma}_0^2 \\ \bar{\sigma}'_{ij} &= \sigma_{ij} - \alpha'_{ij} \end{aligned} \tag{3.3-110}$$

式中，σ'_{ij} 是 σ_{ij} 的偏量。

根据流动法则，塑性应变率与当前屈服面的法向平行，因此式（3.3-85）中的 n_{ij} 为

$$n_{ij} = \frac{\bar{\sigma}'_{ij}}{\sqrt{2/3}\sigma_0} \tag{3.3-111}$$

全应变率 $\dot{\varepsilon}_{ij}$ 可以表示成

$$\dot{\varepsilon}_{ij} = (B^{e}_{ijkl} + \frac{3\alpha}{2\bar{h}\bar{\sigma}_0^2}\bar{\sigma}'_{ij}\bar{\sigma}'_{kl})\overset{\nabla}{S}_{kl} = B^{ep}_{ijkl}\overset{\nabla}{S}_{kl} \tag{3.3-112}$$

式中 $\bar{h}$ 可以由单向拉伸实验确定

$$\frac{1}{\bar{h}} = \frac{3}{2}[(1 - \frac{1-2\nu}{E}\sigma)\frac{1}{E_t} - \frac{1}{E}] \tag{3.3-113}$$

如果材料不可压缩，则

$$\frac{1}{\bar{h}} = \frac{3}{2}(\frac{1}{E_t} - \frac{1}{E}) \tag{3.3-114}$$

式（3.3-112）的逆就是 J_2 随动强化本构关系

$$\overset{\nabla}{S}_{ij} = (D^{e}_{ijkl} + \frac{2G\alpha}{\bar{g}}\bar{\sigma}'_{ij}\bar{\sigma}'_{kl})\dot{\varepsilon}_{kl} \tag{3.3-115}$$

式中

$$\bar{g} = \frac{2}{3}\bar{\sigma}_0^2(1 + \frac{\bar{h}}{2G}) \tag{3.3-116}$$

塑性变形过程中屈服面的移动速度 $\overset{\nabla}{\alpha}_{ij}$ 与应力点所在屈服面中心的相对位置平行，并指向 $\sigma_{ij} - \alpha_{ij}$。因此

$$\overset{\nabla}{\alpha}_{ij} = \dot{\mu}(\sigma_{ij} - \alpha_{ij}) \tag{3.3-117}$$

式中

$$\dot{\mu} = \frac{3\bar{\sigma}'_{ij}\overset{\nabla}{S}_{ij}}{2\bar{\sigma}_0^2} \tag{3.3-118}$$

2.3 各向异性理论

如果我们把各向同性 Mises 屈服函数用于各向异性材料，则屈服函数一般的可以写成

$$f = \frac{1}{2}C_{ijkl}\sigma_{ij}\sigma_{kl} - \frac{1}{3}\bar{\sigma}^2 \tag{3.3-119}$$

其中 C_{ijkl} 是四阶张量。当它是四阶各向同性张量时，形式为

$$C_{ijkl}=\frac{1}{2}\ (\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})\ -\frac{1}{3}\delta_{ij}\delta_{kl} \quad (3.3\text{-}120)$$

C_{ijkl}关于i和j，k和l以及ij和kl对称。$\bar{\sigma}$为等效应力。采用功的互等定理，等效塑性应变$\dot{\bar{\varepsilon}}_p$可以表示为

$$\bar{\sigma}^2=\frac{3}{2}C_{ijkl}\sigma_{ij}\sigma_{kl},\ \dot{\bar{\varepsilon}}_p^2=\frac{2}{3}\bar{C}_{ijkl}\dot{\varepsilon}_{ij}\dot{\varepsilon}_{kl} \quad (3.3\text{-}121)$$

设与屈服曲面f相正交的单位法向量为n_{ij}，则有

$$n_{ij}=\frac{l_{ij}}{\sqrt{2/3}\bar{\sigma}},\ l_{ij}=\frac{\partial f}{\partial \sigma_{ij}} \quad (3.3\text{-}122)$$

以塑性应变率表示的本构方程为

$$\dot{\varepsilon}^p_{ijkl}=\frac{l_{ij}l_{kl}\overset{\nabla}{S}_{kl}}{4/9\bar{\sigma}^2 H} \quad (3.3\text{-}123)$$

其中$\overset{\nabla}{S}_{ij}$为Kirchhoff应力的Jaumann率，$H=\mathrm{d}\bar{\sigma}/\mathrm{d}\bar{\varepsilon}_p$。将弹性应变率本构方程代入，则得到全应变率的本构方程

$$\dot{\varepsilon}_{ij}=(B^e_{ijkl}+\frac{l_{ij}l_{kl}}{4/9\bar{\sigma}^2 H})\overset{\nabla}{S}_{kl} \quad (3.3\text{-}124)$$

它的逆关系为

$$\overset{\nabla}{S}_{ij}=(D^e_{ijkl}-\frac{m_{ij}m_{kl}}{4/9\bar{\sigma}^2 H+m_{rs}l_{rs}})\dot{\varepsilon}_{kl} \quad (3.3\text{-}125)$$

式中

$$m_{ij}=D^e_{ijkl}l_{kl} \quad (3.3\text{-}126)$$

3 各向异性屈服函数

金属薄板在预加工和轧制过程中会产生明显的各向异性，这种结构上的各向异性对其成形规律有显著的影响。在拉深成形过程中，突缘出现制耳、冲压件断裂位置和极限成形高度的改变等现象，都是由于板材的各向异性使其在成形过程中塑性流动发生改变所造成的。

在度量板材的各向异性性质强弱的时候，各向异性参数R值是非常重要的一个参数，是评价板材成型性能的重要指标。R值愈大，材料愈不容易在厚向减薄或增厚；反之，R值愈小，材料愈容易减薄或增厚。当沿与x成φ角对板料施加单向拉伸时，R的定义为

$$R=\frac{\dot{\varepsilon}_t}{\dot{\varepsilon}_z} \quad (3.3\text{-}127)$$

式中，$\dot{\varepsilon}_t$为垂直于拉伸方向的应变率；$\dot{\varepsilon}_z$为板厚方向的应变率。R值是随方向的变化而变化的，不同的方向R值是不一样。

面内各向异性系数ΔR也是一个重要指标

$$\Delta R=\ (R_0+R_{90}-2R_{45})\ /2 \quad (3.3\text{-}128)$$

ΔR值表示厚向各向异性参数R_h值在面内随方向的变化，它的大小决定了圆筒拉深突缘制耳形成的程度，影响材料在面内的塑性流动规律，与板材的成形性能无关。一般来说，ΔR值过大，对冲压成形是不利的。

常用的各向异性屈服函数主要有：Hill正交各向异性函数，Barlat－Lian屈服函数，Gotoh四次方屈服函数，Barlat六参量正交各向异性屈服函数等。

3.1 Hill正交各向异性函数

一般地，若把各向异性主轴作为随体坐标系的x，y，z轴，则Hill屈服函数可以表示成

$$f=\frac{1}{2(F+G+H)}[F(\sigma_{yy}-\sigma_{zz})^2+G(\sigma_{zz}-\sigma_{xx})^2+H(\sigma_{xx}-\sigma_{yy})^2+2L\sigma_{yz}^2+2M\sigma_{zx}^2+2N\sigma_{xy}^2]-\frac{1}{3}\bar{\sigma}^2 \quad (3.3\text{-}129)$$

式中，F，G，H，L，M，N为各向异性参数，由实验确定。

在平面应力状态下，应力$\sigma_{zz}=\sigma_{zx}=\sigma_{yz}=0$，因此式(3.3-129)可简化为

$$f=\frac{1}{2(F+G+H)}[(G+H)\sigma_{xx}^2+(F+H)\sigma_{yy}^2-2H\sigma_{xx}\sigma_{yy}+2N\sigma_{xy}^2]-\frac{1}{3}\bar{\sigma}^2 \quad (3.3\text{-}130)$$

下面讨论R值与各向异性参数F，G，H，N之间的关系。如图3.3-11所示，1方向为单拉方向，则

$$\sigma_{xx}=\sigma\cos^2\phi,\ \sigma_{yy}=\sigma\sin^2\phi,\ \sigma_{xy}=\sigma\sin\phi\cos\phi \quad (3.3\text{-}131)$$

$$\dot{\varepsilon}^p_{11}=\cos^2\phi\dot{\varepsilon}^p_{xx}+\sin^2\phi\dot{\varepsilon}^p_{yy}+2\sin\phi\cos\phi\dot{\varepsilon}^p_{xy} \quad (3.3\text{-}132)$$

$$\dot{\varepsilon}^p_{22}=\sin^2\phi\dot{\varepsilon}^p_{xx}+\cos^2\phi\dot{\varepsilon}^p_{yy}-2\sin\phi\cos\phi\dot{\varepsilon}^p_{xy} \quad (3.3\text{-}133)$$

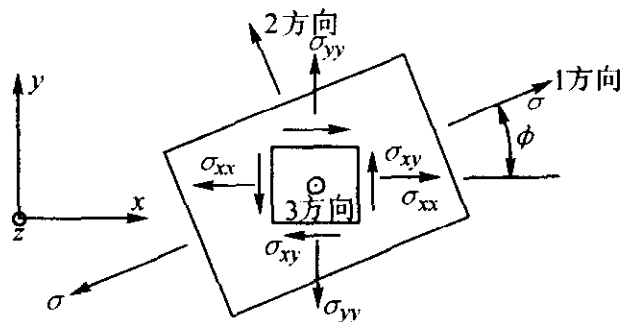


图3.3-11 单向拉伸受力情况

上面两式相加，可知

$$\dot{\varepsilon}^p_{11}+\dot{\varepsilon}^p_{22}=\dot{\varepsilon}^p_{xx}+\dot{\varepsilon}^p_{yy} \quad (3.3\text{-}134)$$

根据体积不可压缩性，有

$$\dot{\varepsilon}^p_{11}+\dot{\varepsilon}^p_{22}+\dot{\varepsilon}^p_{33}=0 \quad (3.3\text{-}135)$$

则任意ϕ角的R值为

$$R_\phi=-\frac{\dot{\varepsilon}^p_{22}}{\dot{\varepsilon}^p_{11}+\dot{\varepsilon}^p_{22}}=\frac{\dot{\varepsilon}^p_{11}}{\dot{\varepsilon}^p_{11}+\dot{\varepsilon}^p_{22}}-1=\frac{\dot{\varepsilon}^p_{11}}{\dot{\varepsilon}^p_{xx}+\dot{\varepsilon}^p_{yy}}-1 \quad (3.3\text{-}136)$$

根据关联流动法则，有

$$\dot{\varepsilon}^p_{xx}=\dot{\lambda}\frac{\partial f}{\partial\sigma_{xx}},\ \dot{\varepsilon}^p_{yy}=\dot{\lambda}\frac{\partial f}{\partial\sigma_{yy}},\ \dot{\varepsilon}^p_{xy}=\dot{\lambda}\frac{\partial f}{\partial\sigma_{xy}} \quad (3.3\text{-}137)$$

和

$$\dot{\varepsilon}^p_{11}=\dot{\lambda}\frac{\partial f^*}{\partial\sigma} \quad (3.3\text{-}138)$$

其中，f^*是将式(3.3-131)代入式(3.3-129)得到的，$\sigma=\sigma_{11}$。

将上面四式代入式(3.3-136)，即可求出任意角R与屈服函数f之间的关系式

$$R_\phi=\frac{\dfrac{\partial f^*}{\partial\sigma}}{\left(\dfrac{\partial f}{\partial\sigma_{xx}}+\dfrac{\partial f}{\partial\sigma_{yy}}\right)}-1 \quad (3.3\text{-}139)$$

$$R_\phi=\frac{H+\ (2N-F-G-4)\ \sin^2\phi\cos^2\phi}{F\sin^2\phi+G\cos^2\phi} \quad (3.3\text{-}140)$$

由于板材从厂家出厂时，一般只标有0°，45°，90°三个方向的R值，而式(3.3-140)中却有F，G，H，N四个参数。因此，一般令$H=1$，则式(3.3-140)可写为

$$R_\phi=\frac{1+\ (2N-F-G-4)\ \sin^2\phi\cos^2\phi}{F\sin^2\phi+G\cos^2\phi} \quad (3.3\text{-}141)$$

将0°，45°，90°方向的R值分别记为R_0，R_{45}，R_{90}，代入式(3.3-141)，得到各向异性参数F，G，N，H与R_0，R_{45}，R_{90}的关系为

$$G=\frac{1}{R_0},F=\frac{1}{R_{90}},N=(R_{45}+\frac{1}{2})(\frac{1}{R_0}+\frac{1}{R_{90}}) \quad (3.3\text{-}142)$$

因此，只要知道了R_0，R_{45}，R_{90}，就能求出F，G，N。

3.2 Barlat - Lian 屈服函数

若把各向异性主轴作为随体坐标系的 x，y，z 轴，则 Barlat - Lian 屈服函数的表达式为

$$f = a|K_1 + K_2|^M + a|K_1 - K_2|^M + c|2K_2|^M - 2\bar{\sigma}^M = 0 \tag{3.3-143}$$

式中

$$K_1 = \frac{1}{2}(\sigma_x + h\sigma_y), K_2 = \sqrt{\left(\frac{\sigma_x - h\sigma_y}{2}\right)^2 + p^2\sigma_{xy}^2}$$
$$h = \sqrt{r_0(1+r_{90})/[(1+r_0)r_{90}]}$$
$$a = 2 - c = 2 - \sqrt{r_0 r_{90}/[(1+r_0)(1+r_{90})]}$$
$$p = \frac{\sigma_p}{\sigma_b} = \left(\frac{\bar{\sigma}}{\tau_{sl}}\right)[2/(2a+2^M c)]^{\frac{1}{M}} \tag{3.3-144}$$

式中，σ_p 为等双拉状态的 Cauchy 主应力 σ_1；σ_b 为单项拉伸状态的 Cauchy 主应力 σ_1；τ_{sl} 为纯剪切状态时的屈服剪应力；M 为非二次屈服函数指数；r_0 和 r_{90} 为板料扎制方向和面内垂直于扎制方向的各向异性参数；p 值可以通过单拉实验的 r_0，r_{45}，r_{90} 求出。

这个屈服函数适用于表现为面内各向异性的各向异性材料。它对适于用 Taylor/Bishop 及 Hill 理论计算计算塑性势的多晶材料能进行很好的模拟，因此，这个公式可以用来研究多晶结构对金属板材冲压成形的影响。由于公式中只包含有 x、y 平面内的三个应力分量，因此上述屈服函数只能应用于平面应力状态，但它能描述各种平面应力状态。

与板料的扎制方向的夹角为 φ 的方向 1 的单拉应力为 σ，则根据应力变换得

$$\begin{aligned} \sigma_x &= \sigma\cos^2\varphi \\ \sigma_y &= \sigma\sin^2\varphi \\ \sigma_{xy} &= \sigma\sin\varphi\cos\varphi \end{aligned} \tag{3.3-145}$$

以及

$$\begin{aligned} \dot{\varepsilon}_{11} &= \cos^2\varphi\,\dot{\varepsilon}_{xx} = \sin^2\varphi\,\dot{\varepsilon}_{yy} + 2\sin\varphi\cos\varphi\,\dot{\varepsilon}_{xy} \\ \dot{\varepsilon}_{22} &= \sin^2\varphi\,\dot{\varepsilon}_{xx} + \cos^2\varphi\,\dot{\varepsilon}_{yy} - 2\sin\varphi\cos\varphi\,\dot{\varepsilon}_{xy} \end{aligned} \tag{3.3-146}$$

式中，$\dot{\varepsilon}_{11}$、$\dot{\varepsilon}_{22}$ 分别是单拉径向和宽度方向的应变率。由塑性不可压缩性，得任意角 φ 的 R 值

$$r_\phi = -\frac{\dot{\varepsilon}_{22}}{\dot{\varepsilon}_{11} + \dot{\varepsilon}_{22}} = \frac{\dot{\varepsilon}_{11}}{\dot{\varepsilon}_{11} + \dot{\varepsilon}_{22}} - 1 = \frac{\dot{\varepsilon}_{22}}{\dot{\varepsilon}_{xx} + \dot{\varepsilon}_{yy}} - 1 \tag{3.3-147}$$

由 Euler 相似方程理论得

$$\frac{\sigma(\cos^2\varphi\,\dot{\varepsilon}_{xx} + \sin^2\varphi\,\dot{\varepsilon}_{yy} + 2\sin\varphi\cos\varphi\,\dot{\varepsilon}_{xy})}{\lambda} = Mf \tag{3.3-148}$$

因此

$$\dot{\varepsilon}_{11} = \frac{\lambda Mf}{\sigma} = \frac{2\lambda M\bar{\sigma}^M}{\sigma} \tag{3.3-149}$$

r_φ 值可以表示为

$$r_\varphi = \frac{2M\bar{\sigma}^M}{\left(\frac{\partial f}{\partial\sigma_x} + \frac{\partial f}{\partial\sigma_y}\right)\sigma} - 1 \tag{3.3-150}$$

利用 r_φ 和屈服函数及相关的流动法则，计算任意方向的 r_φ 是可能的。特别地，这一组方程可用来利用 r_0、r_{90} 确定 a，c 和 h

$$a = 2 - c = 2 - 2\sqrt{\frac{r_0}{1+r_0}\frac{r_{90}}{1+r_{90}}} \tag{3.3-151}$$

$$h = \sqrt{\frac{r_0}{1+r_0}\frac{1+r_{90}}{r_{90}}} \tag{3.3-152}$$

3.3 Barlat 六参量正交各向异性屈服函数

由于以前提出的各种描述多晶体材料的各向异性本构方程多是只能描述材料的平面应力状态，给应用带来了很大的局限性。因此，Barlat 等人后来又提出了一个可以描述任何应力状态的屈服函数。这一描述包含了应力张量中的 6 个应力分量，反映屈服模型的指数 m 和 6 个材料系数 a，b，c，f，g，h。结果表明，Barlat 的屈服函数是正交各向异性材料的各向异性塑性行为的很好描述，由这一屈服函数反映的应力 - 应变响应与多晶体塑性力学的结果是相符合的，可以较准确地描述材料的力学行为，尤其适用于铝及其合金材料。

若把各向异性主轴作为随体坐标系的 x，y，z 轴，则六参量正交各向异性屈服函数的一般表达式为

$$f = f(\sigma_{ij}) - \bar{\sigma}^m = 0 \tag{3.3-153}$$

六参量正交各向异性屈服函数的标准表达式为

$$f = f(\sigma_{ij}) - \bar{\sigma} = 0 \tag{3.3-154}$$

定义

$$\begin{aligned} &A = \sigma_{yy} - \sigma_{zz},\ B = \sigma_{zz} - \sigma_{xx},\ C = \sigma_{xx} - \sigma_{yy} \\ &F = \sigma_{yz},\ G = \sigma_{zx},\ H = \sigma_{xy} \end{aligned} \tag{3.3-155}$$

由于板壳单元理论的假设

$$\sigma_{zz} = 0 \tag{3.3-156}$$

所以上面的几个定义式可以化简为

$$\begin{aligned} &A = \sigma_{yy},\ B = -\sigma_{xx},\ C = \sigma_{xx} - \sigma_{yy} \\ &F = \sigma_{yz},\ G = \sigma_{zx},\ H = \sigma_{xy} \end{aligned} \tag{3.3-157}$$

再定义

$$I_2 = \frac{(fF)^2 + (fG)^2 + (hH)^2}{3} + \frac{(aA - cC)^2 + (cC + bB)^2 + (bB - aA)^2}{54} \tag{3.3-158}$$

$$I_B = \frac{(cC - bB)(aA - cC)(bB - aA)}{54} + fghFGH - \frac{(cC - bB)(fF)^2 + (aA - cC)(gG)^2 + (bB - aA)(hH)^2}{6} \tag{3.3-159}$$

$$\theta = \arccos\left(\frac{I_3}{I_2^{3/2}}\right) \tag{3.3-160}$$

则可以写出 Barlat 六分量各向异性屈服函数的具体表达式

$$\Phi = (3I_2)^{m/2}\left\{\left[2\cos\left(\frac{2\theta+\pi}{6}\right)\right]^m + \left[2\cos\left(\frac{2\theta-3\pi}{6}\right)\right]^m + \left[-2\cos\left(\frac{2\theta+5\pi}{6}\right)\right]^m\right\} = 2\bar{\sigma}^m \tag{3.3-161}$$

$$R_\phi = \frac{2\sin^2\phi\cos^2\phi(9h^2 - a^2 - b^2 - 4c^2 - 2bc - 2ac + ab) + (2c^2 + ac + bc - ab)}{2(b^2 + bc)\cos^2\phi + 2(a^2 + ac)\sin^2\phi + (ab - ac - bc)} \tag{3.3-162}$$

一般而言，轧钢厂提供轧制板材时，厂家只是标定三个方向的 R 值，即 R_0，R_{45}，R_{90}。因此，将 0°，45°，90°分别代人 R_ϕ 得到如下联立方程

$$\begin{aligned} R_0 &= \frac{2c^2 + ac + bc - ab}{2b^2 + bc + ab - ac} \\ R_{45} &= \frac{9h^2 - a^2 - b^2 - ab}{2a^2 + 2b^2 + 2ab} \\ R_{90} &= \frac{2c^2 + ac + bc - ab}{2a^2 + ab + ac - bc} \end{aligned} \tag{3.3-163}$$

解上述三个方程，即可由 R_0，R_{45}，R_{90} 求出 a、b、c、h 四个参数，但三个方程是无法求出四个未知量的。一般来讲 f、g、h 分别代表着各向异性对 σ_{yz}，σ_{zx}，σ_{xy} 这三个应力项的影响，在各向同性的情况下，这六个系数都等于 1。我们可以近似地认为在描述轧制板材的各向异性特性时，也令 f、g、h 等于 1。于是三个方程中只有三个未知量，可以求出这三个参数 a、b、c 了。

4　工艺条件约束处理

接触问题通常是定义为边界值问题，两个物体 B^1 和 B^2 根据连续介质力学原理相互作用。因此，接触问题的最初的公理是动力学公理，即两个物体在任何构形下都不应该相互穿过。这就是所谓的“不可穿过性条件”。其另一个条件是在 B^1 和 B^2 上的边界材料点在物体运动过程中可以贴合也可以分离，这部分边界称为接触表面。在某个表面上的接触条件描述如下。

1）几何条件　在接触点处，两接触表面具有相反的外法线，即

$$\begin{aligned} x^1 &= x^2 \\ n^1 &= -n^2 \end{aligned} \tag{3.3-164}$$

2）动学条件　相互接触的物体在任何构形下都不应该相互穿透，接触点以同样的位移和速度沿接触表面外法线方向运动，即

$$\begin{aligned} u^1 n^1 &= u^2 n^2 \\ \dot{u}^1 n^1 &= \dot{u}^2 n^2 \end{aligned} \tag{3.3-165}$$

3）动力学条件　相互接触的物体在接触表面上沿外法线方向没有拉力作用，接触力的切向分量遵守库仑摩擦定律，即

$$T^1 + T^2 = 0 \tag{3.3-166}$$

$$T^\alpha n^\alpha \leqslant 0,\ \alpha = 1,\ 2 \tag{3.3-167}$$

$$T = T_n n + T_t t,\ 0 \leqslant T_t \leqslant \mu T n \tag{3.3-168}$$

式中，μ 为物体 B^1 和 B^2 之间的摩擦因数。

板材冲压成形过程中，模具可看作是一个刚体，因此模具与板材接触问题被认为是一个变形体和一个任意形状的刚体之间的接触。对于动力显式和静力隐式两种算法来说，对接触和摩擦问题的处理是截然不同的。对于静力隐式算法来说，接触和摩擦问题引发的非线性常常会引起解的不稳定；与之相比，动力显式算法的一个明显优势，就是对接触和摩擦问题的处理非常简便。

动力显式有限元法计算接触力的约束方法主要有两种：罚函数法和拉格朗日乘子法。罚函数法简单，但有引人人为误差和影响显式算法稳定性的缺点。拉格朗日乘子法虽然能够精确求得接触力，但不能直接用于显式算法，因为它涉及未知接触力的联立方程求解，破坏了显式算法不求解联立方程组，计算效率高的优势。正因为这样，在显式算法中，以往只采用罚函数法求接触力。为了避免罚函数法引起的误差和对求解稳定性的影响，有些学者提出了一些特殊算法以便能用拉格朗日乘子法求解接触力，同时避免求解联立方程组。

4.1　罚函数法

在罚函数法中，一个接触点是允许穿透它的接触表面的，并假设作用在接触点上的接触力与它的穿透量成正比。

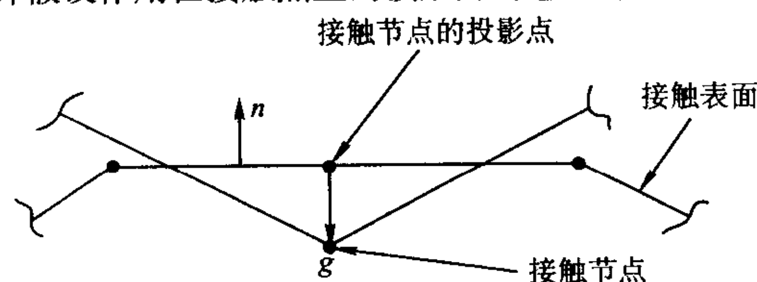


图 3.3-12　接触节点穿透接触表面

如图 3.3-12 所示，设接触节点坐标为 x（ξ_c，η_c），接触节点在工具表面上的投影点坐标为 x^0，则作用在接触节点上的法向接触力计算如下

$$\lambda = \varepsilon g_n \neq 0 \quad \text{仅当} \ |g_n| < 0 \ \text{时} \tag{3.3-169}$$

$$|g_n| = n\left(x(\xi_c,\ \eta_c) - x^0\right) \begin{cases} <0 & \text{穿透} \\ >0 & \text{间隙} \end{cases} \tag{3.3-170}$$

式中，n 为工具表面的外法线；ε 为罚因子，物理意义上相当于接触节点和接触表面之间的弹簧，其量纲与弹性模量相同。

应当指出，尽管罚函数法简单，计算工作量也小，但它有自身的缺陷。首先必须选择一个适当的罚因子，以保证边界穿透量足够小，同时保证求解稳定性不受影响。在罚函数法中，要使边界穿透量小，就要增大罚因子的值。而另一方面，罚因子增大将导致系统的频率增加，从而影响系统的临界时间步长。如果两边界的刚性相近的话，罚因子的选择较好处理。当其中一个边界的刚性远比另一边界的刚性大时，罚因子的选择变得较为棘手。若罚因子太小，边界穿透可能太大；若罚因子太大，临界时间步长可能减小很多，并使系统响应严重失真。为了避免这些情况，罚因子必须根据接触面刚度进行选择。显式算法中，罚因子的选取如下。

1）固体单元

$$\varepsilon_{so} = \alpha \frac{A_s K}{h} = \alpha \frac{A_s^2 K}{V_e} \tag{3.3-171}$$

2）壳单元

$$\varepsilon_{sh} = \frac{A_s K}{l} \tag{3.3-172}$$

式中，A_s 为固体单元面积；h 为固体单元厚度；V_e 为固体单元体积；l 为壳单元最大名义长度；α 为比例因子（≈0.1）；K 为某个固定的大数。

4.2　防御节点法

防御节点法是一种既能精确计算接触力，又能避免求解联立方程组的算法。该算法在冲压成形过程的计算机仿真和汽车碰撞过程的计算机仿真等领域得到了成功的应用。在防御节点法中，每个接触节点和它对应的接触块构成一个接触对，接触节点称为从接触点，接触块上的节点称为主接触点。每个接触对中都增加了一个虚拟的接触节点位于接触块上，即所谓的防御节点，如图 3.3-13 所示。防御节点尽管是一个虚拟节点，但它具有一个普通的接触节点所具有的所有属性，如速度、加速度和力等。

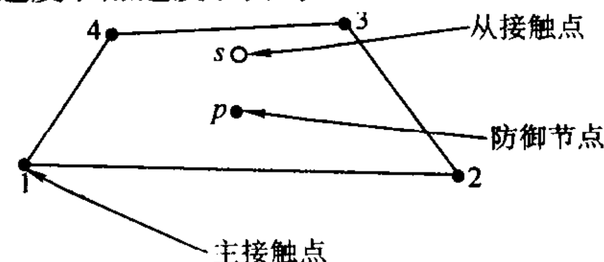


图 3.3-13　防御节点的建立

设防御节点所在的位置为 p（ξ，η）点，那么防御节点的运动参数可通过如下等参插值公式求得

$$\begin{aligned} u &= \phi_i(\xi,\eta) u_i \\ v &= \phi_i(\xi,\eta) v_i \\ a &= \phi_i(\xi,\eta) a_i \end{aligned} \tag{3.3-173}$$

式中，u、v 和 a 分别表示防御节点的位移、速度和加速度；u_i、v_i 和 a_i 分别表示接触块上第 i 个节点的位移、速度和加速度；ϕ_i（ξ，η）表示对应于接触块上第 i 个节点的形函数在 p（ξ，η）点的值。

防御节点的位移、速度和加速度按方程式（3.3-173）计算是比较自然的，但防御节点的质量和节点力的计算就没有这么简单了。为了推导出这些量的计算公式，现假设防御节点的质量为 M，作用在防御节点上的总节点力（接触力除外）为 F，作用在防御节点上的接触力为 f，那么防御节点

的运动方程可写成如下形式

$$Ma = F + f \qquad (3.3\text{-}174)$$

在防御节点法中，接触力 f 是通过联立接触点和防御节点的运动方程来求解的，并将按一定规则分配到接触块的各个节点上。假设接触力 f 分配到接触块第 i 个节点上的那部分为 f_i，那么块节点 i 的运动方程可写成

$$M_i a_i = F_i + f_i \qquad (3.3\text{-}175)$$

式中，M_i、a_i 和 F_i 分别表示块节点 i 的质量、加速度和节点力。将方程式（3.3-173）代入方程式（3.3-174）得

$$M[\phi_i(\xi,\eta)a_i] = F + f \qquad (3.3\text{-}176)$$

将方程式（3.3-175）两边除以 M 并代入方程式（3.3-176）得

$$M[\phi_i(\xi,\eta)(F_i + f_i)/M_i] = F + f \qquad (3.3\text{-}177)$$

很显然，F 只能与 F_i 有关，而 f 只能与 f_i 有关。于是可从方程式（3.3-177）推导出如下关系式

$$F = M[\phi_i(\xi,\eta)F_i/M_i] \qquad (3.3\text{-}178)$$

$$f = M[\phi_i(\xi,\eta)f_i/M_i] \qquad (3.3\text{-}179)$$

方程式（3.3-178）可直接用来计算防御节点上的节点力，因为 F_i 是已知的，但方程式（3.3-179）不能直接用来计算 f，因为 f_i 本身也是未知数，并且与 f 有关。现假设 f_i 可通过下式计算

$$f_i = \varphi_i f \qquad (3.3\text{-}180)$$

式中，φ_i 为待定函数。由于所有的 f_i 都是由接触力 f 所引起的，所以很自然地得出

$$\sum_{i=1}^{m} f_i = f \qquad (3.3\text{-}181)$$

将方程式（3.3-180）代入方程式（3.3-181）并消去 f，可得

$$\sum_{i=1}^{m} \varphi_i = 1 \qquad (3.3\text{-}182)$$

将方程式（3.3-180）代入方程式（3.3-179）并消去 f，可得

$$M[\phi_i(\xi,\eta)\varphi_i/M_i] = 1 \qquad (3.3\text{-}183)$$

为满足方程式（3.3-183），可按下式选择函数 φ_i

$$\varphi_i = M_i(\phi_i)^{k-1}/[M\sum_{i=1}^{m}(\phi_i)^k] \quad (\text{不对 } i \text{ 求和}) \qquad (3.3\text{-}184)$$

式中，$k \geqslant 1$。当 $k=1$ 时，方程式（3.3-184）变为

$$\varphi_i = M_i/[M\sum_{i=1}^{m}\phi_i] = M_i/M \qquad (3.3\text{-}185)$$

将方程式（3.3-185）代入方程式（3.3-182）可得

$$M = \sum_{i=1}^{m} M_i \qquad (3.3\text{-}186)$$

方程式（3.3-186）告诉我们，如果按方程式（3.3-185）选择待定函数，那么防御节点的质量与其位置无关，并且总为接触块节点质量的总和。这显然是不合理的，因为当防御节点正好是接触块上的某一节点时，防御节点应与该节点具有完全相同的性质，因此应与该节点具有相同的质量。这样就排除了使用方程式（3.3-185）的可能性。当 $k=2$ 时，方程式（3.3-184）变为

$$\varphi_i = M_i\phi_i/[M\sum_{j=1}^{m}(\phi_j)^2] \quad (\text{不对 } i \text{ 求和}) \qquad (3.3\text{-}187)$$

将方程式（3.3-187）代入方程式（3.3-182），可得

$$M = M_i\phi_i/\sum_{j=1}^{m}(\phi_j)^2 \quad (\text{对 } i \text{ 求和}) \qquad (3.3\text{-}188)$$

如果防御节点正好处于节点 i 的位置，由上式可得

$$M = M_i \qquad (3.3\text{-}189)$$

这表明防御节点具有与节点 i 的相同的质量，同时从方程式（3.3-173）和方程式（3.3-178）可知，防御节点的其他物理参数也与节点 i 相同。因此可以接受方程式（3.3-187）作为选择待定函数 φ_i 的依据。由于 $k=2$ 已能满足要求，故不再考虑 $k>2$ 的其他情况。为剖析方程（3.3-188）的含义，设

$$m_i = M_i\phi_i/\sum_{j=1}^{m}(\phi_j)^2 \quad (\text{不对 } i \text{ 求和}) \qquad (3.3\text{-}190)$$

那么方程式（3.3-188）可写为

$$M = \sum_{i=1}^{m} m_i \qquad (3.3\text{-}191)$$

从上述方程不难看出，m_i 代表节点 i 对防御节点所贡献的质量，而 M 正好是接触块上 m 个节点贡献的质量的总和。

将方程式（3.3-190）代入方程式（3.3-187）得

$$\varphi_i = m_i/M \qquad (3.3\text{-}192)$$

将上式代入方程式（3.3-180）得

$$f_i = m_i f/M \qquad (3.3\text{-}193)$$

在以上讨论中，假设了一个接触块上只有一个从接触点，如果一个接触块上有多个从接触点，那么只能一个一个地处理。在处理其中一个主接触点时，将其余所有的从接触点用外力来代替，这时上述公式中的节点力 F_i 和 F 也包括其他从接触点上的接触力影响。

当防御节点建立后，就可方便地计算从接触点与防御节点间的接触力。为简明起见，可暂不考虑摩擦力，因此只需考虑法向接触力的计算。假设从接触点的质量、法向加速度、法向节点力和法向接触力分别为 M_1、a_1、F_1 和 f_1，防御节点上的对应量分别为 M_2、a_2、F_2 和 f_2。那么从接触点和防御节点的运动方程可写为

$$\begin{aligned} M_1 a_1 &= F_1 + f_1 \\ M_2 a_2 &= F_2 + f_2 \end{aligned} \qquad (3.3\text{-}194)$$

应用中心差分法，方程式（3.3-194）可写成

$$\begin{aligned} M_1[(u_1^t - u_1^\tau)/\Delta t - v_1^l]/\Delta t &= F_1^\tau + f_1^\tau \\ M_2[(u_2^t - u_2^\tau)/\Delta t - v_2^l]/\Delta t &= F_2^\tau + f_2^\tau \end{aligned} \qquad (3.3\text{-}195)$$

式中，上标 t、τ 和 l 分别表示三个相邻的时刻，如图 3.3-14 所示。

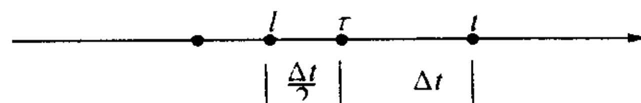


图 3.3-14　时间轴上三个相邻时刻

从接触点和防御节点的法向间距可表示为

$$g^\tau = g^l + u_2^t - u_2^\tau - (u_1^t - u_1^\tau) \qquad (3.3\text{-}196)$$

根据约束条件 $|g^\tau = 0|$，有

$$g^l + u_2^t - u_2^\tau - (u_1^t - u_1^\tau) = 0 \qquad (3.3\text{-}197)$$

由于 $f_1^\tau = -f_2^\tau$，从方程式（3.3-195）、式（3.3-196）、式（3.3-197）可得出

$$f_1^\tau = -f_2^\tau = M_1 M_2(F_2^\tau/M_2 - F_1^\tau/M_1 + v_2^l/\Delta t - v_1^l/\Delta t - g^l/\Delta t)/(M_1 + M_2) \qquad (3.3\text{-}198)$$

从上式可知，采用防御节点法后，接触力的计算不仅可以采用拉格朗日乘子法使约束条件得到精确满足，还可以避免求解联立方程组。

将方程式（3.3-198）代入方程式（3.3-194），可得到从接触点加速度计算公式如下

$$a_1 = [F_2^\tau + F_1^\tau + M_2(v_2^l - v_1^l)/\Delta t - M_2 g^l/\Delta t^2]/(M_1 + M_2) \qquad (3.3\text{-}199)$$

若假设冲头速度不受板料变形影响，即 $M_2 \to \infty$，$a_2 \to 0$，代入方程式（3.3-199）可得

$$a_1 = (v_2^l - v_1^l)/\Delta t - g^l/\Delta t^2 \qquad (3.3\text{-}200)$$

4.3　摩擦力的计算

在冲压成形过程中，摩擦起着举足轻重的作用，在冲压成形工艺中，有时需要加强摩擦作用，如增加压边力，而有时需要削弱摩擦作用，如涂润滑剂。无论是加强摩擦作用还是削弱摩擦作用，其目的都是相同的，即控制材料的流动。

经典的摩擦定律考虑的是两个接触物体间的干摩擦。如图3.3-15a所示，一物体 A 在力 P 的作用下与物体 B 接触。取物体 A 的脱离体可得图3.3-15b所示的受力图，其中 F_t 为作用在物体 A 上的切向力之和，也就是摩擦力的合力，F_n 为法向接触力的合力。

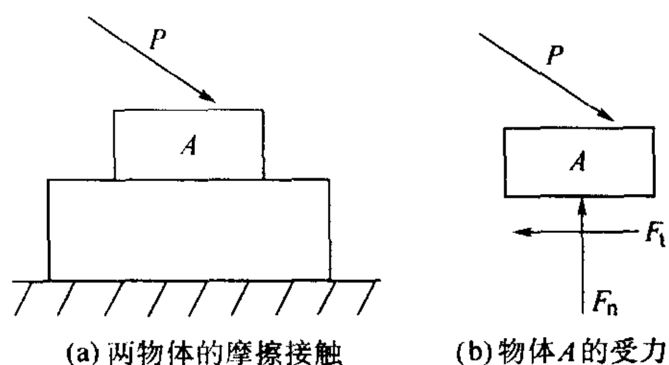


图3.3-15　摩擦接触与摩擦力

摩擦力 F_t 和接触力 F_n 之间的关系用经典摩擦定律描述如下。

1）如果两接触体处于静止状态，或者两接触体处于运动状态但没有相对运动，那么切向力必须达到一个临界值 F_c 才能使两物体产生切向相对运动。这个临界值与法向接触力 F_n 成正比

$$F_c = \mu_s F_n \tag{3.3-201}$$

式中，μ_s 表示静摩擦因数。当两接触体开始产生切向相对运动时，它们间的摩擦力 F_t 就等于临界值 F_c

$$F_t = \mu_s F_n \tag{3.3-202}$$

2）在两个物体的相对滑动中，摩擦力的大小与法向接触力成正比，其方向与切向相对运动方向相反，即

$$F_t = -\mu_d F_n \quad (v_t / |v_t|) \tag{3.3-203}$$

式中，μ_d 为两接触体之间的动摩擦系数；v_t 为两接触体间的相对滑动速度。

3）动摩擦系数小于静摩擦系数，即

$$\mu_d < \mu_s \tag{3.3-204}$$

而且两者都与接触面积和相对滑动的速度无关。

由上述定律可知，如果两个物体间的摩擦力小于临界值 F_c，那么该两物体处于纯黏着状态，即切向相对滑移量 $u_t = 0$。一旦两物体开始产生相对滑移，那么相对滑移量并不影响摩擦力的大小。假设法向接触一定，那么摩擦力和相对滑移量的关系可用图3.3-16所示的曲线来表示。经典摩擦定律中的这种摩擦力与相对滑移的关系假设有两个不足，首先是不符合微观摩擦现象，因为事实上任何小于 F_c 的摩擦力都可产生一定的微小相对滑移。另一个不足是把摩擦因数看成与相对滑移速度和接触面积无关，而实验表明摩擦因数

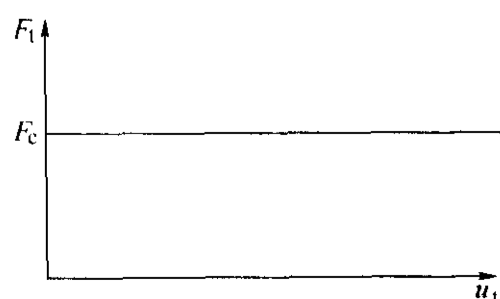


图3.3-16　方向接触力为常数时摩擦力与相对滑移量的关系

既与相对滑移速度有关，也与接触面积有关，而且相对滑移速度越大对摩擦因数的影响也越大。对于薄板冲压成形过程来说，相对滑移的速度通常较小，故它对摩擦因数的影响可以忽略不计。

尽管经典摩擦定律有它的不足之处，但它在工程中仍有广泛的应用。因为人们围绕它做了大量的研究工作，对各种材料的接触表面在不同状态下的摩擦因数做了很多实验，并获得大量有工程意义的数据，同时获得了广泛的应用经验。在应用经典摩擦定律时，不必限于图3.3-15所示的接触合力和摩擦合力。经典摩擦定律的结论可扩展到接触应力和摩擦应力上去。但由于在有限元方法中采用了离散法处理接触问题，因此没有必要去使用接触应力和摩擦应力的概念，而可直接讨论接触节点上的法向接触力和摩擦力的关系。

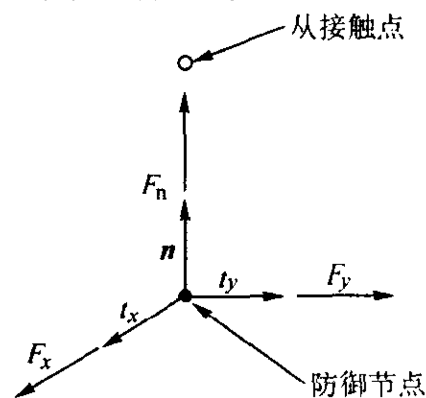


图3.3-17　从接触点和防御节点

如图3.3-17所示，考虑一对从接触点和防御节点。图中，$\boldsymbol{n}$ 为定义在防御节点上的单位法矢，t_x 和 t_y 是经过防御节点的切平面上的单位正交向量，F_n^t、F_x^t 和 F_y^t 分别为法向接触力和摩擦力的分量。单位向量 t_x、t_y 和 n 构成一个右手的局部坐标系。库仑摩擦定律表述如下：

$$\begin{aligned}&\text{定义} f(F_x^t, F_y^t, F_n^t) = \sqrt{(F_x^t)^2 + (F_y^t)^2} - \mu^t F_n^t\\&\text{如果} f(F_x^t, F_y^t, F_n^t) < 0, \text{则 } \mu_x^t t_x + \mu_y^t t_y = 0\\&\text{如果} f(F_x^t, F_y^t, F_n^t) = 0, \text{则 } \mu_x^t t_x + \mu_y^t t_y = \mu_t^t \neq 0\end{aligned} \tag{3.3-205}$$

并有

$$\vec{F}_t^t = F_x^t t_x + F_y^t t_y = -\lambda \mu_t^t \tag{3.3-206}$$

其中，μ^t 为摩擦因数；λ 为一个正的参数。通常来说，摩擦因数是与运动相关的。为简便起见，假定摩擦因数为常数 μ。

设 $t-\Delta t$ 时刻的相关量已知，计算 t 时刻的摩擦。采用基于中心差分的显式方法，计算步骤如下。

1）对于给定的从接触点 i，检查在 $t-\Delta t$ 时刻接触节点是否处于滑移状态。如果是处于滑移状态，则更新从接触点所对应的防御节点的位置，计算 t 时刻新位置下防御节点处的单位向量 t_x、t_y 和 n；否则，防御节点的位置不变。

2）更新防御节点的所有节点量，如节点力和速度。

3）假设从接触点处于黏着状态，计算法向接触力 F_n^t 和摩擦力 F_x^t、F_y^t。

4）检查式（3.3-205）是否满足。如果满足，即黏着接触假设正确，则步骤3计算的摩擦力 F_x^t 和 F_y^t 是准确的，跳到步骤6，否则执行步骤5。

5）计算滑移状态下的摩擦力，由式（3.3-205），将摩擦力作用在与相对滑移方向相反的方向上

$$\vec{F}_t^t = -\mu F_n^t \mu_t^t / |\mu_t^t| \tag{3.3-207}$$

如果从接触点与防御节点之间的相对速度为零，即摩擦力刚好等于临界值，则用步骤3所计算的摩擦力的合力来确定摩擦力的方向。

$$\vec{F}_t^t = \mu F_n^t (F_x^t t_x + \vec{F}_y^t t_y) / |F_x^t t_x + F_y^t t_y| \tag{3.3-208}$$

6）法向接触力和切向摩擦力作用于接触节点，如果还有别的接触点需要计算摩擦力，回到步骤1。

在上述确定接触节点摩擦力的过程中，最重要的一个步骤是计算纯黏着状态下的摩擦力。在运动学上，纯黏着状态相当于在 t_x 和 t_y 方向上的滑动位移增量为零，这种零滑动位移状态与计算法向接触力时的零穿透状态是完全一样的。因此，可以用防御节点法来直接计算摩擦力增量。类似计算法向接触力时法线方向上的零穿透约束，计算摩擦力时，t_x 和 t_y 方向都要施加约束。与上一节计算法向接触力的方法相同，摩擦力分量计算结果如下

$$f_{hx} = -f_{dx} = M_h M_d (F_{dx}^t / M_d - F_{hx}^t / M_h + \mu_{dx}^\tau / \Delta t - \mu_{dx}^\tau / \Delta t - g_x^\tau / \Delta t^2) / (M_h + M_d)$$

$$f_{hy} = -f_{dy} = M_h M_d (F_{dy}^t / M_d - F_{hy}^t / M_h + \mu_{dy}^\tau / \Delta t - \mu_{hy}^\tau / \Delta t - g_y^\tau / \Delta t^2) / (M_h + M_d) \tag{3.3-209}$$

式中，下标 h 和 d 分别对应于从接触点和防御节点；x 和 y 分别表示各物理量在 t_x 和 t_y 方向上的分量；F_{hx}^t 和 F_{dx}^t 分别表示从接触点和防御节点在 t_x 方向上受到的外力（不包括摩擦力）；F_{hy}^t 和 F_{dy}^t 分别表示从接触点和防御节点在 t_y 方向上受到的外力（不包括摩擦力）；g_x^τ 和 g_y^τ 为 τ 时刻相对滑动位移。

假定从接触点在 τ 时刻和 t 时刻都处于黏着接触状态，t 时刻相对滑动位移 g_x^t 和 g_y^t 计算如下

$$\begin{aligned} g_x^t &= (\Delta u_h^t - \Delta u_d^t) t_x \\ g_y^t &= (\Delta u_h^t - \Delta u_d^t) t_y \end{aligned} \tag{3.3-210}$$

式中，Δu_h^t 和 Δu_d^t 分别表示从接触点和防御节点从 τ 时刻到 t 时刻的位移增量。注意由于从接触点从 τ 时刻起处于黏着接触状态，则从 τ 时刻到 t 时刻防御节点的位置保持不变。

4.4　板料冲压成形界面滑动约束处理

采用隐式有限元法模拟板料冲压成形过程时，必须要处理模具与板料间的界面滑动约束。假定模具表面是连续光滑的。在有限元模拟过程中，如果板料节点与模具表面接触或在模具表面切向滑动，并且接触压力大于零，则假设在某一个增量步内板料节点在接触点处模具的切向平面内滑动。

板料节点 N 为接触滑动节点，$\dot{u}_N = \{\dot{u}_N\ \dot{u}_{N+1}\ \dot{u}_{N+2}\}^T$ 为该节点在 X_i 坐标下的滑动速度，$n = (n_1\ n_2\ n_3)$ 为与 N 点接触处冲头的单位法向量。设冲头沿 X_3 方向以 $\dot{w}_p$ 速度行进，则由 $\dot{u}_N$ 与 n 的正交性条件，可得一般性滑动约束方程：

$$\dot{u}_{N+2} = \alpha_1 \dot{u}_N + \alpha_2 \dot{u}_{N+1} + \dot{w}_p \tag{3.3-211}$$

式中，$\alpha_i = -\partial \Psi(X_1\ X_2\ X_3) / \partial X_i (i = 1, 2), \Psi(X_1\ X_2\ X_3) = 0$ 为冲头表面的曲面方程。

将式（3.3-211）代入单元平衡方程，得

$$k_{i1} \dot{u}_1 + \cdots + (k_{iN} + \alpha_1 k_{i,N+2}) \dot{u}_N + (k_{i,N+1} + \alpha_2 k_{i,N+2}) \dot{u}_{N+1} + \cdots + k_{iM} \dot{u}_M = \dot{f}_i - \dot{w}_p k_{i,N+2} \tag{3.3-212}$$

式中，$i = 1, 2, \cdots, M$（M 为单元自由度数）。

经过上面的处理，方程组式（3.3-212）的系数矩阵已经不再对称。为了保持单元平衡方程系数矩阵的对称性，在式（3.3-212）的第 $N+2$ 个方程分别乘以 α_1 和 α_2 后与第 N 和 $N+1$ 个方程相加，再去掉第 $N+2$ 个方程。经过这样的行初等变换后得到的单元矩阵是对称的，因为式（3.3-212）实际上是由有限元方程的系数矩阵做相应的列变换得到的。可以证明，当节点 N 为若干个单元的公共节点时，每个单元都做类似的变换，再组装到构形刚度矩阵之后得到的修正矩阵仍是对称的，而且与在总体平衡方程中直接对 N 节点所在自由度的行、列进行初等变换的结果是相同的。

就计算效率和程序设计考虑，在单元上约束处理更方便。如果求解过程的增量步比较小的话，约束处理是比较精确的。这种方法可以使方程组得到更多符合实际情况的约束处理，使有限元方程组的求解稳定性得到加强。

5　接触搜索判断

5.1　工具形状的定义

板料成形模拟过程中，一般根据压力机的运动方向来选择整体坐标系。将冲压方向定义为 Z 轴方向，向上为正方向。X 和 Y 轴取在与冲压方向相垂直的投影平面内，与 Z 轴一起构成右手直角坐标系。模具的型面用定义域中给定了外法线方向的空间曲面来表示。空间曲面可用如下几种方法表示。

1）点数据描述　将空间曲面用空间中有规则地在该曲面上选取的离散点描述。这些离散点在 XY 平面内的投影落在由平行于 X 轴和 Y 轴的直线所形成的矩形或正方形网格的节点上。这是一种简洁高效的描述方法，但不适于曲面中包含有与 XY 平面垂直或接近垂直的部分的情况。

2）解析曲面　这类曲面包括平面、圆锥面、球面、圆环面、椭圆面等。用它们来定义模具形状时，要注意正确地给出各曲面片的定义域。这种方法无法描述复杂曲面，不是一种通用方法。

3）参数曲面　有些模具型面包含雕塑曲面，在汽车车身零件的冲模中，这是很常见的。这种曲面在 CAD 系统中用贝塞尔曲面、有理 B 样条曲面等参数曲面来描述。考虑到计算效率，在有限元模拟中，一般很少采用参数曲面进行接触判断，它比离散有限元网格法慢数倍。在现有商业软件中都采用离散网格法，很少采用这种方法描述曲面。

4）离散网格　采用有限元网格对曲面进行离散化处理。为了提高接触判断的搜索效率，一般都采用 3 节点三角形单元或 4 节点四边形单元。这种方法描述曲面的精度没有参数曲面高，但是它非常实用，具有普遍性和通用性，与有限元计算流程一致，是成形模拟中最主要的方法。

5.2　接触判断

在成形模拟过程中，板料有限元节点与模具之间的接触状态是不断变化的。在接触区域中要对板料节点施加接触条件约束。因此在每步计算中都要进行接触搜索判断，以便确定板料上的哪些节点与哪个模具上的哪个曲面或单元处于接触状态。

为了提高判断效率，一般将接触判断分为两步。首先进行全局整体搜索，目的是确定位于同一空间子域中的节点和曲面片，将这些节点和曲面片称为可能接触的节点和曲面片。然后再进行局部搜索，以便最终确定可能接触的节点和曲面片是否具的发生接触。

（1）全局搜索

全局搜索是建立在对板料和模具所涉及的空间区域的划分基础之上的。除了对于解析曲面可根据各曲面片的定义域进行划分以外，一般都是采用整体坐标系中的坐标平面或坐标线进行划分的。在一般的三维描述中，用坐标面将曲面的定义域分割成相互连接的立方体。若曲面均不包含 XY 平面的垂直面，也可用坐标线将曲面投影的定义域分割成相互连接的矩形。分割后形成的立方体或矩形可统称为空间子域。若在同一个空间子域同时存在板料上的节点和模具上的曲面片或它们的投影，则它们构形可能接触的节点和曲面片。当采用点数据描述时，全局搜索是很容易的，只需要确定落在同一个投影不矩形格内的节点和曲面片即可。当模具型面不存在垂直面时，无论采用哪种方法定义模具曲面，都可以向 XY 平面投影，在投影面内进行全局搜索。一般情况下不能排除存在垂直面或接近垂直的曲面可能性，这时需要将曲面的定义域分割成立方体。

(2) 局部搜索

局部搜索要完成的任务是：①判断一个可能接触的节点与各模具在同一个空间子域中的诸曲面片中的哪一个最接近，即最有接触的可能性；②该节点是否进入了最接近的曲面片的内部；③如果进入，则求出过该节点法线与曲面片的交点坐标与该节点进入模具的进入量。下面介绍几种局部搜索方法。

1) Z 坐标比较法　当模具型面不存在垂直面时，可以认为与板料上表面接触的模具在每个接触点处，其 Z 坐标是大于或等于板料上节点的 Z 坐标的。

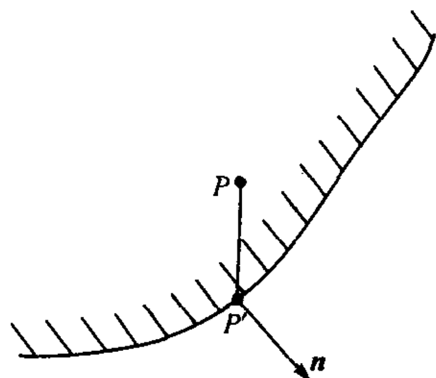


图 3.3-18　Z 坐标比较法

如图 3.3-18 所示，对于板料上的任一节点 P，可以找到 XY 面内与其投影重合的模具曲面上的对应点 P'，当 $Z(P)\geqslant Z(P')$时，认为节点 P 与模具相接触，进入量近似为 $\overline{PP'}$，而对应的工具法矢为 P'点的法矢 n。若 $Z(P)<Z(P')$，认为节点 P 不与模具接触。对于与板料下表面接触的模具可作类似的分析。这种方法简单、计算量少，但所得工具法矢误差较大，而且不能用于有垂直面的情况。

2) 投影法　如图 3.3-19 所示，将节点 P 向曲面片投影。设投影点为 P'，P'点坐标为 X'_0。P 点应该处于通过 P'点且由 P'点的工具法矢 n 所决定的直线上。而 $\left.\frac{\partial X}{\partial \xi}\right|_{x'_0}$ 和 $\left.\frac{\partial X}{\partial \eta}\right|_{x'_0}$ 分别为曲面片上通过 P'点的两条曲线，因此有

$$\begin{aligned}\left.\frac{\partial X}{\partial \xi}\right|_{x'_0}\cdot\overrightarrow{P'P}=0\\ \left.\frac{\partial X}{\partial \eta}\right|_{x'_0}\cdot\overrightarrow{P'P}=0\end{aligned}\tag{3.3-213}$$

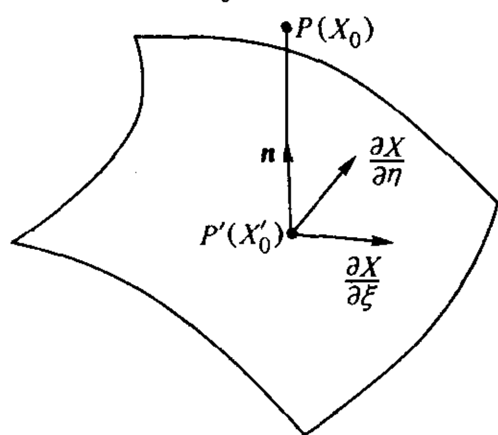


图 3.3-19　投影法

式 (3.3-213) 是关于坐标 x'_0 的非线性方程组，可以用 Newton - Raphson 法求解。求得 x'_0 以后，可以计算 P'点的工具法矢 n。P 点的进入量 g 由下式计算

$$g=n\cdot\overrightarrow{P'P}\tag{3.3-214}$$

当 $g>0$ 时 P 点位于曲面片外侧，不发生接触；当 $g\leqslant 0$ 时发生接触。这种方法是普遍适用的，但式 (3.3-213) 的求解计算量较大。

6　板料成形有限元方法

板料成形有限元方法主要有：

1) 基于 Green 应变和第二类 Kirchhoff 应力能量共轭的虚功原理，以初始时刻为参考构形的全量拉格朗日 (TL) 有限元方法；

2) 基于 Green 应变和第二类 Kirchhoff 应力能量共轭的虚功原理，以当前时刻为参考构形的修正拉格朗日 (UL) 有限元方法；

3) 基于 Lagrange (第一类 Kirchhoff) 应力与速度对物质坐标偏导数能量共轭的虚功原理，以当前时刻为参考构形的虚功率增量型有限元方法；

4) 基于动力学原理和中心差分方法建立的动力显式有限元方法。

6.1　非线性方程组迭代解法

板材冲压成形数值模拟是一个强非线性问题，涉及到几何、材料和边界三重非线性。如果采用隐式有限元法就要求解非线性有限元方程组。非线性方程组一般是采用线性化方法，通过一系列线性解逼近非线性解。但是这种方法是有局限性的，而且有时解的漂移误差很大。因此，一般都采用迭代法求解非线性有限元方程组，即 Newton - Raphson 法。

将非线性有限元方程组写成一般形式，令

$$\begin{aligned}R(u)&=K(u)U\\ F(u)&=R(u)-P(u)\end{aligned}\tag{3.3-215}$$

则一般非线性迭代方程组为

$$F(u)=0\tag{3.3-216}$$

由于$P(u)$中的 $U=0$，所以$F(u)$在u_n 点的一阶导数就是 u_n 点的总体切线刚度矩阵

$$\left.\frac{\partial F(u)}{\partial U}\right|_{u=u_n}=\left.\frac{\partial R(u)}{\partial U}\right|_{u=u_n}=\boldsymbol{K}(u_n)\tag{3.3-217}$$

6.1.1　Newton - Raphson 法

对具有一阶导数连续的函数$F(u)$在 U_n 处作一阶 Taylor 展开，并用 u_n 表 U_n，则它在 U_n 处的线性近似公式为

$$F(u)=F(u_n)+\left.\frac{\partial F(u)}{\partial U}\right|_{u=u_n}(U-U_n)\tag{3.3-218}$$

因此，非线性方程组在 U_n 附近的近似方程组是一个线性方程组，即

$$F(u_n)+\left.\frac{\partial R(u)}{\partial U}\right|_{u=u_n}(U-U_n)=0\tag{3.3-219}$$

假设

$$\begin{aligned}U&=U_{n+1}\\ \Delta U_n&=U_{n+1}-U_n\end{aligned}\tag{3.3-220}$$

因此 Newton - Raphson 法的迭代方程为

$$\begin{aligned}\Delta U_n&=\boldsymbol{K}(u_n)^{-1}(P-R(u_n))\\ U_{n+1}&=U_n+\Delta U_n\end{aligned}\tag{3.3-221}$$

6.1.2　修正的 Newton - Raphson 法

在 Newton - Raphson 法中，刚度矩阵$\boldsymbol{K}(u_n)$是与 u_n 有关的，在每个迭代步都要重新计算一次。为了减少计算量，$\boldsymbol{K}(u_n)$用某一不变的刚度矩阵代替，例如初始刚度矩阵$\boldsymbol{K}(u_0)$或中间某一步刚度矩阵$\boldsymbol{K}(u_k)$。这样就得到了修正的 Newton - Raphson 法迭代方程

$$\begin{aligned}\Delta U_n&=\boldsymbol{K}(u_0)^{-1}(P-R(u_n))\\ U_{n+1}&=U_n+\Delta U_n\end{aligned}\tag{3.3-222}$$

6.1.3　拟 Newton - Raphson 法

Newton - Raphson 法和修正的 Newton - Raphson 法都是用切线刚度矩阵$\boldsymbol{K}(u_n)$进行迭代平衡的。实际计算表明，这种直接迭代法不仅计算量大，而且常常不收敛。因此又提出一种拟 Newton - Raphson 法，用割线刚度矩阵进行迭代平衡。

参考式 (3.3-221) 和式 (3.3-222) 建立刚度矩阵逆矩

阵迭代公式

$$K_s^{-1}(u_n)=K_s^{-1}(u_{n-1})+\Delta K_s^{-1}(u_{n-1}) \quad (3.3\text{-}223)$$

式中，下角标 s 表示割线刚度矩阵。或者简单地记为

$$K_n^{-1}=K_{n-1}^{-1}+\Delta K_{n-1}^{-1} \quad (3.3\text{-}224)$$

1）秩 1 算法　由于任何一个同 $m\times m$ 阶的秩 1 矩阵总可以表示为两个 $m\times 1$ 阶列向量相乘，所以若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 均为 $m\times 1$阶列向量，则

$$\Delta K_{n-1}^{-1}=\boldsymbol{A}\boldsymbol{B}^{\mathrm T} \quad (3.3\text{-}225)$$

由式（3.3-224）可得

$$\boldsymbol{A}\boldsymbol{B}^{\mathrm T}\Delta R_{n-1}=\Delta U_{n-1}-K_{n-1}^{-1}\Delta R_{n-1} \quad (3.3\text{-}226)$$

如果 $\boldsymbol{B}^{\mathrm T}\Delta R_{n-1}\neq 0$，则

$$\boldsymbol{A}=\frac{1}{\boldsymbol{B}^{\mathrm T}\Delta R_{n-1}}(\Delta U_{n-1}-K_{n-1}^{-1}\Delta R_{n-1}) \quad (3.3\text{-}227)$$

则

$$\Delta K_{n-1}^{-1}=\alpha(\Delta U_{n-1}-K_{n-1}^{-1}\Delta R_{n-1})\boldsymbol{B}^{\mathrm T} \quad (3.3\text{-}228)$$

式中

$$\alpha=\begin{cases}\dfrac{1}{\boldsymbol{B}^{\mathrm T}\Delta R_{n-1}} & 当\ \Delta R_{n-1}\neq 0\\ 0 & 当\ \Delta R_{n-1}=0\end{cases} \quad (3.3\text{-}229)$$

如果取

$$\boldsymbol{B}^{\mathrm T}=\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1} \quad (3.3\text{-}230)$$

则

$$\Delta K_{n-1}^{-1}=(\Delta U_{n-1}-K_{n-1}^{-1}\Delta R_{n-1})\frac{\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}}{\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1}}(\Delta R=0) \quad (3.3\text{-}231)$$

如果取

$$\boldsymbol{B}=\Delta U_{n-1}-K_{n-1}^{-1}\Delta R_{n-1} \quad (3.3\text{-}232)$$

则

$$\Delta K_{n-1}^{-1}=(\Delta U_{n-1}-K_{n-1}^{-1}\Delta R_{n-1})\frac{(\Delta U_{n-1}-K_{n-1}^{-1}\Delta R_{n-1})^{\mathrm T}}{(\Delta U_{n-1}-K_{n-1}^{-1}\Delta R_{n-1})^{\mathrm T}\Delta R_{n-1}}$$

$$(\Delta U_{n-1}\neq K_{n-1}^{-1}\Delta R_{n-1}) \quad (3.3\text{-}233)$$

式（3.3-231）是非对称秩 1 算法，式（3.3-233）是对称秩 1 算法。

2）秩 2 算法　由于任何一个 $m\times m$ 阶的秩 2 矩阵可以表示为

$$\Delta K_{n-1}^{-1}=A_1B_1^{\mathrm T}+A_2B_2^{\mathrm T} \quad (3.3\text{-}234)$$

按照秩 1 算法可得

$$\Delta K_{n-1}^{-1}=\alpha_1\Delta U_{n-1}B_1^{\mathrm T}+\alpha_2K_{n-1}^{-1}\Delta R_{n-1}B_2^{\mathrm T} \quad (3.3\text{-}235)$$

式中

$$\alpha_1=\begin{cases}\dfrac{1}{B_1^{\mathrm T}\Delta R_{n-1}} & 当\ \Delta R_{n-1}\neq 0\\ 0 & 当\ \Delta R_{n-1}=0\end{cases} \quad (3.3\text{-}236)$$

$$\alpha_2=\begin{cases}\dfrac{1}{B_2^{\mathrm T}\Delta R_{n-1}} & 当\ \Delta R_{n-1}\neq 0\\ 0 & 当\ \Delta R_{n-1}=0\end{cases} \quad (3.3\text{-}237)$$

如果假设

$$\bar B_1^{\mathrm T}=\frac{B_1^{\mathrm T}}{B_1^{\mathrm T}\Delta R_{n-1}};\ \bar B_2^{\mathrm T}=\frac{B_2^{\mathrm T}}{B_2^{\mathrm T}\Delta R_{n-1}} \quad (3.3\text{-}238)$$

则有

$$\bar B_1^{\mathrm T}\Delta R_{n-1}=\bar B_2^{\mathrm T}\Delta R_{n-1}=1 \quad (3.3\text{-}239)$$

于是可得

$$K_{n-1}^{-1}=\Delta U_{n-1}\bar B_1^{\mathrm T}-K_{n-1}^{-1}\Delta R_{n-1}\bar B_2^{\mathrm T} \quad (3.3\text{-}240)$$

引进参数 β，并把 $\bar B_1^{\mathrm T}$ 和 $\bar B_2^{\mathrm T}$ 取为如下的组合形式

$$\bar B_1^{\mathrm T}=(1+\beta\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1})\frac{\Delta U_{n-1}^{\mathrm T}}{\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}}-\beta\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}$$

$$(\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}\neq 0) \quad (3.3\text{-}241)$$

$$\bar B_2^{\mathrm T}=(1+\beta\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1})\frac{\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}}{\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1}}-\beta\Delta U_{n-1}^{\mathrm T}$$

$$(\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1}\neq 0) \quad (3.3\text{-}242)$$

因此

$$\begin{aligned}\Delta K_{n-1}^{-1}=&\frac{\Delta U_{n-1}\Delta U_{n-1}^{\mathrm T}}{\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}}-\frac{K_{n-1}^{-1}\Delta R_{n-1}\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}}{\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1}}+\\&\beta(\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1}\frac{\Delta U_{n-1}\Delta U_{n-1}^{\mathrm T}}{\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}}+\\&\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}\frac{K_{n-1}^{-1}\Delta R_{n-1}\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}}{\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1}}-\\&\Delta U_{n-1}\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}-K_{n-1}^{-1}\Delta R_{n-1}\Delta U_{n-1}^{\mathrm T})\end{aligned} \quad (3.3\text{-}243)$$

式（3.3-243）是对称秩 2 算法。

参数 β 的不同选法可以得到不同的迭代算法。

Ⅰ）DFP（Davidon Fletcher Powell）算法　如果取 $\beta=0$，就得到 DFP 算法

$$\Delta K_{n-1}^{-1}=\frac{\Delta U_{n-1}\Delta U_{n-1}^{\mathrm T}}{\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}}-\frac{K_{n-1}^{-1}\Delta R_{n-1}\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}}{\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1}} \quad (3.3\text{-}244)$$

Ⅱ）BFS（Broyden Fletcher Shanno）算法　如果取 $\beta=\dfrac{1}{\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}}$，就得到 BFS 算法

$$\Delta K_{n-1}^{-1}=\frac{\alpha\Delta U_{n-1}\Delta U_{n-1}^{\mathrm T}-\Delta U_{n-1}\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}-K_{n-1}^{-1}\Delta R_{n-1}\Delta U_{n-1}^{\mathrm T}}{\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}} \quad (3.3\text{-}245)$$

式中

$$\alpha=1+\frac{\Delta R_{n-1}^{\mathrm T}K_{n-1}^{-1}\Delta R_{n-1}}{\Delta U_{n-1}^{\mathrm T}\Delta R_{n-1}} \quad (3.3\text{-}246)$$

一般 BFS 算法比 DFP 算法有较好的数值稳定性。

6.1.4 迭代收敛准则

求解非线性方程组时，必须要给出迭代收敛准则，判别迭代过程是否收敛，以便终止迭代。如果收敛准则给的不合适，可能计算精度不高或太浪费机时，有时甚至计算失败。

1）位移收敛准则　常用位移收敛准则为

$$\|\Delta U_n\|\leqslant\alpha\|U_n\| \quad (3.3\text{-}247)$$

式中，$\|*\|$ 一般取 Euclid 范数 $\|*\|_2$；α 是位移收敛准则的容差常数，它是一个小量，一般取

$$0.1\%\leqslant\alpha\leqslant 5\% \quad (3.3\text{-}248)$$

当系统含有刚体位移时，$\|\Delta U_n\|$ 会比较大，此时不适合采用位移收敛准则。

2）失衡力收敛准则　常用失衡力收敛准则为

$$\|\Delta F_n\|\leqslant\alpha\|P\| \quad (3.3\text{-}249)$$

式中，ΔF_n 为第 n 迭代步的失衡力

$$\Delta F_n=P-R(u_n) \quad (3.3\text{-}250)$$

当系统处于失稳状态时，失衡力的微小变化将引起位移增量的很大偏差，此时不能采用失衡力收敛准则。

3）能量收敛准则　位移收敛准则和失衡力收敛准则都有一定的缺陷，相对而言能量收敛准则是比较好的，因为它同时控制位移增量和失衡力。能量收敛准则是把每次迭代后的内能增量与初始内能增量相比较，即

$$\Delta U_n^{\mathrm T}(P-R(u_n))\leqslant\alpha\Delta U_1^{\mathrm T}(P-R(u_1)) \quad (3.3\text{-}251)$$

式中，α 为能量收敛准则的容差常数，它是一个小量。

6.2 板料成形全量拉格朗日（TL）有限元方法

6.2.1 全量拉格朗日虚功方程

TL 法是取初始时刻构形作为参考构形，在所有的时间步长内的计算都参照时刻 $t_0=0$ 构形来定义。

则 $t+\Delta t$ 时刻的虚功方程为

$$\int_{V_0}\delta\bar E^{\mathrm T}\bar S\,\mathrm dV_0=\int_{V_0}\delta\bar u^{\mathrm T}\bar p_0\,\mathrm dV_0+\int_{A_0}\delta\bar u^{\mathrm T}\bar q_0\,\mathrm dA_0 \quad (3.3\text{-}252)$$

式中，$\bar{E}$ 为 $t+\Delta t$ 时刻的 Green 应变；$\bar{S}$ 为 $t+\Delta t$ 时刻的第二类 Kirchhoff 应力；$\bar{u}$ 为 $t+\Delta t$ 时刻的可容位移；$\bar{p}_0$ 和 $\bar{q}_0$ 分别为时刻 $t+\Delta t$ 的体力和面力载荷向量，它们都是定义在初始构形上的已知边界条件；V_0 和 A_0 分别为构形初始时刻的体积和表面积。

6.2.2 单元插值关系

单元的插值关系假设为

$$\begin{aligned}\Delta \boldsymbol{u} &= \boldsymbol{N}\Delta \boldsymbol{u}^e \\ \boldsymbol{u} &= \boldsymbol{N}\boldsymbol{u}^e\end{aligned} \tag{3.3-253}$$

式中，$\Delta\boldsymbol{u}$ 和 $\Delta\boldsymbol{u}^e$ 分别为时刻 t 到时刻 $t+\Delta t$ 之间的位移增量和单元节点位移增量向量；$\boldsymbol{u}$ 和 $\boldsymbol{u}^e$ 分别为时刻 t_0 到时刻 t 之间的位移量和单元节点位移量向量；$\boldsymbol{N}$ 为单元形函数矩阵，由单元形函数确定，是当前时刻坐标 X_i 的函数。

6.2.3 单元本构关系

t 时刻单元的本构关系假设为

$$\Delta S = \boldsymbol{D}_T \Delta E \tag{3.3-254}$$

式中，ΔS 为时间增量步 Δt 内第二类 Kirchhoff 应力增量；ΔE 为时间增量步 Δt 内 Green 应变增量；$\boldsymbol{D}_T$ 为切线本构矩阵。

6.2.4 单元几何关系

t 时刻和 $t+\Delta t$ 时刻的单元几何关系分别为

$$E_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial X_i} + \frac{\partial u_i}{\partial X_j} + \frac{\partial u_k}{\partial X_i}\frac{\partial u_k}{\partial X_j}\right) \tag{3.3-255}$$

$$\bar{E}_{ij} = \frac{1}{2}\left(\frac{\partial \bar{u}_j}{\partial X_i} + \frac{\partial \bar{u}_i}{\partial X_j} + \frac{\partial \bar{u}_k}{\partial X_i}\frac{\partial \bar{u}_k}{\partial X_j}\right) \tag{3.3-256}$$

$\bar{E}_{ij}$可分解为

$$\bar{E}_{ij} = E_{ij} + \Delta E_{ij} \tag{3.3-257}$$

式中，ΔE_{ij}为时间增量步 Δt 内的 Green 应变增量，它可以表示为

$$\Delta E_{ij} = \Delta E_{ij}^{L_0} + \Delta E_{ij}^{L_1} + \Delta E_{ij}^{N} \tag{3.3-258}$$

$$\Delta E_{ij}^{L_0} = \frac{1}{2}\left(\frac{\partial \Delta u_j}{\partial X_i} + \frac{\partial \Delta u_i}{\partial X_j}\right) \tag{3.3-259}$$

$$\Delta E_{ij}^{L_1} = \frac{1}{2}\left(\frac{\partial u_k}{\partial X_i}\frac{\partial \Delta u_k}{\partial X_i} + \frac{\partial \Delta u_k}{\partial X_i}\frac{\partial u_k}{\partial X_j}\right) \tag{3.3-260}$$

$$\Delta E_{ij}^{N} = \frac{1}{2}\frac{\partial \Delta u_k}{\partial X_i}\frac{\partial \Delta u_k}{\partial X_j} \tag{3.3-261}$$

在增量求解过程中，时刻 t 的位移 u_i 是已知的，式 (3.3-259)、式 (3.3-260) 中 $\Delta E_{ij}^{L_0}$ 和 $\Delta E_{ij}^{L_1}$ 与未知量 Δu_i 呈线性关系，因此它们是 Green 应变增量的线性部分，而 ΔE_{ij}^{N}是非线性部分。式 (3.3-257) ~式 (3.3-261) 的矩阵表达形式分别为

$$\bar{\boldsymbol{E}} = \boldsymbol{E} + \Delta\boldsymbol{E} \tag{3.3-262}$$

$$\Delta\boldsymbol{E} = \Delta\boldsymbol{E}^{L_0} + \Delta\boldsymbol{E}^{L_1} + \Delta\boldsymbol{E}^{N} \tag{3.3-263}$$

$$\Delta\boldsymbol{E}^{L_0} = \boldsymbol{L}\Delta\boldsymbol{u} \tag{3.3-264}$$

$$\Delta\boldsymbol{E}^{L_1} = \frac{1}{2}\boldsymbol{A}\Delta\boldsymbol{\theta} + \frac{1}{2}\Delta\boldsymbol{A}\boldsymbol{\theta} = \boldsymbol{A}\Delta\boldsymbol{\theta} \tag{3.3-265}$$

$$\Delta\boldsymbol{E}^{N} = \frac{1}{2}\Delta\boldsymbol{A}\Delta\boldsymbol{\theta} \tag{3.3-266}$$

式中

$$\boldsymbol{L} = \begin{bmatrix} \frac{\partial}{\partial X_1} & 0 & 0 \\ 0 & \frac{\partial}{\partial X_2} & 0 \\ 0 & 0 & \frac{\partial}{\partial X_3} \\ 0 & \frac{\partial}{\partial X_3} & \frac{\partial}{\partial X_2} \\ \frac{\partial}{\partial X_3} & 0 & \frac{\partial}{\partial X_1} \\ \frac{\partial}{\partial X_2} & \frac{\partial}{\partial X_1} & 0 \end{bmatrix} \tag{3.3-267}$$

$$\boldsymbol{A} = \begin{bmatrix} \frac{\partial u^T}{\partial X_1} & 0 & 0 \\ 0 & \frac{\partial u^T}{\partial X_2} & 0 \\ 0 & 0 & \frac{\partial u^T}{\partial X_3} \\ 0 & \frac{\partial u^T}{\partial X_3} & \frac{\partial u^T}{\partial X_2} \\ \frac{\partial u^T}{\partial X_3} & 0 & \frac{\partial u^T}{\partial X_1} \\ \frac{\partial u^T}{\partial X_2} & \frac{\partial u^T}{\partial X_1} & 0 \end{bmatrix} \tag{3.3-268}$$

$$\boldsymbol{\theta} = \begin{Bmatrix} \frac{\partial u}{\partial X_1} \\ \frac{\partial u}{\partial X_2} \\ \frac{\partial u}{\partial X_3} \end{Bmatrix} \tag{3.3-269}$$

将单元插值关系式 (3.3-253) 代入式 (3.3-264) ~式 (3.3-266) 可得

$$\Delta\boldsymbol{E}^{L_0} = \boldsymbol{B}_{L_0}\Delta\boldsymbol{u}^e \tag{3.3-270}$$

$$\Delta\boldsymbol{E}^{L_1} = \boldsymbol{B}_{L_1}\Delta\boldsymbol{u}^e \tag{3.3-271}$$

$$\Delta\boldsymbol{E}^{N} = \frac{1}{2}\boldsymbol{B}_N\Delta\boldsymbol{u}^e \tag{3.3-272}$$

式中

$$\boldsymbol{B}_{L_0} = \boldsymbol{L}\boldsymbol{N} \tag{3.3-273}$$

$$\boldsymbol{B}_{L_1} = \boldsymbol{A}\boldsymbol{H}\boldsymbol{N} = \boldsymbol{A}\boldsymbol{G} \tag{3.3-274}$$

$$\boldsymbol{B}_N = \Delta\boldsymbol{A}\boldsymbol{G} \tag{3.3-275}$$

6.2.5 单元平衡方程

在增量求解过程中，时刻 t 的位移 u 和 Green 应变 E 都是已知的，因此虚功方程式 (3.3-252) 中

$$\delta\bar{\boldsymbol{u}} = \delta\Delta\boldsymbol{u} = \boldsymbol{N}\delta\Delta\boldsymbol{u}^e \tag{3.3-276}$$

$$\delta\bar{\boldsymbol{E}} = \delta\Delta\boldsymbol{E} = \boldsymbol{B}\delta\Delta\boldsymbol{u}^e \tag{3.3-277}$$

式中

$$\boldsymbol{B} = \boldsymbol{B}_{L_0} + \boldsymbol{B}_{L_1} + \boldsymbol{B}_N \tag{3.3-278}$$

再考虑可容位移 $\Delta\boldsymbol{u}^e$ 的任意性，虚功方程式 (3.3-252) 可表示为

$$\int_{V_0}\boldsymbol{B}^T\bar{\boldsymbol{S}}\mathrm{d}V_0 = \int_{V_0}\boldsymbol{N}^T\bar{\boldsymbol{p}}_0\mathrm{d}V_0 + \int_{A_0}\boldsymbol{N}^T\bar{\boldsymbol{q}}_0\mathrm{d}A_0 \tag{3.3-279}$$

将 $t+\Delta t$ 时刻的第二类 Kirchhoff 应力 $\bar{S}$ 分别成 t 时刻与 Δt 增量之和，即

$$\bar{\boldsymbol{S}} = \boldsymbol{S} + \Delta\boldsymbol{S} \tag{3.3-280}$$

则虚功方程式 (3.3-279) 可进一步表示为增量形式

$$\int_{V_0}\boldsymbol{B}^T\Delta\boldsymbol{S}\mathrm{d}V_0 + \int_{V_0}\boldsymbol{B}_N^T\boldsymbol{S}\mathrm{d}V_0 + \int_{V_0}(\boldsymbol{B}_{L_0}^T + \boldsymbol{B}_{L_1}^T)\boldsymbol{S}\mathrm{d}V_0 = \bar{\boldsymbol{R}} \tag{3.3-281}$$

式中

$$\bar{R}=\int_{V_0}N^{\mathrm{T}}\bar{p}_0\mathrm{d}V_0+\int_{A_0}N^{\mathrm{T}}\bar{q}_0\mathrm{d}A_0 \tag{3.3-282}$$

令

$$M=\begin{bmatrix} S_{11}I & S_{12}I & S_{13}I \\ S_{21}I & S_{22}I & S_{23}I \\ S_{31}I & S_{32}I & S_{33}I \end{bmatrix} \tag{3.3-283}$$

则

$$B_N^{\mathrm{T}}S=G^{\mathrm{T}}\Delta A^{\mathrm{T}}S=G^{\mathrm{T}}M\Delta\theta=G^{\mathrm{T}}MG\Delta u^{e} \tag{3.3-284}$$

如果忽略几何关系中的二阶小项，则

$$\Delta E=\Delta E^{L_0}+\Delta E^{L_1}+\Delta E^{N}\approx\Delta E^{L_0}+\Delta E^{L_1} \tag{3.3-285}$$

因此，虚功方程式（3.3-281）中第一项积分中应变矩阵 B 可以近似为

$$B\approx B_{L_0}+B_{L_1} \tag{3.3-286}$$

这样，由虚功方程式（3.3-281）可以得到线性化的TL法有限元平衡方程

$$K_{\mathrm{T}}\Delta u^{e}=(K_{\mathrm{L}}+K_{s})\Delta u^{e}=\bar{R}-R_{s} \tag{3.3-287}$$

式中，K_T 为 t 时刻的单元切线刚度矩阵

$$K_{\mathrm{T}}=K_{\mathrm{L}}+K_{s} \tag{3.3-288}$$

K_s 为初应力单元刚度矩阵

$$K_{s}=\int_{V_0}G^{\mathrm{T}}MG\mathrm{d}V_0 \tag{3.3-289}$$

$$K_{\mathrm{L}}=\int_{V_0}(B_{L_0}^{\mathrm{T}}+B_{L_1}^{\mathrm{T}})D_{\mathrm{T}}(B_{L_0}+B_{L_1})\mathrm{d}V_0 \tag{3.3-290}$$

$$R_{s}=\int_{V_0}(B_{L_0}^{\mathrm{T}}+B_{L_1}^{\mathrm{T}})S\mathrm{d}V_0 \tag{3.3-291}$$

6.3 板料成形修正拉格朗日（UL）有限元方法

6.3.1 修正拉格朗日虚功方程

UL法是取当前时刻构形作为参考构形，在所有的时间步长内的计算都参照当前时刻 t 构形来定义。则 $t+\Delta t$ 时刻的虚功方程可以表示为

$$\int_{V}\delta\bar{E}^{\mathrm{T}}\bar{S}\mathrm{d}V=\int_{V}\delta\bar{u}^{\mathrm{T}}\bar{p}\mathrm{d}V+\int_{A}\delta\bar{u}^{\mathrm{T}}\bar{q}\mathrm{d}A \tag{3.3-292}$$

式中，$\bar{E}$ 为 $t+\Delta t$ 时刻的Green应变；$\bar{S}$ 为 $t+\Delta t$ 时刻的第二类Kirchhoff应力；$\bar{u}$ 为 $t+\Delta t$ 时刻的可容位移；$\bar{p}$ 和 $\bar{q}$ 分别为时刻 $t+\Delta t$ 的体力和面力载荷向量，它们都是定义在 t 时刻构形上的已知边界条件；V 和 A 分别为 t 时刻构形的体积和表面积。

6.3.2 单元插值关系

单元的插值关系假设为

$$\Delta u=N\Delta u^{e} \tag{3.3-293}$$

式中，Δu 和 Δu^e 分别为时刻 t 到时刻 $t+\Delta t$ 之间的位移增量和单元节点位移增量向量；N 为单元形函数矩阵，由单元形函数确定，是 t 时刻坐标 x 的函数。

在一个时间增量步 Δt 内的位移增量 Δu 为

$$\Delta u=\bar{x}-x \tag{3.3-294}$$

$t+\Delta t$ 时的位移量 $\bar{u}$ 与 Δu 的关系为

$$\bar{u}=\Delta u \tag{3.3-295}$$

6.3.3 单元本构关系

t 时刻单元的本构关系假设为

$$\Delta S=D_{\mathrm{T}}\Delta E \tag{3.3-296}$$

式中，ΔS 为时间增量步 Δt 内第二类Kirchhoff应力增量；ΔE 为时间增量步 Δt 内Green应变增量；D_T 为切线本构矩阵或弹塑性本构矩阵。本构关系式（3.3-296）与式（3.3-254）形式上是一样的，但它们完全是不同的，式（3.3-254）是以 t_0 时刻为参考构形定义的，而式（3.3-296）是以 t 时刻为参考构形定义的。以 t 时刻为参考构形，则Green应变率 $\dot{E}_{ij}$ 与变形率张量 d_{ij} 相等，即

$$\dot{E}_{ij}=d_{ij} \tag{3.3-297}$$

根据第二类Kirchhoff应力率 $\dot{S}_{ij}$ 与Cauchy应力 σ_{ij} 的关系

$$\dot{S}_{ij}=\overset{\nabla}{\sigma}_{ij}-\sigma_{ik}d_{kj}-\sigma_{jk}d_{ki}+\sigma_{ij}v_{m,m} \tag{3.3-298}$$

式中，$\overset{\nabla}{\sigma}_{ij}$ 为Cauchy应力的Jaumann导数张量。

考虑金属材料的不可压缩性，$v_{m,m}=0$，因此 $\dot{S}_{ij}$ 与 $\dot{E}_{ij}$ 的本构关系为

$$\dot{S}_{ij}=D_{ijkl}\dot{E}_{kl} \tag{3.3-299}$$

式中，本构张量 D_{ijkl} 为

$$D_{ijkl}=D_{ijkl}^{ep}-S_{ik}\delta_{lj}-S_{jk}\delta_{li} \tag{3.3-300}$$

D_{ijkl}^{ep} 为材料弹塑性本构张量

$$\overset{\nabla}{\sigma}_{ij}=D_{ijkl}^{ep}d_{kl} \tag{3.3-301}$$

式（3.3-299）的线性化方程为

$$\Delta S_{ij}=D_{ijkl}^{ep}\Delta E_{kl}-S_{ik}\Delta E_{kj}-S_{jk}\Delta E_{ki} \tag{3.3-302}$$

由于 t 和 $t+\Delta t$ 时刻的第二类Kirchhoff应力都是相对于时刻 t 为参考构形定义的，因此 t 时刻的第二类Kirchhoff应力 S 与 t 时刻的Cauchy应力 σ 相等，即

$$S=\sigma \tag{3.3-303}$$

$t+\Delta t$ 时刻的第二类Kirchhoff应力 $\bar{S}$ 可以分解为

$$\bar{S}=S+\Delta S=\sigma+\Delta S \tag{3.3-304}$$

式（3.3-302）可以进一步表示为

$$\Delta S_{ij}=D_{ijkl}^{ep}\Delta E_{kl}-\sigma_{ik}\Delta E_{kj}-\sigma_{jk}\Delta E_{ki} \tag{3.3-305}$$

式（3.3-305）的矩阵形式为

$$\Delta S=(D^{ep}-F)\Delta E \tag{3.3-306}$$

式中，F 为一个对称的用应力表示的矩阵。如果假设 ΔS 向量和 ΔE 向量分别为

$$\Delta S=(\Delta S_{11}\ \Delta S_{22}\ \Delta S_{33}\ \Delta S_{12}\ \Delta S_{23}\ \Delta S_{31})^{\mathrm{T}}$$
$$\Delta E=(\Delta E_{11}\ \Delta E_{22}\ \Delta E_{33}\ \Delta E_{12}\ \Delta E_{23}\ \Delta E_{31})^{\mathrm{T}} \tag{3.3-307}$$

则

$$F=\begin{bmatrix} 2\sigma_{11} & 0 & 0 & \sigma_{12} & 0 & \sigma_{31} \\ 0 & 2\sigma_{22} & 0 & \sigma_{12} & \sigma_{23} & 0 \\ 0 & 0 & 2\sigma_{33} & 0 & \sigma_{23} & \sigma_{31} \\ \sigma_{12} & \sigma_{12} & 0 & \frac{1}{2}(\sigma_{11}+\sigma_{22}) & \frac{1}{2}\sigma_{31} & \frac{1}{2}\sigma_{23} \\ 0 & \sigma_{23} & \sigma_{23} & \frac{1}{2}\sigma_{31} & \frac{1}{2}(\sigma_{22}+\sigma_{33}) & \frac{1}{2}\sigma_{12} \\ \sigma_{31} & 0 & \sigma_{31} & \frac{1}{2}\sigma_{23} & \frac{1}{2}\sigma_{12} & \frac{1}{2}(\sigma_{11}+\sigma_{33}) \end{bmatrix} \tag{3.3-308}$$

6.3.4 单元几何关系

t 时刻和 $t+\Delta t$ 时刻的单元几何关系分别为

$$E_{ij}=0$$

$$\bar{E}_{ij}=\frac{1}{2}\left(\frac{\partial\Delta u_j}{\partial x_i}+\frac{\partial\Delta u_i}{\partial x_j}+\frac{\partial\Delta u_k}{\partial x_i}\frac{\partial\Delta u_k}{\partial x_j}\right) \tag{3.3-309}$$

因此，Green应变增量 ΔE_{ij} 可以表示为

$$\Delta E_{ij} = \bar{E}_{ij} = \Delta E_{ij}^{L} + \Delta E_{ij}^{N} \tag{3.3-310}$$

式中

$$\Delta E_{ij}^{L} = \frac{1}{2}\left(\frac{\partial \Delta u_j}{\partial x_i} + \frac{\partial \Delta u_i}{\partial x_j}\right)$$

$$\Delta E_{ij}^{N} = \frac{1}{2}\frac{\partial \Delta u_k}{\partial x_i}\frac{\partial \Delta u_k}{\partial x_j} \tag{3.3-311}$$

式（3.3-310）和式（3.3-311）用矩阵形式表示为

$$\Delta \boldsymbol{E} = \Delta \boldsymbol{E}_{L} + \Delta \boldsymbol{E}_{N} \tag{3.3-312}$$

$$\Delta \boldsymbol{E}_{L} = \boldsymbol{L}\Delta u \tag{3.3-313}$$

$$\Delta \boldsymbol{E}_{N} = \frac{1}{2}\Delta \boldsymbol{A}\Delta \boldsymbol{\theta} = \frac{1}{2}\Delta \boldsymbol{A}\boldsymbol{H}\Delta \boldsymbol{u} \tag{3.3-314}$$

式中

$$\boldsymbol{L} = \begin{bmatrix} \frac{\partial}{\partial x_1} & 0 & 0 \\ 0 & \frac{\partial}{\partial x_2} & 0 \\ 0 & 0 & \frac{\partial}{\partial x_3} \\ 0 & \frac{\partial}{\partial x_3} & \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} & 0 & \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1} & 0 \end{bmatrix} \tag{3.3-315}$$

$$\Delta \boldsymbol{A} = \begin{bmatrix} \frac{\partial \Delta u^{T}}{\partial x_1} & 0 & 0 \\ 0 & \frac{\partial \Delta u^{T}}{\partial x_2} & 0 \\ 0 & 0 & \frac{\partial \Delta u^{T}}{\partial x_3} \\ 0 & \frac{\partial \Delta u^{T}}{\partial x_3} & \frac{\partial \Delta u^{T}}{\partial x_2} \\ \frac{\partial \Delta u^{T}}{\partial x_3} & 0 & \frac{\partial \Delta u^{T}}{\partial x_1} \\ \frac{\partial \Delta u^{T}}{\partial x_2} & \frac{\partial \Delta u^{T}}{\partial x_1} & 0 \end{bmatrix} \tag{3.3-316}$$

$$\Delta \boldsymbol{\theta} = \begin{Bmatrix} \frac{\partial \Delta u}{\partial x_1} \\ \frac{\partial \Delta u}{\partial x_2} \\ \frac{\partial \Delta u}{\partial x_3} \end{Bmatrix} \tag{3.3-317}$$

$$\boldsymbol{H} = \begin{Bmatrix} I\frac{\partial}{\partial x_1} \\ I\frac{\partial}{\partial x_2} \\ I\frac{\partial}{\partial x_3} \end{Bmatrix} \tag{3.3-318}$$

将单元插值关系式（3.3-293）代入单元几何关系式（3.3-313）和式（3.3-314）中，得

$$\Delta \boldsymbol{E}_{L} = \boldsymbol{B}_{L}\Delta \boldsymbol{u}^{e} \tag{3.3-319}$$

$$\Delta \boldsymbol{E}_{N} = \tilde{\boldsymbol{B}}_{N}\Delta \boldsymbol{u}^{e} \tag{3.3-320}$$

式中

$$\boldsymbol{B}_{L} = \boldsymbol{L}\boldsymbol{N} \tag{3.3-321}$$

$$\tilde{\boldsymbol{B}}_{N} = \frac{1}{2}\Delta \boldsymbol{A}\boldsymbol{H}\boldsymbol{N} = \frac{1}{2}\Delta \boldsymbol{A}\boldsymbol{G} \tag{3.3-322}$$

6.3.5 单元平衡方程

由式（3.3-319）、式（3.3-320）和式（3.3-322）可得

$$\delta\Delta \boldsymbol{E}_{L} = \boldsymbol{B}_{L}\delta\Delta \boldsymbol{u}^{e} \tag{3.3-323}$$

$$\delta\Delta \boldsymbol{E}_{N} = 2\tilde{\boldsymbol{B}}_{N}\delta\Delta \boldsymbol{u}^{e} = \boldsymbol{B}_{N}\delta\Delta \boldsymbol{u}^{e} \tag{3.3-324}$$

$$\delta\Delta \boldsymbol{E} = \boldsymbol{B}\Delta \boldsymbol{u}^{e} = (\boldsymbol{B}_{L} + \boldsymbol{B}_{N})\,\delta\Delta \boldsymbol{u}^{e} \tag{3.3-325}$$

将式（3.3-293）、式（3.3-325）、式（3.3-302）和式（3.3-304）代入虚功方程式（3.3-292），得

$$(\boldsymbol{K}_{a} + \boldsymbol{K}_{s})\Delta \boldsymbol{u}^{e} + \boldsymbol{R}_{s} - \bar{\boldsymbol{R}} = 0 \tag{3.3-326}$$

式中，$\boldsymbol{K}_s$ 为初应力单元刚度矩阵

$$\boldsymbol{K}_{s} = \int_{V}\boldsymbol{G}^{T}\boldsymbol{M}\boldsymbol{G}\mathrm{d}V \tag{3.3-327}$$

$\boldsymbol{R}_s$ 为初应力节点力向量

$$\boldsymbol{R}_{s} = \int_{V}\boldsymbol{B}_{L}^{T}\sigma\mathrm{d}V \tag{3.3-328}$$

$$\bar{\boldsymbol{R}} = \int_{V}\boldsymbol{N}^{T}\bar{\boldsymbol{p}}\mathrm{d}V + \int_{A}\boldsymbol{N}^{T}\bar{\boldsymbol{q}}\mathrm{d}A \tag{3.3-329}$$

$$\boldsymbol{K}_{a} = \int_{V}\boldsymbol{B}^{T}(\boldsymbol{D}^{ep} - \boldsymbol{F})\boldsymbol{B}\mathrm{d}V \tag{3.3-330}$$

6.4 板料成形虚功率增量型有限元方法

6.4.1 虚功率方程

由于 Lagrange（第一类 Kirchhoff）应力与速度对物质坐标偏导数是能量共轭的，可得时刻 t 现实构形，以初始时刻 t_0 为参考构形的弹塑性大变形虚功方程为

$$\int_{V}t_{Ij}\delta\left(\frac{\partial v_j}{\partial X_I}\right)\mathrm{d}V = \int_{V}P_I\delta v_i\mathrm{d}V + \int_{A}\bar{P}_I\delta v_i\mathrm{d}A \tag{3.3-331}$$

式中，V、A 分别表示 t 时刻构形的体积和表面积；$\bar{P}_I$、P_I 分别表示 t 时刻构形的面积率、体积力率；t_{Ij} 表示 t 时刻构形的第一类 Kirchhoff 应力率。

如果采用逐级更新的持续平衡有限元法，取当前时刻 t 为参考构形，则

$$t_{Ij}\to t_{ij},\ X_I\to x_i,\ P_I\to p_i$$
$$\bar{P}_I\to\bar{p}_i,\ V\to v,\ A\to a \tag{3.3-332}$$

将式（3.3-331）应用于 t 时刻和 $t+\Delta t$ 时刻，对于离散后构形上的某任意单元 e 有

$$\int_{v_e}t_{ij}\delta\left(\frac{\partial v_j}{\partial x_i}\right)\mathrm{d}v_e = \int_{v_e}p_i\delta v_i\mathrm{d}v_e + \int_{a_e}\bar{p}_i\delta v_i\mathrm{d}a_e$$

$$\int_{v_e}(t_{ij}+\mathrm{d}t_{ij})\delta\left(\frac{\partial v_j}{\partial x_i}\right)\mathrm{d}v_e = \int_{v_e}(p_i+\mathrm{d}p_i)\delta v_i\mathrm{d}v_e + \int_{a_e}(\bar{p}_i+\mathrm{d}\bar{p}_i)\,\delta v_i\mathrm{d}a_e \tag{3.3-333}$$

式中，v_e、a_e 分别表示时刻 t 单元 e 的体积和表面积。

式（3.3-331）和式（3.3-333）相减得

$$\int_{v_e}\mathrm{d}t_{ij}\delta\left(\frac{\partial v_j}{\partial x_i}\right)\mathrm{d}v_e = \int_{v_e}\mathrm{d}p_i\delta v_i\mathrm{d}v_e + \int_{a_e}\mathrm{d}\bar{p}_i\delta v_i\mathrm{d}a_e \tag{3.3-334}$$

即弹塑性大变形虚功率方程为

$$\int_{v_e}\dot{t}_{ij}\delta\left(\frac{\partial v_j}{\partial x_i}\right)\mathrm{d}v_e = \int_{v_e}\dot{p}_i\delta v_i\mathrm{d}v_e + \int_{a_e}\dot{\bar{p}}_i\delta v_i\mathrm{d}a_e \tag{3.3-335}$$

6.4.2 单元插值关系

单元内任意一点的速度 v 与单元节点速度向量 $\boldsymbol{v}^e$ 的关系为

$$v = \boldsymbol{N}\boldsymbol{v}^{e} \tag{3.3-336}$$

式中，$\boldsymbol{N}$ 为单元形函数矩阵，由单元形函数确定，是 t 时刻坐标 x 的函数。

6.4.3 单元本构方程

式（3.3-299）给出弹塑性大变形本构方程的一般形式。$\overset{\nabla}{\sigma}_{ij}$ 与 $\dot{t}_{ij}$ 之间满足如下关系

$$\dot{t}_{ij} = \overset{\nabla}{\sigma}_{ij} - \sigma_{ik}d_{kj} - \sigma_{kj}d_{ki} + \sigma_{ik}v_{j,k} \tag{3.3-337}$$

具体推导过程如下。

设 t 时刻为参考构形，下一个邻近的 τ 时刻为现实构形，t 时刻位于 x_i 的质点在 τ 时刻位于 ξ_i，则

$$\sigma_{ij}(\tau) = J^{-1}\left(\frac{\partial \xi_i}{\partial x_k}\right)t_{kj}(\tau) \tag{3.3-338}$$

对于式（3.3-338）取物质导数，得

$$\dot{\sigma}_{ij}(\tau) = J^{-1}\left(\frac{\partial \xi_i}{\partial x_k}\right)\dot{t}_{kj}(\tau) + t_{kj}(\tau)\left[\dot{J}^{-1}\left(\frac{\partial \xi_i}{\partial x_k}\right) + J^{-1}\left(\frac{\partial \dot{\xi}_i}{\partial x_k}\right)\right] \tag{3.3-339}$$

在 Lagrange 描述中，物体质点的坐标 X_I 不随时间变化，在求某力学量对时间 t 的导数时，应固定 X_I 不变，这种导数称为物质时间导数，简称为物质导数。

当 $\tau \to t$ 时，$\xi_i \to x_i$、$\dot{\xi}_i \to v_i$、$t_{ij}(\tau) \to \sigma_{ij}$、$J \to 1$ 以及 $\dot{J} = Jv_{k,k}$。因此，

$$\dot{\sigma}_{ij}(t) = \delta_{ik}\dot{t}_{kj}(t) + \sigma_{kj}(-v_{k,k}\delta_{ik} + v_{i,k}) \tag{3.3-340}$$

可得

$$\begin{aligned}\dot{t}_{ij}(t) &= \dot{\sigma}_{ij} + \sigma_{ij}v_{k,k} - \sigma_{kj}v_{i,k} \\ &= \overset{\nabla}{\sigma}_{ij} - \sigma_{ik}d_{kj} - \sigma_{kj}d_{ki} + \sigma_{ik}v_{j,k} + \sigma_{ij}v_{k,k}\end{aligned} \tag{3.3-341}$$

利用

$$t_{ij}(\tau) = \left(\frac{\partial \xi_i}{\partial x_k}\right)S_{kj}(\tau) \tag{3.3-342}$$

可得

$$\dot{t}_{ij}(t) = \dot{S}_{ij}(t) + \sigma_{ik}v_{j,k} \tag{3.3-343}$$

式中，$\dot{S}_{ij}$ 为第二类 Kirchhoff 应力。考虑金属材料不可压缩性，$v_{k,k} \approx 0$，所以

$$\dot{t}_{ij}(t) = \overset{\nabla}{\sigma}_{ij} - \sigma_{ik}d_{kj} - \sigma_{kj}d_{ki} + \sigma_{ik}v_{j,k} \tag{3.3-344}$$

6.4.4 单元几何关系

变形率张量 d_{ij} 为

$$d_{ij} = \frac{1}{2}(v_{i,j} + v_{j,i}) \tag{3.3-345}$$

式中，$v_{i,j}$ 为速度梯度张量。以当前时刻 t 为参考构形时，变形率张量 d_{ij} 等于应变率张量 $\dot{\varepsilon}_{ij}$。因此单元的几何方程为

$$\dot{\varepsilon}_{ij} = \frac{1}{2}(v_{i,j} + v_{j,i}) \tag{3.3-346}$$

6.4.5 单元平衡方程

将式（3.3-344）代入虚功率方程式（3.3-335）得

$$\int_{v_e}(\overset{\nabla}{\sigma}_{ij} - \sigma_{ik}d_{kj} - \sigma_{kj}d_{ki} + \sigma_{ik}v_{j,k})\delta v_{j,i}\mathrm{d}v_e = \int_{v_e}\dot{p}_i(t)\delta v_i\mathrm{d}v_e + \int_{a_e}\dot{\bar{p}}_i(t)\delta v_i\mathrm{d}a_e \tag{3.3-347}$$

容易证明：任意的对称张量 B_{ij} 都有 $B_{ij}v_{i,j} = B_{ij}d_{ij}$。即

$$\begin{aligned}B_{ij}d_{ij} &= \frac{1}{2}B_{ij}(v_{i,j} + v_{j,i}) = \frac{1}{2}(B_{ij}v_{i,j} + B_{ij}v_{j,i}) \\ &= \frac{1}{2}(B_{ij}v_{i,j} + B_{ji}v_{i,j}) = B_{ij}v_{i,j}\end{aligned}$$

证毕。

利用上式，代入本构方程式（3.3-336），则式（3.3-347）为

$$\begin{aligned}&\int_{v_e}(D^{ep}_{ijkl}d_{kl} - \sigma_{ik}d_{kj} - \sigma_{kj}d_{ki} + \sigma_{ik}v_{j,k})\delta v_{j,i}\mathrm{d}v_e \\ &= \int_{v_e}[D^{ep}_{ijkl}d_{kl}\delta d_{ij} - (\sigma_{ik}d_{kj} + \sigma_{kj}d_{ki})\delta d_{ij} + \sigma_{ik}v_{j,k}\delta v_{j,i}]\mathrm{d}v_e \\ &= \int_{v_e}\dot{p}_i(t)\delta v_i\mathrm{d}v_e + \int_{a_e}\dot{\bar{p}}_i(t)\delta v_i\mathrm{d}a_e\end{aligned} \tag{3.3-348}$$

进一步推导，则

$$\int_{v_e}(D^{ep}_{ijkl} - F_{ijkl})d_{kl}\delta d_{ij}\mathrm{d}v_e + \int_{v_e}\sigma_{ik}v_{j,k}\delta v_{i,j}\mathrm{d}v_e = \int_{v_e}\dot{p}_i\delta v_i\mathrm{d}v_e + \int_{a_e}\dot{\bar{p}}_i\delta v_i\mathrm{d}a_e \tag{3.3-349}$$

式中

$$F_{ijkl} = \frac{1}{2}(\sigma_{lj}\delta_{ki} + \sigma_{kj}\delta_{li} + \sigma_{li}\delta_{kj} + \sigma_{ki}\delta_{lj}) \tag{3.3-350}$$

为了描述方便，下面改用矩阵推导：

$$\int_{v_e}[\delta\boldsymbol{d}^{\mathrm{T}}(\boldsymbol{D}^{ep} - \boldsymbol{F})\boldsymbol{d} + \delta\boldsymbol{q}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{q}]\mathrm{d}v_e = \int_{v_e}\delta\boldsymbol{v}^{\mathrm{T}}\dot{\boldsymbol{p}}\mathrm{d}v_e + \int_{a_e}\delta\boldsymbol{v}^{\mathrm{T}}\dot{\bar{\boldsymbol{p}}}\mathrm{d}a_e \tag{3.3-351}$$

式中，$\boldsymbol{D}^{ep}$ 为弹本构矩阵。

将单元几何关系式（3.3-345）和单元插值关系式（3.3-336）代入虚功率方程式（3.3-351），得

$$(\delta\boldsymbol{v}^e)^{\mathrm{T}}\left\{\left[\int_{v_e}\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{D}^{ep} - \boldsymbol{F})\boldsymbol{B}\mathrm{d}v_e\right]\boldsymbol{v}^e + \left(\int_{v_e}\boldsymbol{B}_v^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{B}_v\mathrm{d}v_e\right)\boldsymbol{v}^e - \int_{v_e}\boldsymbol{N}^{\mathrm{T}}\dot{\boldsymbol{p}}\mathrm{d}v_e - \int_{a_e}\boldsymbol{N}^{\mathrm{T}}\dot{\bar{\boldsymbol{p}}}\mathrm{d}a_e\right\} = 0 \tag{3.3-352}$$

由于 $(\delta\boldsymbol{v}^e)^{\mathrm{T}}$ 是任意可能存在的速度向量变分，所以单元平衡方程为

$$(k_0^e - k_F^e + k_\sigma^e)v^e = \dot{p}^e + \dot{\bar{p}}^e \quad 或 \quad k^e v^e = f^e \tag{3.3-353}$$

式中

$$k_0^e = \int_{v_e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}^{ep}\boldsymbol{B}\mathrm{d}v_e$$
$$k_F^e = \int_{v_e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{B}\mathrm{d}v_e \tag{3.3-354}$$
$$k_\sigma^e = \int_{v_e}\boldsymbol{B}_v^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{B}_v\mathrm{d}v_e$$

$$\dot{p}^e = \int_{v_e}\boldsymbol{N}^{\mathrm{T}}\dot{\boldsymbol{p}}\mathrm{d}v_e$$
$$\dot{\bar{p}}^e = \int_{a_e}\boldsymbol{N}^{\mathrm{T}}\dot{\bar{\boldsymbol{p}}}\mathrm{d}a_e \tag{3.3-355}$$

对于弹塑性小应变情况，$k_F^e = 0$，$k_\sigma^e = 0$。总体速率平衡方程由单元平衡方程式（3.3-353）组装得到

$$K\dot{U} = \dot{F} \tag{3.3-356}$$

假设时间增量步很小，在任意时间增量步 Δt 内总体速率平衡方程式（3.3-356）都是线性的，由此可得增量型总体平衡方程

$$K\Delta U = \Delta F \tag{3.3-357}$$

增量型总体平衡方程式（3.3-357）是一个通常的线性方程组，可直接求解，而不需要进行迭代求解。

6.4.6 应力计算与变换

在任意时间增量步 Δt 内，单元内任意点的第二类 Kirchhoff 应力增量为

$$\Delta S_{ij} = (D^{ep}_{ijkl} - F_{ijkl})\Delta\varepsilon_{kl} \tag{3.3-358}$$

$t+\Delta t$ 时刻的第二类 Kirchhoff 应力为

$$S_{ij}^{t+\Delta t} = \sigma_{ij}^t + \Delta S_{ij} \tag{3.3-359}$$

则单元内任意点 $t+\Delta t$ 时刻的 Cauchy 应力 $\sigma_{ij}^{t+\Delta t}$ 为

$$\sigma_{ij}^{t+\Delta t} = \frac{1}{|J|}\frac{\partial x_i^{t+\Delta t}}{\partial x_k^t}\frac{\partial x_j^{t+\Delta t}}{\partial x_l^t}S_{kl}^{t+\Delta t} \tag{3.3-360}$$

式中，$|J|$ 为变形梯度 F_{ij} 的行列式

$$|J| = \left|\frac{\partial x_i^{t+\Delta t}}{\partial x_j^t}\right| \tag{3.3-361}$$

6.5 板料成形数值模拟算例

静力隐式有限元法从理论上说比动力显式方法更适合冲

压成形过程模拟，而且模拟的精度高。但是目前它仍存在一些问题，求解规模和求解稳定性是主要问题。随着工程技术的需要以及计算机速度和容量的成倍提高，这些问题将逐渐消失，静力隐式有限元法会被广泛地应用。

图 3.3-20 所示是一个轿车外覆盖件的有限元模拟模具装配图，采用双动压机形式模拟。根据对称性取其中的半计算，图 3.3-21 所示是有限元模拟最终结果的厚度分布图。

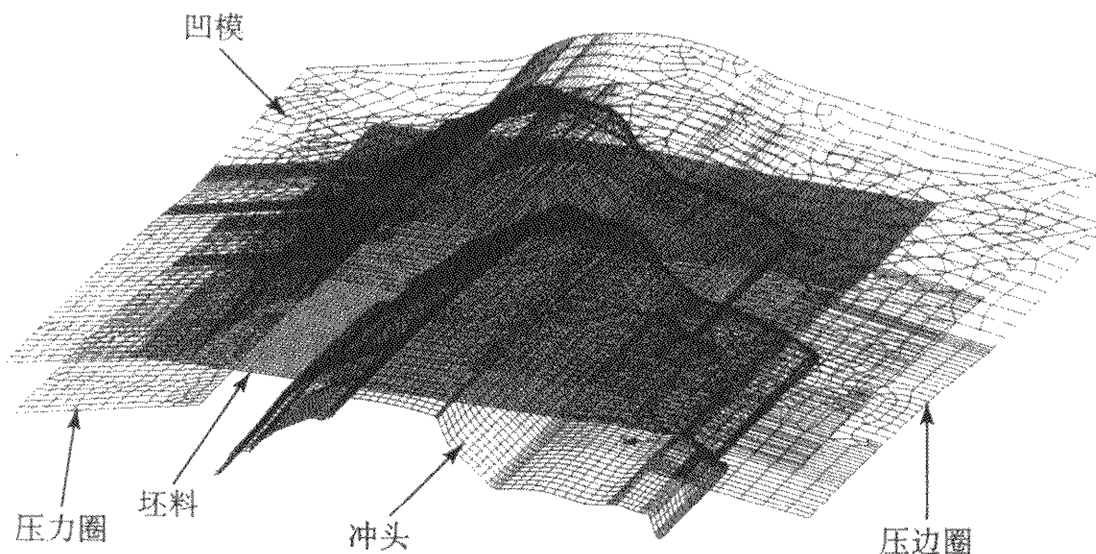


图 3.3-20　有限元模拟模具装配图

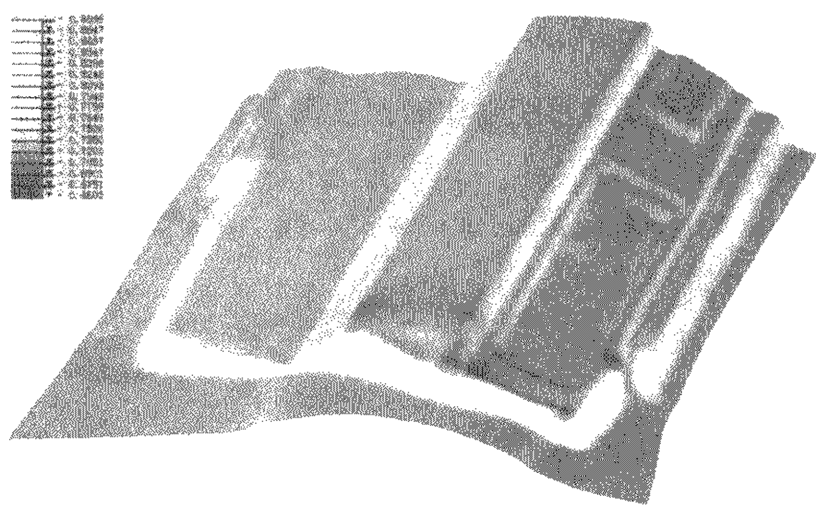


图 3.3-21　有限元模拟厚度分布结果

7　板料成形显式有限元方法

动力显式算法的有限元模型利用时间的中心差分，显式向前计算技术，回避了由于高度非线性引起的计算收敛性问题。

7.1　动力分析的虚功（率）方程

采用动力显式算法，板料的运动学微分方程为

$$\frac{\delta\sigma_{ij}}{\delta x_j}+p_i-\rho\ddot{u}_i-c\dot{u}_i=0 \tag{3.3-362}$$

式中，ρ 为材料的质量密度；c 为阻尼系数；$\dot{u}_i$ 和 $\ddot{u}_i$ 分别为材料内任一点的速度和加速度；p_i 为作用在该点上的外力；σ_{ij} 为该点处的 Cauchy 应力。

根据散度定理以及边界条件，由（3.3-362）式，可以得到系统的虚功方程为

$$\int_V\rho\ddot{u}_i\delta\dot{u}_i\,\mathrm{d}V+\int_V c\dot{u}_i\delta\dot{u}_i\,\mathrm{d}V=\int_V p_i\delta\dot{u}_i\,\mathrm{d}V+\int_\Gamma q_i\delta\dot{u}_i\,\mathrm{d}\Gamma-\int_V\sigma_{ij}\delta\dot{\varepsilon}_{ij}\,\mathrm{d}V \tag{3.3-363}$$

式中，$\delta\dot{u}_i$ 为虚速度；$\delta\dot{\varepsilon}_{ij}$ 为对应于 Cauchy 应力 σ_{ij} 的虚应变速率。

7.2　动力显式积分算法有限元方程

把物体离散成 m 个单元，对于任一单元，有 α 个节点，取其形函数为 N^α，单元内任意点的位移分量 u_i、速度分量 $\dot{u}_i$ 和加速度分量 $\ddot{u}_i$ 的插值关系分别为

$$\begin{cases}u_i=N^\alpha u_i^\alpha\\ \dot{u}_i=N^\alpha\dot{u}_i^\alpha\\ \ddot{u}_i=N^\alpha\ddot{u}_i^\alpha\end{cases} \tag{3.3-364}$$

单元的几何方程为

$$\dot{\varepsilon}_{ij}=B_j^\alpha\dot{u}_i^\alpha \tag{3.3-365}$$

式中，u_i^α、$\dot{u}_i^\alpha$ 和 $\ddot{u}_i^\alpha$ 分别为节点 α 的位移分量、速度分量和加速度分量；B_j^α 为应变几何张量。

将式（3.3-364）和式（3.3-365）代入式（3.3-363）可得：

$$\int_{V_e}\rho N^\alpha\ddot{u}_i^\alpha N^\beta\delta\dot{u}_i^\beta\,\mathrm{d}V+\int_{V_e}cN^\alpha\dot{u}_i^\alpha N^\beta\delta\dot{u}_i^\beta\,\mathrm{d}V=\int_{V_e}p_iN^\beta\delta\dot{u}_i^\beta\,\mathrm{d}V+\int_{\Gamma_e}q_iN^\beta\delta\dot{u}_i^\beta\,\mathrm{d}\Gamma-\int_{V_e}\sigma_{ij}B_j^\beta\delta\dot{u}_i^\beta\,\mathrm{d}V \tag{3.3-366}$$

式中，$\delta\dot{u}_i^\beta$ 为节点 β 的虚速度。式（3.3-366）可以写成

$$\int_{V_e}\rho N^\alpha N^\beta\,\mathrm{d}V\ddot{u}_i^\alpha+\int_{V_e}cN^\alpha N^\beta\,\mathrm{d}V\dot{u}_i^\alpha=\int_{V_e}p_iN^\beta\,\mathrm{d}V+\int_{\Gamma_e}q_iN^\beta\,\mathrm{d}\Gamma-\int_{V_e}\sigma_{ij}B_j^\beta\,\mathrm{d}V \tag{3.3-367}$$

写成矩阵形式为

$$\int_{V_e}\rho\boldsymbol{N}^{\mathrm{T}}\boldsymbol{N}\,\mathrm{d}V\ddot{u}+\int_{V_e}c\boldsymbol{N}^{\mathrm{T}}\boldsymbol{N}\,\mathrm{d}V\dot{u}=\int_{V_e}\boldsymbol{N}^{\mathrm{T}}\boldsymbol{p}\,\mathrm{d}V+\int_{\Gamma_e}\boldsymbol{N}^{\mathrm{T}}\boldsymbol{q}\,\mathrm{d}\Gamma-\int_{V_e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\sigma}\,\mathrm{d}V \tag{3.3-368}$$

将单元方程集合，即得整体有限元方程

$$\sum\left(\int_{V_e}\rho\boldsymbol{N}^{\mathrm{T}}\boldsymbol{N}\mathrm{d}V\right)\ddot{\boldsymbol{u}}+\sum\left(\int_{V_e}c\boldsymbol{N}^{\mathrm{T}}\boldsymbol{N}\mathrm{d}V\right)\dot{\boldsymbol{u}}=\sum\int_{V_e}\boldsymbol{N}^{\mathrm{T}}\boldsymbol{p}\,\mathrm{d}V+\sum\int_{\Gamma_e}\boldsymbol{N}^{\mathrm{T}}\boldsymbol{q}\,\mathrm{d}\Gamma-\sum\int_{V_e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\sigma}\,\mathrm{d}V \tag{3.3-369}$$

式（3.3-369）可简写为

$$\boldsymbol{M}\ddot{\boldsymbol{u}}+\boldsymbol{C}\dot{\boldsymbol{u}}=\boldsymbol{P}-\boldsymbol{F} \tag{3.3-370}$$

式中，$\boldsymbol{M}$ 为质量矩阵

$$\boldsymbol{M}=\sum\int_{V_e}\rho\boldsymbol{N}^{\mathrm{T}}\boldsymbol{N}\,\mathrm{d}V \tag{3.3-371}$$

$\boldsymbol{C}$ 为阻尼矩阵

$$\boldsymbol{C}=\sum\int_{V_e}c\boldsymbol{N}^{\mathrm{T}}\boldsymbol{N}\,\mathrm{d}V \tag{3.3-372}$$

$\boldsymbol{P}$ 是节点外力向量

$$\boldsymbol{P}=\sum\int_{V_e}\boldsymbol{N}^{\mathrm{T}}\boldsymbol{p}\,\mathrm{d}V+\sum\int_{\Gamma_e}\boldsymbol{N}^{\mathrm{T}}\boldsymbol{q}\,\mathrm{d}\Gamma \tag{3.3-373}$$

$\boldsymbol{F}$ 是节点内力向量

$$\boldsymbol{F}=\sum\int_{V_e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\sigma}\,\mathrm{d}V \tag{3.3-374}$$

通常，动力显式积分算法采用集中质量矩阵，即 $\boldsymbol{M}$ 是一对角矩阵，并取 $\boldsymbol{C}=\alpha\boldsymbol{M}$。则式（3.3-370）表示的联立方程组变成（节点数×节点自由度数）个相互独立的方程

$$m_i\ddot{u}_i + c_i\dot{u}_i = P_i - F_i \tag{3.3-375}$$

或

$$m_i\ddot{u}_i + \alpha m_i\dot{u}_i = P_i - F_i \tag{3.3-376}$$

7.3　显式时间积分的中心差分算法

设 t 时刻的状态为 n，t 时刻及 t 时刻之前的力学量已知，且定义 $t-\Delta t$ 为 $n-1$ 状态，$t-\frac{1}{2}\Delta t$ 为 $n-\frac{1}{2}$ 状态，$t+\Delta t$ 为 $n+1$ 状态，$t+\frac{1}{2}\Delta t$ 为 $n+\frac{1}{2}$ 状态。设 t 时刻前后两时间增量步长不同，即 $\Delta t_n \neq \Delta t_{n-1}$，令 $\beta=\frac{\Delta t_n}{\Delta t_{n-1}}$。将节点速度和加速度用差分格式写成

$$\dot{u}_n = \frac{\beta}{1+\beta}\dot{u}_{n+1/2} + \frac{1}{1+\beta}\dot{u}_{n-1/2} \tag{3.3-377}$$

$$\ddot{u}_n = \frac{2}{(1+\beta)\ \Delta t_{n-1}}\ (\dot{u}_{n+1/2} - \dot{u}_{n-1/2}) \tag{3.3-378}$$

而 $t+\Delta t$ 时刻（$n+1$ 状态）的总位移可由下式累加得出

$$u_{n+1} = u_n + \dot{u}_{n+1/2}\Delta t_n \tag{3.3-379}$$

将式（3.3-377）和式（3.3-378）代入式（3.3-376），可得

$$\frac{2m_i}{(1+\beta)\Delta t_{n-1}}(\dot{u}_{n+1/2} - \dot{u}_{n-1/2}) + \alpha m_i\left(\frac{\beta}{1+\beta}\dot{u}_{n+1/2} + \frac{1}{1+\beta}\dot{u}_{n-1/2}\right) = P_n - F_n \tag{3.3-380}$$

整理得

$$\left[\frac{2m_i}{(1+\beta)\Delta t_{n-1}} + \frac{\alpha\beta m_i}{1+\beta}\right]\dot{u}_{n+1/2} + \left[\frac{\alpha m_i}{1+\beta} - \frac{2m_i}{(1+\beta)\Delta t_{n-1}}\right]\dot{u}_{n-1/2} = G_n \tag{3.3-381}$$

$$A_i\dot{u}_{n+1/2} = B_i\dot{u}_{n-1/2} + G_n \tag{3.3-382}$$

$$\dot{u}_{n+1/2} = \frac{B_i}{A_i}\dot{u}_{n-1/2} + \frac{1}{A_i}G_n \tag{3.3-383}$$

式中，$A_i=\frac{2m_i+\alpha\beta m_i\Delta t_{n-1}}{(1+\beta)\Delta t_{n-1}}$，$B_i=\frac{2m_i-\alpha m_i\Delta t_{n-1}}{(1+\beta)\Delta t_{n-1}}$，最后写成

$$\dot{u}_{n+1/2} = \frac{2-\alpha\Delta t_{n-1}}{2+\alpha\beta\Delta t_{n-1}}\dot{u}_{n-1/2} + \frac{(1+\beta)\ \Delta t_{n-1}}{2+\alpha\beta\Delta t_{n-1}}\ (P_n - F_n) \tag{3.3-384}$$

式（3.3-379）和式（3.3-384）给出了节点位移和速度的显式计算格式，前提条件是已知前两步的位移和速度。在第一步计算时，因为 $t-\frac{1}{2}\Delta t$ 时刻的速度 $\dot{u}_{n-1/2}$ 未知，不能直接用式（3.3-379）和式（3.3-384）进行计算。但是，通常板料成形前的位移和速度的初始条件是已知的，即

$$u_0 = 0,\ \dot{u}_0 = 0 \tag{3.3-385}$$

式中，0 为零向量。令 $\Delta t_0=\Delta t_{0-1}$，即 $\beta=1$，由上述的速度初始条件以及式（3.3-377），即可得到 $0-\frac{1}{2}$ 时刻的速度向量

$$\dot{u}_{0-1/2} = -\dot{u}_{0+1/2} \tag{3.3-386}$$

将上式代入式（3.3-383），即可得到第一个增量步中节点速度的计算表达式

$$\dot{u}_{0+1/2} = \frac{1}{2}\frac{\Delta t_0}{m_i}(P_0 - F_0) \tag{3.3-387}$$

综上所述，采用中心差分显式计算的步骤为：

1）确定板料的初始条件，给出初始时刻的节点位移和速度，求出 $0-\frac{1}{2}$ 时刻的节点速度；

2）当 $t=0$ 时，由式（3.3-387），计算 $0+\frac{1}{2}$ 时刻的节点速度；

3）当 $t\neq0$ 时，由式（3.3-384），计算 $t+\frac{1}{2}\Delta t$ 时刻的节点速度；

4）由式（3.3-379），计算 $t+\Delta t$ 时刻的节点位移；

5）反复 3）、4）步骤，直到计算终止。

7.4　临界时间步长的确定

由于中心差分算法是条件稳定的，为了保证系统计算的稳定性，对时间增量步长 Δt 的大小必须加以限制。稳定性条件通常由系统的最高频率 ω_{max} 决定，满足稳定性条件的时间增量步长为

$$\Delta t \leqslant \frac{2}{\omega_{max}}\ (\sqrt{1+\xi^2} - \xi) \tag{3.3-388}$$

其中 ξ 是最高模态中的临界阻尼。与工程直觉相反，阻尼的引入实际上降低了系统的临界稳定性条件。系统的最高频率由网格中最大的单元膨胀模式决定。

满足稳定性条件的时间增量步长可以由膨胀波沿网格中任意单元的最小穿越时间近似得到

$$\Delta t_n \leqslant \gamma\frac{L_n^e}{c} \tag{3.3-389}$$

其中，$\gamma=0.5\sim0.8$；c 为膨胀波在材料中的传播速度；L_n^e 为第 n 状态单元 e 的名义长度。

稳定性条件可以保证在一个时间增量步内，扰动只传播网格中的一个单元。如果系统只包括一种材料，则满足稳定性条件的时间增量步长与网格中最小的单元尺寸成正比；如果系统划分的单元网格尺寸比较均匀，但包括多种不同材料，则具有最高膨胀波速的材料中网格尺寸最小的单元决定系统的稳定时间步长。

对于一个简单的桁架单元，在团聚质量矩阵的情况下，稳定性准则给出一个临界时间步长：$\Delta t\leqslant\frac{l}{c}$，式中，$c$ 为材料声速，l 为单元长度，Δt 表示膨胀波穿越长度为 l 的单元所需要的时间。这就是所谓的 Courant - Friedrichs - Lewy（CFL）稳定性条件。

对于三角形单元和四边形板单元来说，临界时间步长的选取依赖于单元名义长度的确定，一般按照图 3.3-22 的原则来确定单元的名义长度 l_{crit}。

对于高阶单元来说，临界时间步长远比低阶单元的临界时间步长小。这一事实使得高阶单元对于显式积分算法相当不合适。

虽然上面给出的稳定性准则严格来说是对线性系统而言的，但对于非线性问题也给出了有用的稳定性估计。对线性问题时间步长的 80%～90% 缩小，对于大多数非线性问题保持其系统稳定性是足够的。然而，十分重要的是：在整个计算过程中，要不断地检查能量的平衡问题。任何总能量的增加或损失（5%或更多）都将导致失稳。Belytschko 指出：常增量时间步不能保持解的稳定性，即使系统的最高频率 ω_{max} 不断减小。

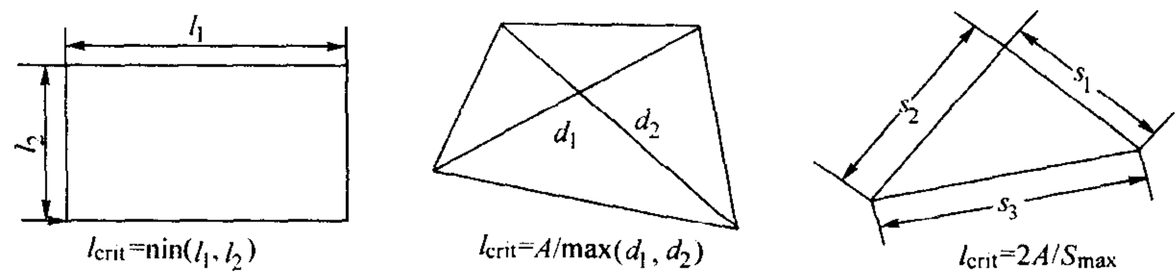


图 3.3-22　单元名义长度

8　板料成形有限元逆算法

坯料形状反算与优化是模具工艺设计过程中一个重要的问题，不仅可以节省材料，也可以改善板料成形过程中的塑性流动规律，减小或避免修边工艺，降低成本，提高产品质量。传统方法主要有经验法、滑移线法、几何映射法、电模拟法等。这些方法在理论方面都或多或少地存在一定的缺陷，应用范围受到限制，计算精度也不高，尤其对于复杂冲压件来说它们都很难预测坯料形状。

有限元逆算法已应用于板料冲压成形坯料形状和应变分布的预测。这种方法也叫一步成形有限元法，根据产品零件或已经工艺补充的冲压件几何形状来预测它的坯料形状和可成形性。由于这种算法模拟速度非常快，数据准备量少，因此它可以在产品设计阶段和模具工艺补充设计就进行快速成形性分析，优化工艺参数和工艺设计方案。

8.1　有限元逆算法基本理论

板料在冲压成形过程中假设是比例加载的变形过程，并且材料不可压缩，模拟过程中采用塑性全量形变本构模型。有限元逆算法的基本思想是在成形后的冲压件上建立有限元方程进行迭代求解，坯料与冲压件之间几何尺寸和物理量情况如表 3.3-1 所示。

表 3.3-1　坯料与冲压件的几何尺寸和物理量比较

项　目	坯料 C_0	冲压件 C
几何尺寸	未知	已知
板厚	已知	未知
应力、应变	已知	未知
工艺条件、边界条件	已知	已知

从表 3.3-1 中可以发现，推导有限元逆算法所需要的基本条件和物理在坯料或冲压件中是已知的，其中 3 个未知的量是有限元逆算法要求解的。

8.1.1　运动方程与几何关系

假设板料变形满足 Kirchhoff 薄板理论。初使坯料状态为 C_0，最终冲压件状态为 C，则在两构形间板内任意点 p 的运动方程为

$$\boldsymbol{x}_0 = \boldsymbol{x} - [\boldsymbol{u} + z(\boldsymbol{n} - \frac{h}{h_0}\boldsymbol{n}_0)] \qquad (3.3\text{-}390)$$

式中，$\boldsymbol{x}_0$ 和 $\boldsymbol{x}$ 为点 p 在 C_0 和 C 状态的坐标向量；$\boldsymbol{n}_0$ 和 $\boldsymbol{n}$ 为点 p 在 C_0 和 C 状态时的单位法向量；h_0 和 h 为点 p 在 C_0 和 C 状态时的板厚；$\boldsymbol{u}$ 为点 p 所在板中性面处的位移向量；z 为板法向坐标值。

x_0 和 x 满足如下变换关系

$$\mathrm{d}\boldsymbol{x}_0 = \boldsymbol{F}\mathrm{d}\boldsymbol{x} \qquad (3.3\text{-}391)$$

式中，$\boldsymbol{F}$ 为变形梯度张量。

因此左 Cauchy - Green 张量 $\boldsymbol{B}$ 为

$$\boldsymbol{B}^{-1} = \boldsymbol{F}^{-\mathrm{T}}\boldsymbol{F}^{-1} \qquad (3.3\text{-}392)$$

由式（3.3-390）～式（3.3-392）可得 p 点所在中性面处的对数应变 ε_{xx}，ε_{yy}，ε_{xy} 分别为

$$\begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{xy} \end{Bmatrix} = \begin{bmatrix} \ln\lambda_1\cos^2\theta + \ln\lambda_2\sin^2\theta \\ \ln\lambda_1\sin^2\theta + \ln\lambda_2\cos^2\theta \\ (\ln\lambda_1 - \ln\lambda_2)\ \cos\theta\sin\theta \end{bmatrix} \qquad (3.3\text{-}393)$$

式中，λ_1^{-2} 和 λ_2^{-2} 为 $\boldsymbol{B}^{-1}$ 张量的特征值；θ 为第一主应变方向。

8.1.2　屈服准则与本构方程

在平面应力状态下，把各向异性主轴作为 x，y 轴的话，Hill 各向异性屈服准则为

$$f = \frac{1}{2(F+G+H)}[(G+H)\sigma_x^2 + (F+H)\sigma_y^2 - 2H\sigma_x\sigma_y + 2N\sigma_{xy}^2] - \frac{1}{3}\bar{\sigma}^2 \qquad (3.3\text{-}394)$$

将式（3.3-394）代入 Hencky 全量形变理论可得本构方程

$$\sigma = E_s\boldsymbol{P}^{-1}\varepsilon = D_s\varepsilon \qquad (3.3\text{-}395)$$

式中

$$E_s = \frac{\bar{\sigma}}{\bar{\varepsilon}} \qquad (3.3\text{-}396)$$

$$\boldsymbol{P} = \frac{3R_{90}\ (1+R_0)}{2\ (R_0+R_{90}+R_0R_{90})}\begin{bmatrix} 1 & -\frac{R_0}{1+R_0} & 0 \\ -\frac{R_0}{1+R_0} & \frac{R_0\ (1+R_{90})}{R_{90}\ (1+R_0)} & 0 \\ 0 & 0 & \frac{(1+2R_{45})\ (R_0+R_{90})}{R_{90}\ (1+R_0)} \end{bmatrix} \qquad (3.3\text{-}397)$$

式中，σ 为 Cauchy 应力张量；$\bar{\sigma}$ 和 $\bar{\varepsilon}$ 分别为等效应力和等效应变；R_0、R_{45}、R_{90} 为板料的各向异性参数。

8.1.3　虚功方程

在 C 状态上建立虚功方程

$$W = W_{\mathrm{int}} - W_{\mathrm{ext}} = \int_V \varepsilon^{\mathrm{T}}\boldsymbol{\sigma}\mathrm{d}V - \int_V \boldsymbol{u}^{\mathrm{T}}\boldsymbol{f}\mathrm{d}V = 0 \qquad (3.3\text{-}398)$$

式中，$\boldsymbol{f}$ 为外力向量；$\boldsymbol{u}^{\mathrm{T}}$ 为虚位移向量。

冲压件（C 状态）经过单元离散化

$$W = \sum_e (\boldsymbol{u}^e)^{\mathrm{T}}\ (F_{\mathrm{int}}^e - F_{\mathrm{ext}}^e) = -\sum_e (\boldsymbol{u}^e)^{\mathrm{T}}\boldsymbol{R}^e = 0 \qquad (3.3\text{-}399)$$

或者

$$W = \boldsymbol{U}^{\mathrm{T}}\ (F_{\mathrm{int}} - F_{\mathrm{ext}}) = -\boldsymbol{U}^{\mathrm{T}}\boldsymbol{R} = 0 \qquad (3.3\text{-}400)$$

采用 Newton - Raphson 方法求解式（3.3-400）非线性方程组，对于第 i 个迭代步

$$R(U^i) = F_{\mathrm{ext}}(U^i) - F_{\mathrm{int}}(U^i) \neq 0 \qquad (3.3\text{-}401)$$

$$K_{\mathrm{T}}^i\Delta U = R\ (U^i) \qquad (3.3\text{-}402)$$

$$U^{i+1} = U^i + \Delta U \quad (3.3\text{-}403)$$

$$K_T^i = \left[-\frac{\partial R(U)}{\partial U} \right]_{U=U^i} \quad (3.3\text{-}404)$$

单元内力部分切线刚度矩阵为

$$\frac{\partial \boldsymbol{F}_{\text{int}}^{\text{e}}}{\partial \boldsymbol{U}^{\text{e}}} = \boldsymbol{T}^{\text{T}} h \frac{\partial\left(\int_A \boldsymbol{B}^{\text{T}} \sigma \text{d}A\right)}{\partial u^{\text{e}}} = \boldsymbol{T}^{\text{T}}\left(h\int_A \boldsymbol{B}^{\text{T}} \boldsymbol{D}_{\text{s}} \boldsymbol{B} \text{d}A\right) \boldsymbol{T}$$

$$= \boldsymbol{T}^{\text{T}} \boldsymbol{K}_{\text{T}}^{\text{int}} \boldsymbol{T} \quad (3.3\text{-}405)$$

式中，$\boldsymbol{T}^{\text{T}}$ 为单元局部坐标系与整体坐标系之间的转换矩阵；$\boldsymbol{B}$ 为单元应变矩阵；A 为 C 状态的单元面积；h 为单元厚度。

8.2　初始场猜测

迭代方程式（3.3-401）需要给一个初始解才能进行迭代计算，这个初始解是根据冲压件结构和可能的变形情况人为猜测的，它直接影响方程组迭代收敛速度和收敛性，是有限元逆算法的关键技术之一。

8.2.1　Z 向平面投影法

Z 向平面投影法是最简单的一种位移初始场猜测方法。如图 3.3-23 所示，冲压件的有限元节点 1、2、3、4…沿整体坐标 Z 向投影到平面上 1′、2′、3′、4′…点。这种方法非常简单，很容易实现，但是当冲压件存在垂直面或接近垂直时，这种方法就失效了。

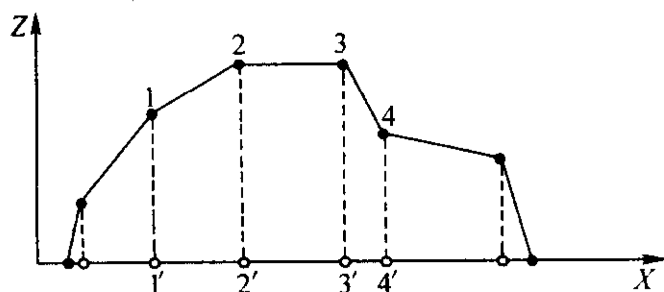


图 3.3-23　Z 向平面投影法

8.2.2　Z 向曲面投影法

为了避免存在垂直面的冲压件在进行 Z 向平面投影时失效，采用 Z 向曲面投影法，即将平面的投影面改成曲面，这个曲面必须是不存在垂直面的。如图 3.3-24 所示，冲压件的有限元节点 1、2、3、4…首先按节点最小距离法投影到 P 曲面上①、②、③、④点，然后再将①、②、③、④点按 Z 向平面投影法投影到 XY 平面上 1′、2′、3′、4′…点。

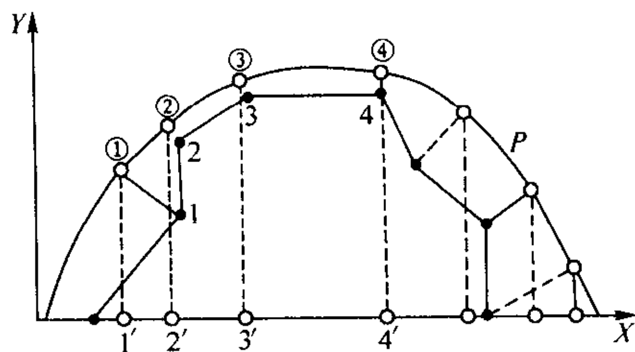


图 3.3-24　Z 向曲面投影法

实际应用时可以取一个半球面作为投影 P 曲面，这样计算相对比较简单。但是这种方法猜测的位移初始场有时与真实的坯料形状差别很大，会造成方程组式（3.3-401）迭代收敛速度很慢甚至迭代发散。

8.2.3　弹性变形投影法

假设冲压件在投影过程中是线弹性变形。如图 3.3-25 所示，投影过程要进行两次变换，具体作法如下。

1）从最终状态到初始状态，对每个单元进行两种变换，如图 3.3-25 所示绕 P 旋转和 Z 向投影变换。计算两个变换间的单元应变 ε_0。

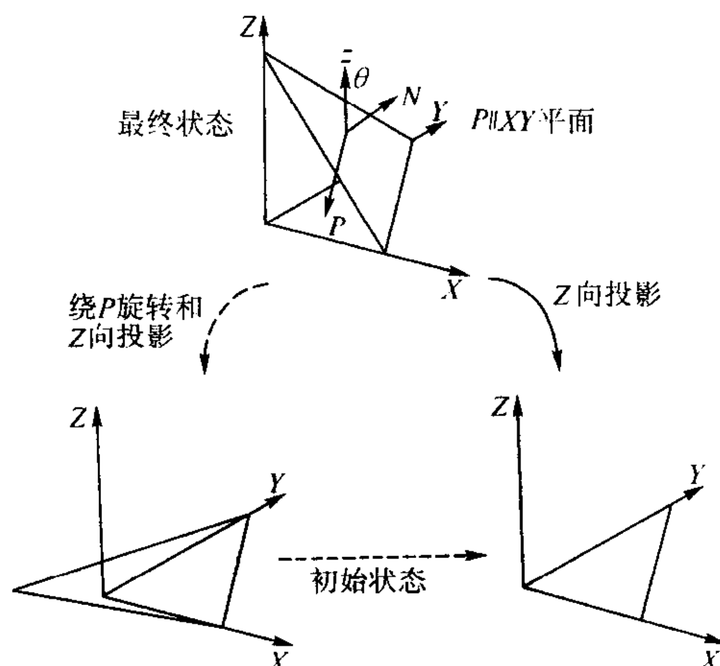


图 3.3-25　弹性变形投影法

2）按式（3.3-406）的冲压件有限元平衡方程求解单元两次变换的弹性变形。

$$\sum_{i=1}^{N_e} \int_{V_e} \boldsymbol{B}^{\text{T}} \boldsymbol{D}^{\text{e}} \boldsymbol{B} \boldsymbol{U} \text{d}V_e = \sum_{i=1}^{N_e} \int_{V_e} \boldsymbol{B}^{\text{T}} \boldsymbol{D}^{\text{e}} \varepsilon_0 \text{d}V_e \quad (3.3\text{-}406)$$

3）如果式（3.3-406）收敛，转到 4），否则转到 2）。

4）根据初始状态和最终状态的面积比，按比例适当调整由式（3.3-406）得到的初始状态。

这几种位移初始场猜测方法，Z 向平面投影法和 Z 向曲面投影法比较简单，计算量小，但是它们在应用时都有局限性。弹性变形投影法是一种通用方法，由于要反复迭代求解初始状态，计算量相对于其他两种方法比较大。实际上，式（3.3-406）有限元平衡方程依然是每个节点两个自由度，与逆算法迭代方程式（3.3-401）相同，而且式（3.3-406）的计算时间相对于式（3.3-401）来说还是比较小的。弹性变形投影法还有一个优势是可以考虑一些实际工艺条件，所猜测的初始场与实际坯料形状比较接近。当冲压件变形很小或是弯曲为主的成形时，它所猜测的初始场与实际坯料形状是一致的，没有必要再采用逆算法进行迭代计算。因此弹性变形投影法被广泛应用，AUTOFORM 中的 One Step 以及 FASTAMP 都采用了这种方法。

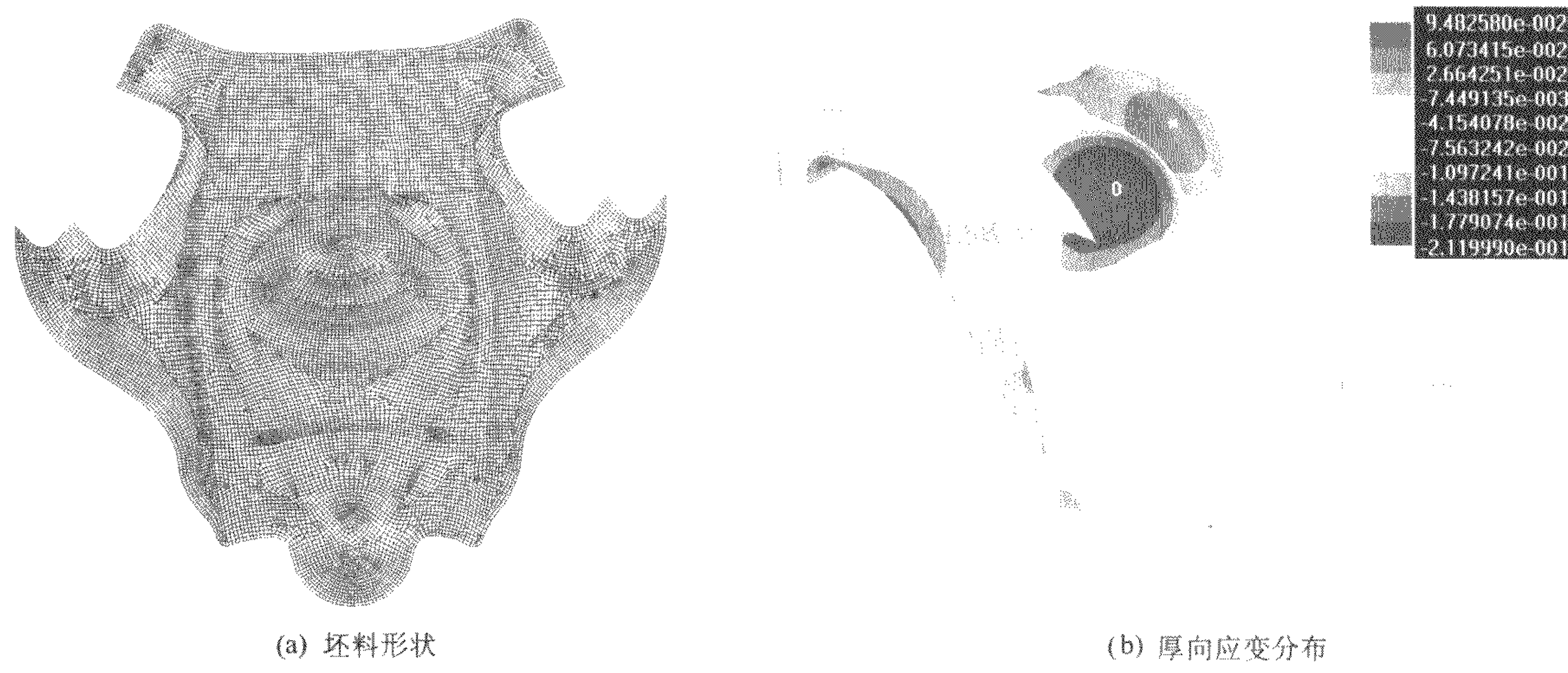
8.3　逆算法算例

8.3.1　级进模

级进模对零件的坯料形状的尺寸精度要求比较高，传统方法一般采用试错法，反复调试确定毛坯形状，这种方法工作效率低，需要比较长的调试时间。逆算法在这方面具有很大的优势，它可在很短的时间内，比较精确地计算出毛坯尺寸，如图 3.3-26a 所示。使模具调试次数成倍减少，缩短调试时间，降低成本。而且逆算法还能计算出冲压件的厚度减薄情况，如图 3.3-26b 所示，为模具工艺设计提供定量化参考。

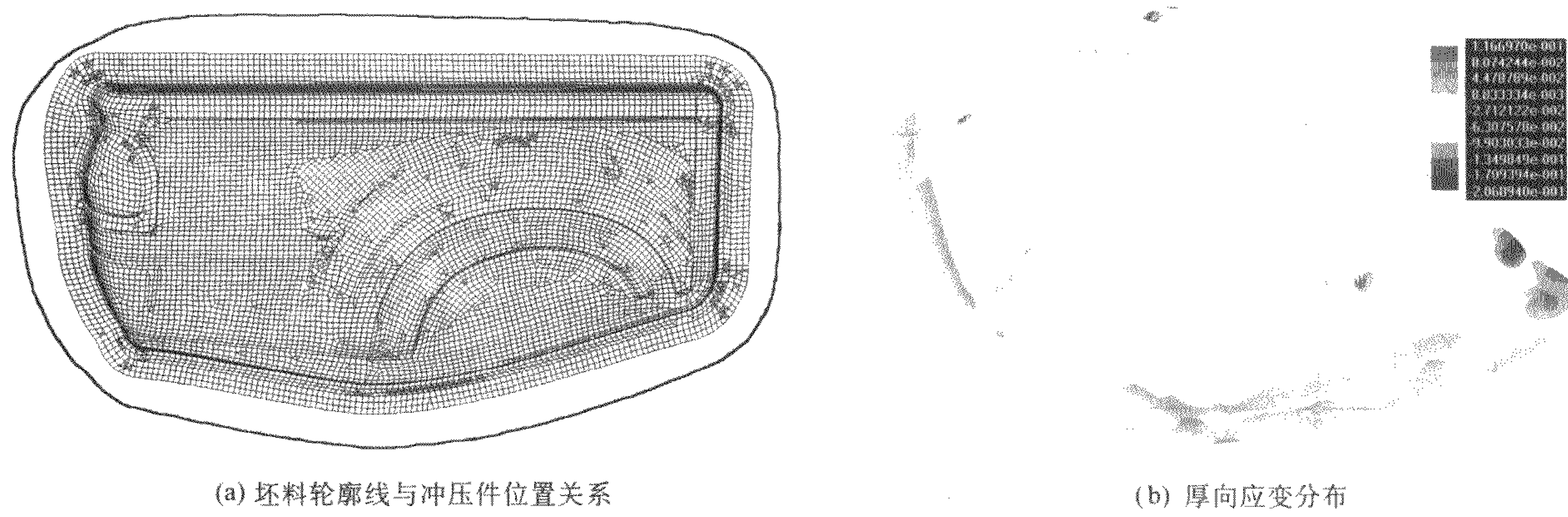
8.3.2　大型覆盖件

图 3.3-27 所示为轿车后翼子板，是比较大型的外覆盖件。采用逆算法可以比较好地预测冲压件的最佳坯料形状以及坯料与冲压件的位置关系，如图 3.3-27a 所示。根据厚向应变分布情况，在工艺设计阶段就可以发现其中存在的问题，改进工艺设计方案，如图 3.3-27b 所示。这一点相对于传统增量有限元法，逆算法有很大的优势。而且逆算法计算速度快，计算准备数据量少，可以反复计算优化工艺方案。



(a) 坯料形状　　(b) 厚向应变分布

图 3.3-26　级进模



(a) 坯料轮廓线与冲压件位置关系　　(b) 厚向应变分布

图 3.3-27　轿车后翼子板

编写：柳玉起（华中科技大学）

第 4 章　体积成形数值模拟方法

1　刚塑性有限元法

1.1　概述

金属塑性加工根据材料的变形特点，可以分为体积成形和板料成形。体积成形工艺包括锻造与轧制等，材料在成形过程中产生较大的塑性变形，相比之下弹性变形很小，因此可以忽略不计。20 世纪 70 年代初，德国的 Lung 和美国的 Lee 及 Kobayashi 分别提出了刚塑性有限元法。由于不需考虑材料弹性变形，所以可采用比弹塑性有限元法大的增量步长，从而减少计算时间。实际上，对于绝大多数体积成形过程，金属材料的塑性变形量很大；由于温度的影响，材料的弹性参数难以确定；同时回弹的影响不大，采用刚塑性小变形增量分析法是非常有效的，求解精度也完全满足工程应用。因此，刚塑性有限元法已成为分析材料体积成形的重要数值方法之一。但由于忽略了成形过程的弹性变形，采用刚塑性有限元法不能确定刚性区的应力、应变分布，不能处理卸载问题，因此也不能进行回弹和残余应力的分析。

从刚塑性有限元理论提出后，各国学者就对其进行了广泛深入的研究。Kobayashi 及其合作者先后用刚塑性有限元和刚黏塑性有限元分析了锻造、镦粗、挤压、轧制等体积成形以及板料的拉延、胀形、弯曲、缩口等成形问题。Zienkiewicz 于 1974 年提出了刚黏塑性有限元的罚函数法。1982 年 Mori 和 Osakada 提出了刚塑性有限元中的材料可压缩法，并对各类轧制和挤压工艺进行了模拟分析。近年来，一些学者对三维金属塑性成形问题作了刚塑性有限元分析。随着有限元理论的发展，研究的重点转向了模拟过程中遇到的各种问题，其中包括网格自动生成算法研究，成形过程模拟中网格再划分技术的研究，摩擦问题的研究，任意模具边界的处理等。这些研究都丰富和发展了塑性有限元的具体处理技术和应用范围。

随着有限元理论和相关技术的发展，刚塑性有限元法的应用经历了从二维到三维问题的分析，由简单工艺的模拟发展到对复杂成形工艺的优化预测，同时还能够进行成形过程的热力耦合分析。现在，更是应用到反向变形的模拟分析，即从成品的形状尺寸，通过计算反向模拟出合理的毛坯形状和尺寸。在人机交互方面，也集成了友好的前后处理模块。

1.1.1　离散化

在有限元计算时，最基本的就是要将材料进行离散化。所谓离散化就是将连续的材料划分成有限个单元组成的离散体，并把作用力按等效原则移置到各个节点上。而选择适当的单元和形函数是离散化中的关键问题。一般来说，单元形状的选择依赖于结构或总体求解域的几何特点以及求解所希望的精度等因素，而有限元插值函数则取决于单元的形状、节点数目和类型等因素。

1.1.2　单元与形函数

(1) 平面 4 节点四边形单元

对于二维问题，通常选用 4 节点四边形单元。在四边形单元上建立一个局部坐标，在局部坐标中四边形可以变换成一个正方形，形函数可直接在局部坐标中建立，4 节点四边形单元如图 3.4-1 所示。

单元形函数以参数形式定义，定义域为 $-1\leq\xi\leq1$，$-1\leq\eta\leq1$，形函数为

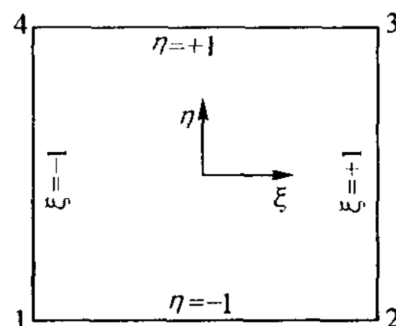


图 3.4-1　4 节点四边形单元

$$N_i(\xi,\eta)=\frac{1}{4}(1+\xi_i\xi)(1+\eta_i\eta)(i=1,2,3,4) \quad (3.4\text{-}1)$$

式中，ξ_i，η_i 为节点 i 的局部坐标值。

单元内的容许速度场可以由节点速度矢量表示为

$$\boldsymbol{v}=\boldsymbol{N}\dot{\boldsymbol{u}}^e \quad (3.4\text{-}2)$$

式中，$\boldsymbol{v}$ 为单元任一点速度矢量；$\dot{\boldsymbol{u}}^e$ 为单元的节点速度矢量；$\boldsymbol{N}$ 为单元形函数矩阵。

单元内局部坐标系下任一点 $(\xi,\ \eta)$ 可以映射到整个坐标 $(x,\ y)$ 内，定义如下

$$\left.\begin{aligned}x(\xi,\eta)&=\sum_i N_i(\xi,\eta)x_i\\ y(\xi,\eta)&=\sum_i N_i(\xi,\eta)y_i\end{aligned}\right\} \quad (3.4\text{-}3)$$

式中，x_i，y_i 为节点 i 的整体坐标。

(2) 三维 8 节点六面体单元

对于三维问题，通常选用 8 节点六面体单元，如图 3.4-2 所示。

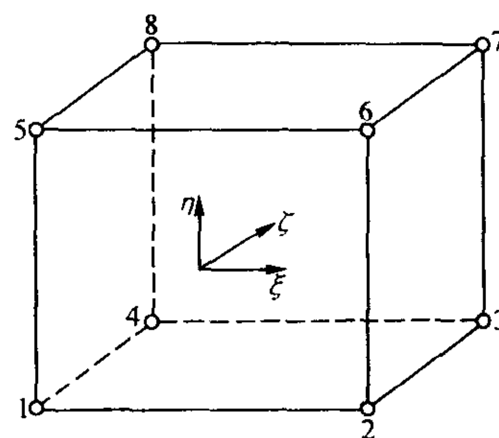


图 3.4-2　8 节点六面体单元

其形函数定义域为 $-1\leq\xi\leq1$、$-1\leq\eta\leq1$ 和 $-1\leq\zeta\leq1$，形函数如下

$$N_i(\xi,\eta,\zeta)=\frac{1}{8}(1+\xi_i\xi)(1+\eta_i\eta)(1+\zeta_i\zeta)$$
$$(i=1,2,\cdots,7,8) \quad (3.4\text{-}4)$$

单元速度场表示为

$$\dot{\boldsymbol{u}}=\boldsymbol{N}\dot{\boldsymbol{u}}^e \quad (3.4\text{-}5)$$

其坐标变换形式为

$$\left.\begin{aligned}x(\xi,\eta,\zeta)&=\sum_i N_i(\xi,\eta,\zeta)x_i\\ y(\xi,\eta,\zeta)&=\sum_i N_i(\xi,\eta,\zeta)y_i\\ z(\xi,\eta,\zeta)&=\sum_i N_i(\xi,\eta,\zeta)z_i\end{aligned}\right\} \quad (3.4\text{-}6)$$

式中，x_i、y_i、z_i 为节点 i 的整体坐标。

1.1.3　单元应变率矩阵

通过形函数建立单元速度场与节点速度矢量之间关系，在此基础上可以导出用节点速度表示的单元内任一点应变速

率的表达式。

(1) 四节点等参单元

单元内的应变率常用几何方程计算，由于平面变形问题和轴对称问题采用相同的单元而且单元自由度相同，所以可以一并推导出统一的表达式。若令平面变形和轴对称问题中的应变率矢量分别为

$$\dot{\boldsymbol{\varepsilon}} = (\dot{\varepsilon}_x \quad \dot{\varepsilon}_y \quad \dot{\varepsilon}_z \quad \dot{\gamma}_{xy})^{\mathrm{T}} \text{ 和 } \dot{\boldsymbol{\varepsilon}} = (\dot{\varepsilon}_r \quad \dot{\varepsilon}_z \quad \dot{\varepsilon}_\theta \quad \dot{\gamma}_{rz})^{\mathrm{T}} \tag{3.4-7}$$

则

$$\dot{\boldsymbol{\varepsilon}} = \boldsymbol{L}\dot{\boldsymbol{u}} = \boldsymbol{L}\boldsymbol{N}\dot{\boldsymbol{u}}^e = \boldsymbol{B}\dot{\boldsymbol{u}}^e \tag{3.4-8}$$

式中，$\boldsymbol{L}$ 为微分算子矩阵；$\boldsymbol{B}$ 为应变率矩阵或几何矩阵。

而对应于平面应变和轴对称的微分算子矩阵分别为

$$\boldsymbol{L} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & \frac{\partial}{\partial y} \\ 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial x} \end{bmatrix}^{\mathrm{T}} \text{ 和 } \boldsymbol{L} = \begin{bmatrix} \frac{\partial}{\partial r} & 0 & \frac{1}{r} & \frac{\partial}{\partial z} \\ 0 & \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial r} \end{bmatrix}^{\mathrm{T}} \tag{3.4-9}$$

可以看出，平面变形问题与轴对称问题应变率矢量除第三分量不同外，其他分量只需将各有关下标 $(x,\ y)$ 与 $(r,\ z)$ 对调即可。为此，下面先给出轴对称问题的 $\boldsymbol{B}$ 矩阵公式。

$$\boldsymbol{B} = \boldsymbol{L}\boldsymbol{N} = (B_1 \quad B_2 \quad B_3 \quad B_4) \tag{3.4-10}$$

式中

$$\boldsymbol{B}_i = \begin{bmatrix} X_i & 0 \\ 0 & Y_i \\ K_i & 0 \\ Y_i & X_i \end{bmatrix} \tag{3.4-11}$$

$$\begin{Bmatrix} X_i \\ Y_i \end{Bmatrix} = \begin{Bmatrix} \frac{\partial N_i}{\partial r} \\ \frac{\partial N_i}{\partial z} \end{Bmatrix} = \boldsymbol{J}^{-1} \begin{Bmatrix} \frac{\partial N_i}{\partial \xi} \\ \frac{\partial N_i}{\partial \eta} \end{Bmatrix} \tag{3.4-12}$$

其中，$\boldsymbol{J}^{-1}$ 为雅可比矩阵 $\boldsymbol{J}$ 的逆矩阵，分别为

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial r}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial r}{\partial \eta} & \frac{\partial z}{\partial \eta} \end{bmatrix} \tag{3.4-13}$$

$$\boldsymbol{J}^{-1} = \frac{1}{|\boldsymbol{J}|} = \begin{bmatrix} \frac{\partial z}{\partial \eta} & -\frac{\partial z}{\partial \xi} \\ -\frac{\partial r}{\partial \eta} & \frac{\partial r}{\partial \xi} \end{bmatrix} \tag{3.4-14}$$

而 $|\boldsymbol{J}|$ 为雅可比（Jacobian）矩阵行列式，即

$$|\boldsymbol{J}| = \frac{\partial r \partial z}{\partial \xi \partial \eta} - \frac{\partial r \partial z}{\partial \eta \partial \xi} \tag{3.4-15}$$

而式（3.4-11）中的 K_i 项，对于平面变形问题恒为零，对式（3.4-11）只需将坐标量 $(r,\ z)$ 换为 $(x,\ y)$ 即可，

$$K_i = \begin{cases} 0 & \text{（平面变形）} \\ \frac{N_i}{r} & \text{（轴对称）} \end{cases} \tag{3.4-16}$$

(2) 三维单元

在直角坐标系下，三维问题的应变率矢量定义为

$$\dot{\boldsymbol{\varepsilon}} = (\dot{\varepsilon}_x \quad \dot{\varepsilon}_y \quad \dot{\varepsilon}_z \quad \dot{\gamma}_{xy} \quad \dot{\gamma}_{yz} \quad \dot{\gamma}_{zx})^{\mathrm{T}} \tag{3.4-17}$$

由几何方程，则

$$\dot{\boldsymbol{\varepsilon}} = \boldsymbol{L}\dot{\boldsymbol{u}} = \boldsymbol{L}\boldsymbol{N}\dot{\boldsymbol{u}}^e = \boldsymbol{B}\dot{\boldsymbol{u}}^e$$

式中，$\boldsymbol{L}$ 为三维微分算子矩阵

$$\boldsymbol{L} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial z} \\ 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial x} & \frac{\partial}{\partial z} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{bmatrix}^{\mathrm{T}} \tag{3.4-18}$$

应变率矩阵 $\boldsymbol{B}$

$$\boldsymbol{B} = \boldsymbol{L}\boldsymbol{N} = (B_1 \quad B_2 \quad \cdots \quad B_7 \quad B_8) \tag{3.4-19}$$

$$\boldsymbol{B}_i = \begin{bmatrix} X_i & 0 & 0 \\ 0 & Y_i & 0 \\ 0 & 0 & Z_i \\ Y_i & X_i & 0 \\ 0 & Z_i & Y_i \\ Z_i & 0 & X_i \end{bmatrix} \tag{3.4-20}$$

$$\begin{Bmatrix} X_i \\ Y_i \\ Z_i \end{Bmatrix} = \boldsymbol{J}^{-1} \begin{Bmatrix} \partial N_i / \partial \xi \\ \partial N_i / \partial \eta \\ \partial N_i / \partial \zeta \end{Bmatrix} \tag{3.4-21}$$

三维 Jacobian 变换矩阵 $\boldsymbol{J}$ 及逆矩阵 $\boldsymbol{J}^{-1}$ 分别为

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} & \frac{\partial z}{\partial \eta} \\ \frac{\partial x}{\partial \zeta} & \frac{\partial y}{\partial \zeta} & \frac{\partial z}{\partial \zeta} \end{bmatrix} \tag{3.4-22}$$

$$\boldsymbol{J}^{-1} = \frac{1}{|\boldsymbol{J}|} \begin{bmatrix} \frac{\partial y \partial z}{\partial \eta \partial \zeta} - \frac{\partial y \partial z}{\partial \zeta \partial \eta} & \frac{\partial y \partial z}{\partial \xi \partial \zeta} + \frac{\partial y \partial z}{\partial \zeta \partial \xi} & \frac{\partial y \partial z}{\partial \xi \partial \eta} - \frac{\partial y \partial z}{\partial \eta \partial \xi} \\ \frac{\partial x \partial z}{\partial \eta \partial \zeta} + \frac{\partial x \partial z}{\partial \zeta \partial \eta} & \frac{\partial x \partial z}{\partial \xi \partial \zeta} - \frac{\partial x \partial z}{\partial \zeta \partial \xi} & \frac{\partial x \partial z}{\partial \xi \partial \eta} + \frac{\partial x \partial z}{\partial \eta \partial \xi} \\ \frac{\partial x \partial y}{\partial \eta \partial \zeta} + \frac{\partial x \partial y}{\partial \zeta \partial \eta} & \frac{\partial x \partial y}{\partial \zeta \partial \xi} - \frac{\partial x \partial y}{\partial \xi \partial \zeta} & \frac{\partial x \partial y}{\partial \xi \partial \eta} - \frac{\partial x \partial y}{\partial \eta \partial \xi} \end{bmatrix} \tag{3.4-23}$$

(3) 等效应变率和体积应变率矩阵形式

由等效应变率公式，即

$$\dot{\bar{\varepsilon}} = \left(\frac{2}{3}\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}\right)^{\frac{1}{2}} = (\dot{\boldsymbol{\varepsilon}}^{\mathrm{T}}\boldsymbol{D}\dot{\boldsymbol{\varepsilon}})^{\frac{1}{2}} = ((\dot{\boldsymbol{u}}^e)^{\mathrm{T}}\boldsymbol{A}\dot{\boldsymbol{u}}^e)^{\frac{1}{2}} \tag{3.4-24}$$

式中矩阵 $\boldsymbol{D}$ 为对角阵，对角元素值取 2/3 或 1/3，分别对应于应变率矢量 $\dot{\boldsymbol{\varepsilon}}$ 中的正应变率和切应变率分量，其阶数与 $\dot{\boldsymbol{\varepsilon}}$ 中分量数相同。

矩阵 $\boldsymbol{A}$ 定义为

$$\boldsymbol{A} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{B} \tag{3.4-25}$$

由体积应变率公式，得

$$\dot{\boldsymbol{\varepsilon}}_V = \boldsymbol{C}^{\mathrm{T}}\dot{\boldsymbol{\varepsilon}} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{B}\dot{\boldsymbol{u}}^e \tag{3.4-26}$$

其中，矢量 $\boldsymbol{C}$ 定义为

$$\boldsymbol{C} = \begin{cases} (1 \quad 1 \quad 0 \quad 0)^{\mathrm{T}} & \text{（平面变形）} \\ (1 \quad 1 \quad 1 \quad 0)^{\mathrm{T}} & \text{（轴对称）} \\ (1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0)^{\mathrm{T}} & \text{（三维问题）} \end{cases} \tag{3.4-27}$$

1.1.4　泛函离散化

将变形体 V 在定义域上分割成 n 个节点和 m 个单元，建立上述各泛函的离散形式。设 Π^e 表示单元 e 上的泛函，则总体泛函 Π 为

$$\Pi = \sum_e \Pi^e \tag{3.4-28}$$

1.1.5　模拟分析步骤

在材料变形过程中，塑性区的大小和形状及其内部场变量（如速度场、应力场、应变场和温度场等）一般都随变形过程的进行而变化。采用刚塑性有限元法模拟变形过程的主要步骤如下：

1) 建立有限元模拟初始模型（包括模具型腔描述、网格划分、材料模型和边界条件等）；

2) 构造或生成初始速度场；

3) 计算各单元刚度矩阵和残余力矢量，并进行斜约束处理；

4) 形成整体刚度矩阵和残余力矢量，并引入速度约束条件消除奇异性；

5) 解整体刚度方程并检查收敛情况，若收敛转入第 6 步，反之重复第 3 步～第 5 步；

6）由几何方程和塑性本构关系求出应变率和应变场；

7）确定增量变形时间步长 Δt_m，更新工件构形、应变场和材料性能及接触边界信息；

8）若预定变形未完成，则重复第 3）~第 7）步直到结束。

1.2 刚塑性变分原理

为了确定塑性加工过程中的力能参数、形变参数以及应力应变场，必须在一定的初始条件和边界条件下求解有关方程，也就是解塑性加工力学的边值问题。变分原理是求解刚塑性有限元法的理论基础，把塑性偏微分方程组的求解问题变成了泛函极值问题，建立有限元法的基本方程。

1.2.1 刚塑性变形的边值问题

固体力学中的塑性变形问题是一个边值问题，刚塑性变形的边值问题可作如下描述：设一刚塑性体，体积为 V，表面边界为 S，在表面力 p_i 作用下整个变形体处于塑性状态，表面 S 分为 S_p 和 S_v 两部分，其中 S_p 上给定表面力 p_i，S_v 上给定速度 v_i。该问题称之为刚塑性边值问题，它由以下塑性方程和边界条件定义，即

1）静力平衡方程

$$\sigma_{ij,j}=0 \qquad (3.4\text{-}29)$$

2）应变协调方程（几何方程）

$$\dot{\varepsilon}_{ij}=\frac{1}{2}(v_{i,j}+v_{j,i}) \qquad (3.4\text{-}30)$$

3）塑性本构方程

$$\dot{\varepsilon}_{ij}=\frac{3\dot{\bar{\varepsilon}}}{2\bar{\sigma}}\sigma'_{ij} \qquad (3.4\text{-}31)$$

式中，$\bar{\sigma}$ 为等效应力，$\bar{\sigma}=\sqrt{\frac{3}{2}\sigma'_{ij}\sigma'_{ij}}$；$\dot{\bar{\varepsilon}}$ 为等效应变率，$\dot{\bar{\varepsilon}}=\sqrt{\frac{2}{3}\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}}$。

4）Mises 屈服条件

$$\bar{\sigma}=\bar{\sigma}(\bar{\varepsilon},\ T) \qquad (3.4\text{-}32)$$

式中，$\bar{\sigma}$ 为材料屈服应力；$\bar{\varepsilon}$ 为等效应变；T 为温度。

5）体积不可压缩条件

$$\dot{\varepsilon}_V=\dot{\varepsilon}_{ii}=0 \qquad (3.4\text{-}33)$$

6）边界条件

应力边界条件　$\sigma_{ij}n_j=p_i \quad S\in S_p$　(3.4-34)

速度边界条件　$v_i=v_i^0 \quad S\in S_v$　(3.4-35)

式中，n_j 为 S_p 表面上任一点处单位外法线矢量的分量。

1.2.2 刚塑性变形的变分原理

刚塑性有限元的理论基础是马克夫变分原理（Makov Principle），可表述为：对于刚塑性边值问题，在满足变形几何方程、体积不可压缩条件和边界位移条件的一切运动容许速度场 v_i^* 中，使泛函

$$\Pi=\int_V\bar{\sigma}\dot{\bar{\varepsilon}}^*\,dV-\int_{S_p}p_iv_i^*\,dS \qquad (3.4\text{-}36)$$

取驻值，即一阶变分 $\delta\Pi=0$ 时的 v_i^* 为问题的真实解。

应用 Markov 变分原理求解的难点之一就是构造容许速度场 v_i^*。一般来说，选取满足位移速度边界条件式（3.4-35）的容许速度比较容易，而要满足体积不可压缩条件式（3.4-33）则非常困难。同时，由于刚塑性材料模型不计弹性变形部分，并采用体积不可压缩假设，就难以确定静压力 σ_m，因而求不出变形体内的应力分布 σ_{ij}。因此，在采用变分法求解刚塑性问题时，通常采用某种方法将体积不变条件引入到泛函表达式中，作为对体积变化的一个约束项。典型的处理方法有拉格朗日乘子法、罚函数法和可压缩特性法。

（1）拉格朗日乘子法

刚塑性有限元法中的拉格朗日乘子法的数学基础是数学分析中多元函数的条件极值理论，对于 Markov 变分原理，把体积不可压缩条件式（3.4-33）用拉格朗日乘子 λ 引入泛函式（3.4-36），构造新的泛函如下

$$\Pi=\int_V\bar{\sigma}\dot{\bar{\varepsilon}}\,dV+\int_V\lambda\dot{\varepsilon}_V\,dV-\int_{S_p}p_iv_i\,dS \qquad (3.4\text{-}37)$$

对于单元 e，其体积为 V^e，力边界为 S_p^e，则单元泛函为

$$\Pi^e=\int_{V^e}\bar{\sigma}[(\dot{\boldsymbol{u}}^e)^T\boldsymbol{A}\dot{\boldsymbol{u}}^e]^{\frac{1}{2}}\,dV+\int_{V^e}\lambda^e\boldsymbol{C}^T\boldsymbol{B}\dot{\boldsymbol{u}}^e\,dV-\int_{S_p^e}\boldsymbol{P}^T\boldsymbol{N}\dot{\boldsymbol{u}}^e\,dV \qquad (3.4\text{-}38)$$

对于一切满足几何方程和位移速度边界条件的容许速度场，其真实解使式（3.4-37）取极值，即满足

$$\delta\Pi=\int_V\bar{\sigma}\delta\dot{\bar{\varepsilon}}\,dV+\int_V\lambda\delta\dot{\varepsilon}_V\,dV+\int_V\delta\lambda\dot{\varepsilon}_V\,dV-\int_{S_p}p_i\delta v_i\,dS=0 \qquad (3.4\text{-}39)$$

速度场为真解时的拉格朗日乘子 λ 为静水压力

$$\lambda=\sigma_m \qquad (3.4\text{-}40)$$

（2）罚函数法

罚函数法的基本思想是用一个足够大的正数 α，把体积不可压缩条件引入泛函式（3.4-36），构造出一个新泛函，即

$$\Pi=\int_V\bar{\sigma}\dot{\bar{\varepsilon}}\,dV+\frac{\alpha}{2}\int_V\dot{\varepsilon}_V^2\,dV-\int_{S_p}p_iv_i\,dS \qquad (3.4\text{-}41)$$

对于单元 e，其体积为 V^e，力边界为 S_p^e，则单元泛函为

$$\Pi^e=\int_{V^e}\bar{\sigma}[(\dot{\boldsymbol{u}}^e)^T\boldsymbol{A}\dot{\boldsymbol{u}}^e]^{\frac{1}{2}}\,dV+\frac{\alpha}{2}\int_{V^e}(\boldsymbol{C}^T\boldsymbol{B}\dot{\boldsymbol{u}}^e)^2\,dV-\int_{S_p^e}\boldsymbol{P}^T\boldsymbol{N}\dot{\boldsymbol{u}}^e\,dS \qquad (3.4\text{-}42)$$

则对于一切满足几何方程和位移速度边界条件的容许速度场，其真实解使式（3.4-41）取极值，即满足

$$\delta\Pi=\int_V\bar{\sigma}\delta\dot{\bar{\varepsilon}}\,dV+\alpha\int_V\dot{\varepsilon}_V\delta\dot{\varepsilon}_V\,dV-\int_{S_p}p_i\delta v_i\,dS=0 \qquad (3.4\text{-}43)$$

这里的罚函数法源于最优化原理中的罚函数法，具有数值解法的特征。它的作用原理是，当 v_i 偏离真实解时，惩罚项值很大，相当于对速度解违反约束条件施加一种“惩罚”作用；而随着 v_i 接近真解，罚项的作用也随之降低。应当指出，从理论上讲拉格朗日乘子法是精确的。对于罚函数法，只有当 α 趋于无穷大时，$\dot{\varepsilon}_V$ 才趋于零。但在实际应用时，α 不可能取无穷大。计算实践表明，α 的取值大小对解有很大的影响。若 α 取值太小，则体积不可压缩条件施加不当，以致降低计算精度；若 α 取值过大，则有限元刚度方程会出现病态，甚至不能求解。因此，α 取值应适宜，通常取 $\alpha=10^5\sim10^7$ 较好。

将式（3.4-43）与式（3.4-40）比较，容易得出，当 $\dot{v}_i$ 为真实解时，静水压力 σ_m 为

$$\sigma_m=\alpha\dot{\varepsilon}_V \qquad (3.4\text{-}44)$$

泛函式（3.4-41）罚项的被积函数采用 $\dot{\varepsilon}_V^2$ 形式，它要求 $|\dot{\varepsilon}_V|$ 在域内处处满足体积不可压缩条件，才能保证罚项总值很小。而实际应用中发现这样的约束条件过于严格而不易达到，可通过适当放松约束条件处理。目前常用的方法有简化积分法和修正罚函数法。

所谓简化积分法，只要求单元形心点 $|\dot{\varepsilon}_V|$ 很小，而对其他点不做要求。对于平面问题简化积分值与积分原值相同。所以，这种简化积分反映了单元内部体积变化的平均效应，即只需单元整体满足体积不可压缩条件就可。修正罚函数法该方法则是通过对罚项构造形式的修改，来达到放松约束的

目的。修改后罚项的泛函表示为

$$\Pi = \int_V \bar{\sigma}\dot{\bar{\varepsilon}} \mathrm{d}V + \frac{\alpha}{2V}\left(\int_V \dot{\varepsilon}_V \mathrm{d}V\right)^2 - \int_{S_p} p_i v_i \mathrm{d}S \tag{3.4-45}$$

式中的第二项即为修改后的罚项，它的直观意义是要求单元体积变化的平均值很小。因此，尽管修正的罚项与前者在形式上不同，但它们的内部是类似的。

同理，对于泛函式（3.4-45），若 v_i 取真实解时，静水压力 σ_m 为

$$\sigma_m = \frac{\alpha}{V}\int_V \dot{\varepsilon}_V \mathrm{d}V \tag{3.4-46}$$

实际应用时，两种放松约束的方法都能达到同样的目的。对于修正罚函数法，各体积积分运算在同一数值积分格式下进行，所以程序可适当简化。而对于简化积分法，由于减少了积分点，因而降低了运算次数，提高了计算效率。

(3) 可压缩特性法

可压缩特性法是假设材料的相对密度不是100%，具有一定的可压缩性。因此，它可以直接从应变率求出应力分量，达到与采用拉格朗日乘子法或罚函数法相同的目的。

刚塑性可压缩材料的变分原理仍采用Markov变分原理的形式，即真实速度场使以下泛函取极值

$$\Pi = \int_V \bar{\sigma}\dot{\bar{\varepsilon}} \mathrm{d}V - \int_{S_p} p_i v_i \mathrm{d}S \tag{3.4-47}$$

对于单元 e，其体积为 V^e，力边界为 S_p^e，则单元泛函为

$$\Pi^e = \int_{V^e} \bar{\sigma}\left[(\dot{\boldsymbol{u}}^e)^{\mathrm{T}}\boldsymbol{A}\dot{\boldsymbol{u}}^e + \frac{1}{g}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{B}\dot{\boldsymbol{u}}^e)^2\right]^{\frac{1}{2}} \mathrm{d}V - \int_{S_p^e} \boldsymbol{P}^{\mathrm{T}}\boldsymbol{N}\dot{\boldsymbol{u}}^e \mathrm{d}S \tag{3.4-48}$$

可压缩特性法考虑材料的屈服应力对静压力有少许依赖性，其等效应力定义为

$$\bar{\sigma} = \sqrt{3J'_2 + g\sigma_m^2} = \sqrt{\frac{3}{2}\sigma'_{ij}\sigma'_{ij} + g\sigma_m^2} \tag{3.4-49}$$

式中，g 为可压缩参数，取 $g = 0.01 \sim 0.000\ 1$。

与Mises屈服准则类似，可以在主应力空间表示刚塑性可压缩材料屈服准则的屈服表面形状，其屈服表面是一个内切于Mises圆柱的椭球面。可压缩参数 g 的值越小，就越接近Mises屈服条件，如果 $g=0$，二者就完全一致。

采塑性势定义可以推导出刚塑性可压缩材料的本构关系，即

$$\dot{\varepsilon}_{ij} = \frac{3\dot{\bar{\varepsilon}}}{2\bar{\sigma}}\left(\sigma'_{ij} + \frac{2}{9}g\sigma_m\delta_{ij}\right) \tag{3.4-50}$$

或者

$$\sigma_{ij} = \frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}}\left[\frac{2}{3}\dot{\varepsilon}_{ij} + \dot{\varepsilon}_V\left(\frac{1}{g} - \frac{2}{9}\right)\sigma_{ij}\right] \tag{3.4-51}$$

式中，$\dot{\bar{\varepsilon}}$ 为刚塑性可压缩材料的等效应变率，$\dot{\bar{\varepsilon}} = \sqrt{\frac{2}{3}\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij} + \frac{1}{g}\dot{\varepsilon}_V^2}$；$\dot{\varepsilon}_V$ 为刚塑性可压缩材料的体积应变率，$\dot{\varepsilon}_V = \frac{\dot{\bar{\varepsilon}}}{\bar{\sigma}}g\sigma_m$。

这样，若 g 值给定，就可以直接由应变率 $\dot{\varepsilon}_{ij}$ 从式（3.4-51）计算出应力场 σ_{ij}。

比较三种体积不可压缩约束处理方法可知，拉格朗日乘子法引入了一个具有物理意义的未知参数 λ，增加了方程未知量的数量，因而增加了有限元刚度方程数，使得刚度矩阵半带宽增大。与罚函数法和可压缩特性法相比，对同样问题要大大增加计算时间。在刚度矩阵方面，拉格朗日乘子法的非零元素不呈带状分布，而罚函数法和可压缩特性法都为明显带状分布。故拉格朗日乘子法会增加计算机贮存空间，降低计算效率，但是却有较好的收敛性和精度。

可压缩法考虑了静水压力对体积变化率的影响，因而比较适合于多孔的可压缩材料的计算。在罚函数法中，如果 α 选得过大，有可能使刚度方程组出现病态，可能得不到收敛解。同样，如果可压缩特性法中的 g 值选得过小，也有可能不收敛。相比之下，拉格朗日乘子法的乘子 λ 是经过求解优化得到的，精确地计入了体积不可压缩条件，不用预先设定。

总的来说，这三种方法都带有某种程度的近似性，在工程应用当中，根据实际情况恰当运用，都能够达到满意的效果。

1.3　刚塑性有限元列式

利用有限元法可以将变形体分为若干个单元，要求在单元内保持场函数连续性，依次建立单元泛函，将单元泛函集成得到整体的泛函，然后对整体泛函求驻值，就可以得到问题的数值解。求出速度场后，即可根据各塑性方程得到应变速率场、应力应变场以及位移场等等。

1.3.1　拉格朗日乘子法有限元列式

将单元泛函 Π^e 集成为总体泛函 Π，有

$$\Pi = \sum_e \Pi^e(\dot{u}^e, \lambda^e) = \Pi(\dot{\boldsymbol{u}}, \boldsymbol{\lambda}) \tag{3.4-52}$$

式中 $\dot{\boldsymbol{u}}$ 和 $\boldsymbol{\lambda}$ 分别为总体节点速度向量和拉格朗日乘子向量，定义为

$$\begin{aligned} \dot{\boldsymbol{u}} &= (\dot{u}_1 \quad \dot{u}_2 \quad \dot{u}_3 \quad \cdots \quad \dot{u}_{k-1} \quad \dot{u}_k)^{\mathrm{T}} \\ \boldsymbol{\lambda} &= (\lambda^1,\ \lambda^2,\ \cdots,\ \lambda^m)^{\mathrm{T}} \end{aligned} \tag{3.4-53}$$

式中，k 等于总节点数 n 与问题维数之积。

经过上述的离散化处理，泛函已变为节点速度 $\dot{\boldsymbol{u}}$ 和拉格朗日乘子 $\boldsymbol{\lambda}$ 的函数，其驻值条件为

$$\begin{aligned} \partial\Pi &= \sum_e \left[(\delta\dot{\boldsymbol{u}}^e)^{\mathrm{T}}\frac{\partial\Pi^e}{\partial\dot{\boldsymbol{u}}^e} + (\delta\boldsymbol{\lambda}^e)^{\mathrm{T}}\frac{\partial\Pi^e}{\partial\boldsymbol{\lambda}^e}\right] \\ &= (\delta\dot{\boldsymbol{u}})^{\mathrm{T}}\frac{\partial\Pi}{\partial\dot{\boldsymbol{u}}} + (\delta\boldsymbol{\lambda})^{\mathrm{T}}\frac{\partial\Pi}{\partial\boldsymbol{\lambda}} = 0 \end{aligned} \tag{3.4-54}$$

由于 $\partial\dot{\boldsymbol{u}}$ 和 $\partial\boldsymbol{\lambda}$ 的任意性，显然上式成立的条件是

$$\begin{cases} \dfrac{\partial\Pi}{\partial\dot{u}} = \sum_e \dfrac{\partial\Pi^e}{\partial\dot{u}^e} = 0 \\ \dfrac{\partial\Pi}{\partial\boldsymbol{\lambda}} = \sum_e \dfrac{\partial\Pi^e}{\partial\lambda^e} = 0 \end{cases} \tag{3.4-55}$$

这样，问题就归结为单元泛函对本单元的场变量求导，并按式（3.4-55）集成后得到总体刚度方程，即离散形式的平衡方程如下

$$\begin{Bmatrix} \dfrac{\partial\Pi}{\partial\dot{u}} \\ \dfrac{\partial\Pi}{\partial\lambda} \end{Bmatrix} = \begin{bmatrix} \sum_e\int_{V^e}\dfrac{\bar{\sigma}}{\dot{\bar{\varepsilon}}}\boldsymbol{A}\mathrm{d}V & \sum_e\boldsymbol{Q} \\ \sum_e\boldsymbol{Q}^{\mathrm{T}} & 0 \end{bmatrix}\begin{Bmatrix}\dot{\boldsymbol{u}} \\ \boldsymbol{\lambda}\end{Bmatrix} - \begin{Bmatrix}\sum_e\boldsymbol{P} \\ 0\end{Bmatrix} = \{0\} \tag{3.4-56}$$

其中，若外力为非摩擦力，矢量 $\boldsymbol{P}$ 可以用下式表示

$$\begin{aligned} \boldsymbol{P} &= \int_{S_p^e}\boldsymbol{N}^{\mathrm{T}}\boldsymbol{P}\mathrm{d}S \\ \boldsymbol{Q} &= \int_{V^e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}\mathrm{d}V \end{aligned} \tag{3.4-57}$$

式（3.4-56）是一个关于节点速度矢量 $\dot{\boldsymbol{u}}$ 和拉格朗日乘子矢量 $\boldsymbol{\lambda}$ 的非线性方程组，即系数矩阵与节点速度相关，所以方程不能直接求解。实际上，塑性变形的本质决定了任何一种塑性有限元方程都是非线性的，必须线性化处理，用迭代法求解。

通常采用Newton－Raphson法线性化后迭代求解，其要点是：假定泛函在求解区域内连续，并存在对节点速度的各阶导数，若 $\dot{u}_{n-1}$ 为第 $n-1$ 次迭代的近似解，则将泛函的一

阶偏导数在 $\dot{u}_{n-1}$ 的邻域展成泰勒级数，并仅取线性项，从而得到以 $\Delta\dot{u}_n$ 为未知数的线性方程组，其迭代递推公式如下

$$\begin{cases}\left[\dfrac{\partial^2 \Pi}{\partial\dot{\boldsymbol{u}}\partial\dot{\boldsymbol{u}}^{\mathrm{T}}}\right]_{n-1}\Delta\dot{u}_n = -\left\{\dfrac{\partial \Pi}{\partial\dot{\boldsymbol{u}}}\right\}_{n-1} \\ \left[\dfrac{\partial^2 \Pi}{\partial\boldsymbol{\lambda}\partial\dot{\boldsymbol{u}}^{\mathrm{T}}}\right]_{n-1}\Delta\dot{u}_n = -\left\{\dfrac{\partial \Pi}{\partial\boldsymbol{\lambda}}\right\}_{n-1}\end{cases} \tag{3.4-58}$$

或

$$\begin{cases}\left[\sum_e \dfrac{\partial^2 \Pi^e}{\partial\dot{\boldsymbol{u}}^e\partial(\dot{u}^e)^{\mathrm{T}}}\right]_{n-1}\sum_e \Delta\dot{u}_n^e = -\left\{\sum_e \dfrac{\partial \Pi^e}{\partial\dot{\boldsymbol{u}}^e}\right\}_{n-1} \\ \left[\sum_e \dfrac{\partial^2 \Pi^e}{\partial\boldsymbol{\lambda}^e\partial(\dot{\boldsymbol{u}}^e)^{\mathrm{T}}}\right]_{n-1}\sum_e \Delta\dot{u}_n^e = -\left\{\sum_e \dfrac{\partial \Pi^e}{\partial\boldsymbol{\lambda}^e}\right\}_{n-1}\end{cases}$$

$$\dot{u}_n = \dot{u}_{n-1} + \beta\Delta\dot{u}_n$$

式中，n 表示迭代次数，β 成为减速稀疏或阻尼因子，$0<\beta\leqslant 1$。β 的作用在于提高 Newton 迭代法的收敛性，但会降低其收敛速度，亦即增加迭代次数。若同时兼顾迭代求解的收敛及其速度，通常迭代初期 β 值应小些，随着求解过程的进行 β 值逐渐增大直至收敛为止。

$$\begin{bmatrix}\sum_e \bar{\sigma}\boldsymbol{K}_1 & \sum_e \boldsymbol{Q} \\ \sum_e \boldsymbol{Q}^{\mathrm{T}} & 0\end{bmatrix}_{n-1}\begin{Bmatrix}\Delta\dot{\boldsymbol{u}} \\ \boldsymbol{\lambda}\end{Bmatrix}_n = \begin{Bmatrix}\sum_e (\boldsymbol{P}-\bar{\sigma}\boldsymbol{H}_1) \\ \sum_e \boldsymbol{Q}^{\mathrm{T}}\dot{\boldsymbol{u}}^e\end{Bmatrix}_{n-1} \tag{3.4-59}$$

式中

$$\boldsymbol{K}_1 = \int_{V^e}\frac{1}{\dot{\bar{\varepsilon}}}\left[\boldsymbol{A} - \frac{1}{\dot{\bar{\varepsilon}}^2}\boldsymbol{b}\boldsymbol{b}^{\mathrm{T}}\right]\mathrm{d}V$$

$$\boldsymbol{H}_1 = \int_{V^e}\frac{1}{\dot{\bar{\varepsilon}}}\boldsymbol{b}\,\mathrm{d}V$$

$$\boldsymbol{b} = \boldsymbol{A}\dot{\boldsymbol{u}}^e$$

应当注意，拉格朗日法迭代公式（3.4-58）说明，泛函一阶变分方程仅是 $\dot{\boldsymbol{u}}$ 的非线性方程组，它关于 $\boldsymbol{\lambda}$ 是线性的，因此迭代时 $\dot{\boldsymbol{u}}$ 进行逐步修正计算，而 $\boldsymbol{\lambda}$ 则伴随求解。

1.3.2　罚函数法和可压缩特性法有限元列式

根据前述相同的方法，可以推导出罚函数法和可压缩特性法的有限元公式。

(1) 罚函数法

$$\left[\sum_e(\bar{\sigma}\boldsymbol{K}_1 + \alpha\boldsymbol{M})\right]_{n-1}\Delta\dot{u}_n = \left\{\sum_e(\boldsymbol{P} - \bar{\sigma}\boldsymbol{H}_1 - \alpha\boldsymbol{H}_2)\right\}_{n-1} \tag{3.4-60}$$

式中

$$\boldsymbol{M} = \int_{V^e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C})^{\mathrm{T}}\mathrm{d}V$$

$$\boldsymbol{H}_2 = \int_{V^e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C})^{\mathrm{T}}\dot{\boldsymbol{u}}^e\mathrm{d}V$$

$\boldsymbol{M}$ 和 $\boldsymbol{H}_2$ 是与体积变化相关的量，而 $\boldsymbol{K}_1$、$\boldsymbol{P}$、$\boldsymbol{H}_1$ 则与拉格朗日乘子法相同。

(2) 可压缩特性法

$$\left[\sum_e(\bar{\sigma}\boldsymbol{K}_2)\right]_{n-1}\Delta\dot{u}_n = \left\{\sum_e(\boldsymbol{P} - \bar{\sigma}\boldsymbol{H}_3)\right\}_{n-1} \tag{3.4-61}$$

式中

$$\boldsymbol{K}_2 = \int_{V^e}\frac{1}{\dot{\bar{\varepsilon}}}\left[\boldsymbol{A} + \frac{1}{g}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C})^{\mathrm{T}} - \frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}}{\dot{\bar{\varepsilon}}^2}\right]\mathrm{d}V$$

$$\boldsymbol{H}_3 = \int_{V^e}\frac{1}{\dot{\bar{\varepsilon}}}\boldsymbol{w}\,\mathrm{d}V$$

$$\boldsymbol{w} = \left[\boldsymbol{A} + \frac{1}{g}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C})^{\mathrm{T}}\right]\dot{\boldsymbol{u}}^e$$

$$\boldsymbol{A} = \frac{2}{3}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \frac{1}{g}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C})(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C})^{\mathrm{T}}$$

$\boldsymbol{P}$ 与拉格朗日乘子法相同。

1.4　刚黏塑性有限元法

许多材料的塑性变形时，流动应力与变形速度有明显的相关性，如在高温变形和超塑性成形时，就要考虑时间和变形速率的影响。

1.4.1　刚黏塑性变分原理

刚黏塑性边值问题可以叙述如下：设在准静态变形的某一阶段，变形体的形状、温度及材料参数等的瞬时值已确定。设该变形体体积为 V，表面为 S，并且 S_p 和 S_v 上分别给定了应力边界条件和位移速度边界条件，则应力场和速度场的解应该满足平衡方程、协调方程和体积不可压缩条件。此时，变形体处于黏塑性状态，可以采用 Hill 提出的黏塑性材料的变分原理解此边值问题。

刚黏塑性边值问题的描述方程和边界条件与刚塑性问题的方程相同，但是，黏塑性材料的塑性屈服应力却更多地依赖于应变速率 $\dot{\varepsilon}$ 的大小，所以其屈服应力可以表示为

$$\bar{\sigma} = \bar{\sigma}(\bar{\varepsilon}, \dot{\bar{\varepsilon}}, T) = \sqrt{J_2'} - K \tag{3.4-62}$$

对于刚粘塑性边值问题，在满足几何方程、体积不可压缩条件及位移速度边界条件的一切容许速度场中，其真实解使下列泛函取极小值。

$$\Pi = \int_V E(\dot{\varepsilon}_{ij})\mathrm{d}V - \int_{S_p} p_i v_i \mathrm{d}S \tag{3.4-63}$$

式中，$E(\dot{\varepsilon}_{ij})$ 为塑性变形功率函数，它与材料模型公式密切相关。

功函数 $E(\dot{\varepsilon}_{ij})$ 可按照下式计算得到

$$E(\dot{\varepsilon}_{ij}) = \int_0^{\dot{\varepsilon}_{ij}}\sigma_{ij}'\mathrm{d}\dot{\varepsilon}_{ij} = \int_0^{\dot{\bar{\varepsilon}}}\bar{\sigma}\mathrm{d}\dot{\bar{\varepsilon}} \tag{3.4-64}$$

$E(\dot{\varepsilon}_{ij})$ 的含义表示应变速率从零增加到 $\dot{\varepsilon}$ 时，流动应力 $\bar{\sigma}$ 在单位体积内做的功。这样，只要给定刚粘塑性材料模型的公式，就可以由上式得到相应的功函数。

对于体积不可压缩条件的引入，同样可以采用拉格朗日乘子法、罚函数法和可压缩特性法中的任何一种方法，相应的修正泛函分别如下。

1) 拉格朗日乘子法

$$\Pi = \int_V E(\dot{\varepsilon})\mathrm{d}V + \int_V \lambda\dot{\varepsilon}_V\mathrm{d}V - \int_{S_p} p_i v_i \mathrm{d}S \tag{3.4-65}$$

2) 罚函数法

$$\Pi = \int_V E(\dot{\varepsilon})\mathrm{d}V + \frac{a}{2}\int_V \lambda\dot{\varepsilon}_V^2\mathrm{d}V - \int_{S_p} p_i v_i \mathrm{d}S \tag{3.4-66}$$

3) 可压缩特性法

$$\Pi = \int_V E(\dot{\varepsilon})\mathrm{d}V - \int_{S_p} p_i v_i \mathrm{d}S \tag{3.4-67}$$

1.4.2　刚黏塑性有限元列式

以拉格朗日乘子法为例，推导刚黏塑性有限元公式。根据修正的泛函，经离散化，单元 e 的泛函为

$$\Pi^e = \int_{V^e}E(\dot{\varepsilon})\mathrm{d}V + \lambda^e\int_{V^e}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{B}\dot{\boldsymbol{u}}^e\mathrm{d}V - \int_{S_p^e}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{N}\dot{\boldsymbol{u}}^e\mathrm{d}S \tag{3.4-68}$$

求出单元泛函的一阶、二阶偏导数为

$$\frac{\partial \Pi^e}{\partial \dot{u}^e} = \int_{V^e}\frac{\partial E}{\partial \dot{\varepsilon}}\frac{\partial \dot{\bar{\varepsilon}}}{\partial \dot{u}^e}\mathrm{d}V + \lambda^e\int_{V^e}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}\mathrm{d}V - \int_{S_p^e}\boldsymbol{N}^{\mathrm{T}}\boldsymbol{p}\mathrm{d}S$$
$$= \int_{V^e}\frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}}\boldsymbol{A}\dot{\boldsymbol{u}}\mathrm{d}V + \boldsymbol{\lambda}^e\boldsymbol{Q} - \boldsymbol{P} = \boldsymbol{H}_1' + \boldsymbol{\lambda}^e\boldsymbol{Q} - \boldsymbol{P} \tag{3.4-68a}$$

$$\frac{\partial \Pi^e}{\partial \lambda^e} = \int_{V^e} C^T B \dot{u}^e \mathrm{d}V = Q^T \dot{u}^e \quad (3.4\text{-}68b)$$

$$\frac{\partial^2 \Pi^e}{\partial \dot{u}^e \partial (\dot{u}^e)^T} = \int_{V^e} \frac{1}{\bar{\dot{\varepsilon}}}[\bar{\sigma} A + \frac{1}{\bar{\dot{\varepsilon}}}(\frac{\partial \bar{\sigma}}{\partial \bar{\dot{\varepsilon}}} - \frac{\bar{\sigma}}{\bar{\dot{\varepsilon}}}) b b^T] \mathrm{d}V = K'_1 \quad (3.4\text{-}68c)$$

$$\frac{\partial^2 \Pi^e}{\partial \lambda^e \partial (\dot{u}^e)^T} = Q^T \quad (3.4\text{-}68d)$$

利用上述式（3.4-68a）～式（3.4-68d），代入 Newton-Raphson 法展开的公式中，最后得到总刚方程，即

$$\begin{bmatrix} \sum_e K'_1 & \sum_e Q \\ \sum_e Q^T & 0 \end{bmatrix}_{n-1} \begin{Bmatrix} \Delta \dot{u} \\ \lambda \end{Bmatrix}_n = \begin{Bmatrix} \sum_e (P - H'_1) \\ \sum_e Q^T \dot{u}^e \end{Bmatrix}_{n-1} \quad (3.4\text{-}69)$$

式中

$$K'_1 = \int_{V^e} \frac{1}{\bar{\dot{\varepsilon}}}[\bar{\sigma} A + \frac{1}{\bar{\dot{\varepsilon}}}(\frac{\partial \bar{\sigma}}{\partial \bar{\dot{\varepsilon}}} - \frac{\bar{\sigma}}{\bar{\dot{\varepsilon}}}) b b^T] \mathrm{d}V$$

$$H'_1 = \int_{V^e} \frac{\bar{\sigma}}{\bar{\dot{\varepsilon}}} b \mathrm{d}V$$

$$b = B^T B \dot{u}^e$$

罚函数法和可压缩特性法中泛函均只是节点速度 $\dot{u}$ 的函数，采用上述同样的方法可以推导出它们的有限元公式。

1.5　体积成形过程关键技术

在体积成形过程中，材料在模具的作用下经过连续变形，由初始简单的形状成形为轮廓复杂的工件，变形体内各个节点的相对位置不断改变，材料与模具型腔表面也存在较大的相对运动。因此，要准确模拟整个变形过程，必须解决几个关键技术，如模具型腔的几何描述、动态边界条件的处理、网格的划分与重新划分，另外还包括刚性区的处理、初始速度场生成、摩擦条件的处理等。

1.5.1　模具型腔的几何描述

在体积成形中，对模具型腔的描述是后续相关计算的基础。由于模具的形状种类繁多，型腔曲面复杂，因此模具型腔的几何描述方法对于计算精度和效率影响很大。目前模具型腔的描述方法有点阵法、代数法、网格法以及参数曲面和实体造型相结合的方法等。

点阵方式是采用一些排列规则的模具型腔表面上的点来近似描述模具型腔表面。该方法使塑性成形有限元求解过程中的动态接触处理变得容易，但由于描述精度低，且只能近似描述表面的几何信息，现已不常用。

代数法是采用解析式来描述简单的模具型腔曲面，动态接触处理转变为对简单方程组的求解。但对于型腔曲面比较复杂的模具，该方法只能用圆柱面、圆锥面及球面等简单的解析曲面来逼近模具的复杂表面，这使得描述精度降低且复杂度增加。

网格法则是将模具型腔表面离散为空间平面三角形或者四边形有限元单元面片，用空间平面近似代替曲面，有限元求解过程中的动态接触处理归结为直线和平面求交或直线和曲面求交问题。显然，离散的面片越多，离散精度越高，其所描述的信息就越多。这种方法的优点是所描述的模具型腔表面的几何信息可以准确地在不同的有限元分析系统间进行数据交换，缺点是描述精度不够高。对于精密塑性成形过程的数值模拟，这种方法必须把模具表面离散为数量很多的细小单元，从而降低了动态接触处理时的求交搜索效率。并且，离散后的模具表面不再光滑连续，这将导致法向矢量不连续，容易使有限元计算不能正常进行。

参数曲面和实体造型相结合的方法通常采用 Bezier 曲面、B 样条曲面或 NURBS 曲面等参数曲面来描述模具型腔的复杂曲面，整个模块采用实体造型方法，从而准确全面地描述模具的几何特征。其中，样条曲面中的 B 样条描述方法解决了整体形状控制与局部控制问题，在曲面描述中得到了广泛的应用。采用 B 样条方法对模具型腔曲面进行几何描述，能够提供光滑连续的曲面块和曲面片，为界面节点几何位置的判断和动态接触处理带来了极大的方便。同时，这种描述方法可进一步同模具 CAD 相结合，为工件和模具在成形过程中的同步数值分析奠定了基础。这种方法的优点是容易把塑性成形数值模拟软件与 CAD 软件集成起来，缺点是动态接触处理时要计算直线与复杂曲面的交点以及直线与复杂曲面的最短距离，实现难度较大。

下面具体介绍参数曲面和实体造型相结合的方法来描述模具型腔的过程。

采用控制顶点控制的 B 样条曲线、曲面，具有直观的几何特性、形状易于控制等优点，在曲线和曲面的构造及设计中获得了广泛的应用，B 样条方法已经成为模具型腔曲面几何描述的主流方法。

(1) 三次均匀 B 样条基本理论

设给定 $n+1$ 个空间有序位置点 V_i（$i=0, 1, \cdots, n$），依次用直线段连接这些点形成一个开多边形，称其为控制多边形。点 V_i（$i=0, 1, \cdots, n$）称为此控制多边形的顶点，简称为控制顶点，如图 3.4-3 所示。

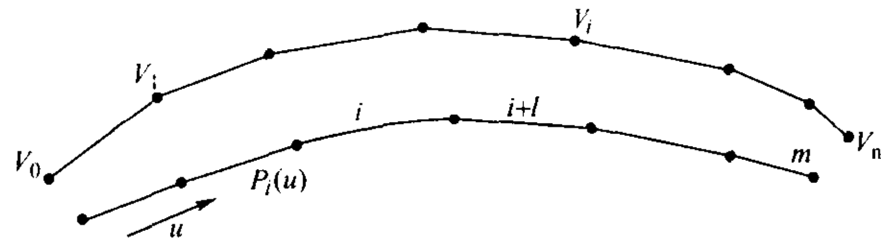


图 3.4-3　B 样条曲线及其控制多边形

现顺次把相邻的 $(k+1)(k \leqslant n)$ 个控制顶点 V_i，V_{i+1}，…，V_{i+k}（$i=0, 1, \cdots, m-1$，$m=n-k+1$）作为一组，做线性组合

$$P_i(u) = \sum_{j=0}^{k} N_{j,k}(u) V_{i+j} \quad (0 \leqslant u \leqslant 1, i = 0,1,\cdots,m-1) \quad (3.4\text{-}70)$$

式中，$N_{j,k}$（u）为 k 次规范 B 样条基函数，其中每一个称为规范 B 样条，其递推定义为

$$\begin{cases} N_{j,0} = 1 \quad 若\ u_j \leqslant u \leqslant u_{j+1} \\ N_{j,0} = 0 \quad 其他 \\ N_{j,k}(u) = \dfrac{u - u_j}{u_{j+k} - u_j} N_{j,k-1}(u) + \dfrac{u_{j+k+1} - u}{u_{j+k+1} - u_{j+1}} N_{j+1,k-1}(u) \\ 规定 \dfrac{0}{0} = 0 \end{cases} \quad (3.4\text{-}71)$$

$N_{j,k}(u)$ 中的 j 为序号，k 为次数。对于固定的 i，式(3.4-70) 为一段 k 次曲线。这样总共有 m 段曲线依次首尾相连，且在连接点处为 C^{k-1} 阶连续，称这 m 段曲线连成的整根曲线为 k 次均匀 B 样条曲线。当 $k=3$ 时为三次均匀 B 样条曲线，三次均匀 B 样条曲线具有 C^2 阶连续性，其方程为

$$P_i(u) = \sum_{j=0}^{3} N_{j,3}(u) V_{i+j} \quad (0 \leqslant u \leqslant 1, i = 0,1,\cdots,n-3) \quad (3.4\text{-}72)$$

用矩阵形式可表示为

$$P_i(u) = UBV \quad (3.4\text{-}73)$$

式中，$U = [1 \quad u \quad u^2 \quad u^3]$

$V = [V_i \quad V_{i+1} \quad V_{i+2} \quad V_{i+3}]$

$$\boldsymbol{B}=\frac{1}{6}\begin{bmatrix}1 & 4 & 1 & 0\\ -3 & 0 & 3 & 0\\ 3 & -6 & 3 & 0\\ -1 & 3 & -3 & 1\end{bmatrix} \tag{3.4-74}$$

(2) 双三次均匀 B 样条曲面

可以在三次均匀 B 样条曲线的基础上构造双三次均匀 B 样条曲面。设给定空间中的十六个位置点 $V_{i,j}$（i，$j=1$，2，3，4），并按顺序排成一个四阶方阵 $\boldsymbol{V}$

$$\boldsymbol{V}=\begin{bmatrix}V_{1,1} & V_{1,2} & V_{1,3} & V_{1,4}\\ V_{2,1} & V_{2,2} & V_{2,3} & V_{2,4}\\ V_{3,1} & V_{3,2} & V_{3,3} & V_{3,4}\\ V_{4,1} & V_{4,2} & V_{4,3} & V_{4,4}\end{bmatrix} \tag{3.4-75}$$

把四阶方阵中的每一列看作是一个特征多边形的四个顶点，按式（3.4-72）可以构造四条三次均匀 B 样条曲线

$$[P_1(u)\quad P_2(u)\quad P_3(u)\quad P_4(u)]=$$
$$[N_{0,3}(u)\quad N_{1,3}(u)\quad N_{2,3}(u)\quad N_{3,3}(u)]$$
$$\begin{bmatrix}V_{1,j}\\ V_{2,j}\\ V_{3,j}\\ V_{4,j}\end{bmatrix}\quad (0\leqslant u\leqslant 1, j=1,2,3,4) \tag{3.4-76}$$

对于 $[0,\ 1]$ 之间的每个 u 值，再把 $P_1(u)$，$P_2(u)$，$P_3(u)$，$P_4(u)$看作是一个特征多边形的四个顶点，可以构造关于参数 w 的三次均匀 B 样条曲线

$$P(u,w)=[N_{0,3}(w)\quad N_{1,3}(w)\quad N_{2,3}(w)\quad N_{3,3}(w)]\begin{bmatrix}P_1(u)\\ P_2(u)\\ P_3(u)\\ P_4(u)\end{bmatrix} \tag{3.4-77}$$

当参数 u，w 在区间 $[0,\ 1]$ 上相互独立地变化时，式（3.4-77）即为一双三次均匀 B 样条曲面片的方程，用矩阵形式可表示为

$$P(u,w)=\boldsymbol{UBVB}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}} \tag{3.4-78}$$

式中，$\boldsymbol{W}=[1\quad w\quad w^2\quad w^3]$，$\boldsymbol{V}$ 称为 B 样条曲面片的控制网格阵，$V_{i,j}$（i，$j=1$，2，3，4）称为控制网格顶点，简称控制顶点。双三次均匀 B 样条曲面片由这十六个控制顶点惟一确定，如图 3.4-4 所示。

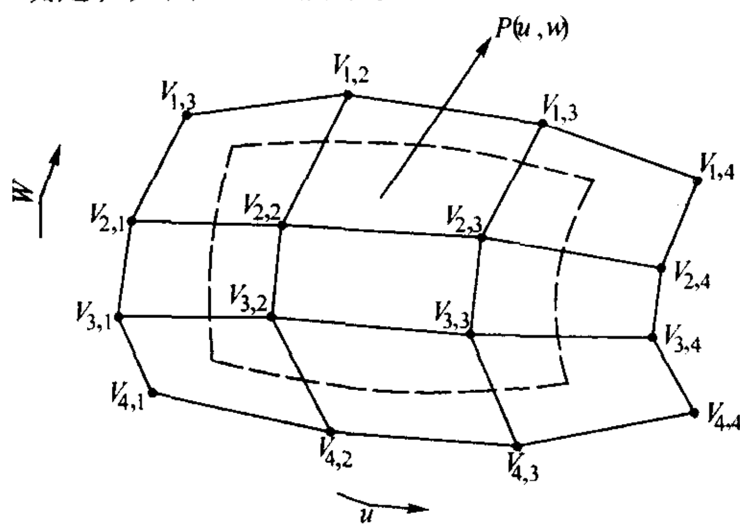


图 3.4-4　双三次均匀 B 样条曲面及其控制网格

实际空间曲面是由若干个曲面片连接而成的，称作曲面块。设在三维空间内给定（$n+1$）×（$m+1$）个位置点，它们排成一个（$n+1$）×（$m+1$）阶矩阵 $V_{i,j}$（$i=0$，1，⋯，n；$j=0$，1，⋯，m）

$$\begin{bmatrix}V_{0,0} & V_{0,1} & V_{0,2} & \cdots & V_{0,m}\\ V_{1,0} & V_{1,1} & V_{1,2} & \cdots & V_{1,m}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ V_{n,0} & V_{n,1} & V_{n,2} & \cdots & V_{n,m}\end{bmatrix} \tag{3.4-79}$$

构成一控制网格，相应的双三次 B 样条曲面方程为

$$P(u,w)=P_{i,j}(u,w)=\boldsymbol{UBVB}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}$$
$$(0\leqslant u,\ w\leqslant 1;\ i=0,\ 1,\ \cdots,\ n-3;\ j=0,\ 1,\ \cdots,\ m-3) \tag{3.4-80}$$

它是由（$n-2$）×（$m-2$）个双三次 B 样条曲面片连接而成的，每两个相邻面片之间都具有 C^2 阶连续性。

对确定的双三次 B 样条曲面，其法矢量 $\boldsymbol{N}$ 可由对参数 u 和 w 的偏导矢确定

$$\frac{\partial \boldsymbol{P}(u,w)}{\partial u}=\boldsymbol{U'BVB}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}} \tag{3.4-81}$$

$$\frac{\partial \boldsymbol{P}(u,w)}{\partial w}=\boldsymbol{UBVB}^{\mathrm{T}}\boldsymbol{W'}^{\mathrm{T}} \tag{3.4-82}$$

$$\boldsymbol{N}=\frac{\partial \boldsymbol{P}(u,w)}{\partial u}\times\frac{\partial \boldsymbol{P}(u,w)}{\partial w} \tag{3.4-83}$$

实际构造曲线曲面时，一般有两种情况：一是已知控制顶点构造曲线曲面，称其为正算过程；另一种是已知曲线或曲面上的点（型值点）构造曲线或曲面，称其为反算过程。

(3) 双三次 B 样条曲面控制网格顶点的反算

在工程实践中，往往已知曲面上的一些型值点。因此，必须首先根据型值点来反算出 B 样条曲面控制网格的顶点，然后根据这些控制顶点来构造 B 样条曲面。可采用双向 B 样条曲线反算法来求得双三次 B 样条曲面的控制顶点列 $V_{i,j}$（$i=0$，1，⋯，n；$j=0$，1，⋯，m），其步骤如下。

首先对 u 向的 $n+1$ 组型值点，按三次 B 样条曲线的反算方法，得到各条插值曲线的特征多边形顶点 $P_{i,j}$（$i=0$，1，⋯，$n+2$；$j=0$，1，⋯，m）。然后，将 $P_{i,j}$看作 w 方向上的 $m+1$ 组型值点列，再按三次 B 样条曲线的反算法得到双三次 B 样条曲面的特征网格顶点 $V_{i,j}$（$i=0$，1，⋯$n+2$；$j=0$，1，⋯，$m+2$）。

对双三次 B 样条曲面，在反算过程中要在每条曲线上增加两个边界条件，主要有切矢边界条件和自由端点边界条件两种。

1.5.2　动态边界条件的处理

金属塑性成形过程伴随着变形工件与模具接触的动态变化过程。在变形的任一瞬时，工件外表面总是可以分为自由表面和接触表面两部分，随着变形过程的进行，由于与模具的相对运动，自由表面上的质点会与模具表面接触，而接触表面也有可能脱开模具表面。对于自由表面，其边界条件十分简单，即为给定载荷为零的特殊力面；而接触表面上是给定了外力或位移速度的混合边界条件，即切向分力（摩擦力）和法向分速度给定。动态接触边界条件的处理，主要是对接触问题的几何机制——触模与脱模和速度约束的处理。

(1) 自由节点触模判断

设第 m 步增量变形区间为 $[t_{m-1},\ t_m]$，采用牛顿－拉夫森法解出该区间收敛的速度场后，各自由节点就以该速度场的速度在空间运动，它们与模具是否接触是由本身与模具的相对运动性质所决定的，而接触条件可以归结为任一自由节点与模具接触时间的确定。

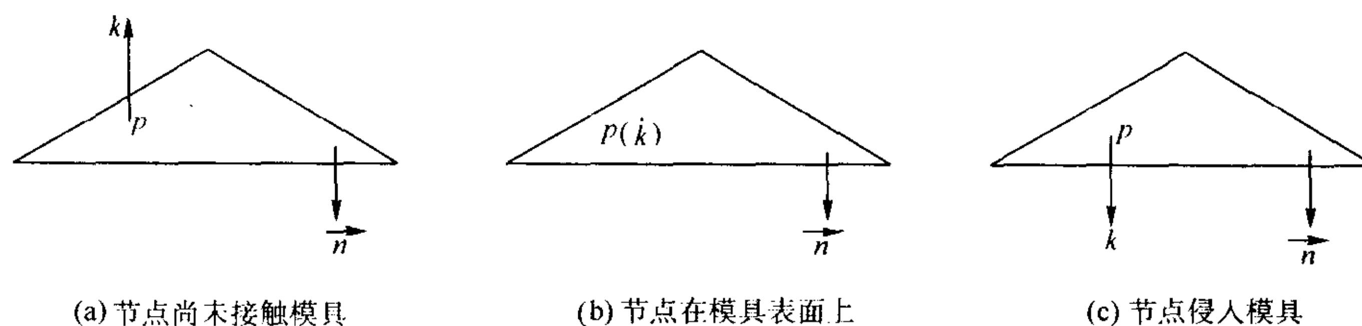
模具型腔表面可以根据需要以三角形面片来描述。对每一个三角形面片，已知的是其三个顶点的坐标及其法向矢量 $\boldsymbol{n}$。对变形体的任一个边界节点 k，均可以向描述模具型腔曲面的小三角形面片所在的平面作垂线，进而求得垂足 p。然后判断垂足 p 是否在三角形面片之内，如果在，则连接 p 点和 k 点形成一矢量 $\boldsymbol{pk}$，其方向由 p 指向 k。利用矢量 $\boldsymbol{pk}$ 与 $\boldsymbol{n}$ 的相互关系可以准确地判断出节点 k 与模具型腔曲面之间的相互位置关系，即

1）$\boldsymbol{pk}\cdot\boldsymbol{n}<0$　　节点 k 尚未接触模具；

2）$\boldsymbol{pk}\cdot\boldsymbol{n}=0$　　节点 k 处在模具表面上；

3）$\boldsymbol{pk}\cdot\boldsymbol{n}>0$　　节点 k 侵入模具。

其位置关系如图 3.4-5 所示。



(a) 节点尚未接触模具 (b) 节点在模具表面上 (c) 节点侵入模具

图3.4-5 边界节点与模具的三角形面片间的位置关系

上述即为边界节点与模具型腔曲面之间相互位置关系的动态识别过程。这个过程要在每一次加载修正后对每个边界节点都进行一次，以此来判断边界节点的状态，修正其相应的边界条件。对那些侵入模具的节点，应对其位置进行修正，将其"拉"回到模具表面上。

(2) 接触节点脱模检验与位置修正

1) 接触节点脱模检验 对于接触节点脱离模具，常用接触力检验法检验，即检查接触节点的法向是否受力，若为拉力时，认为原节点脱离模具。

接触节点脱离模具的接触判据为

$$F_n > 0 \text{ 或 } \sigma_n > 0 \quad (3.4\text{-}84)$$

式中，F_n、σ_n 为接触节点的法向力和应力。

若接触节点满足式（3.4-84），则该节点在下一增量计算时解除接触约束，变为自由节点。

2) 接触节点位置修正 前已述及，采用增量计算会造成接触节点偏离模具表面。这主要是由于计算时接触节点的运动与实际变形不完全相同而造成的，实际变形时接触点是贴合在模具表面上并沿其形状变化滑动，而有限元计算时接触节点在任一增量区间速度近似认为不变。这种处理误差引起的偏差有以下两种情况，即接触节点穿透模具表面嵌入到模具内或离开模具表面与之分离。

由于模具型腔表面的分片描述，因而首先要确定接触节点在空间上所对应的曲面片，即接触节点与曲面的相关性判断，然后再进行位置修正。

对于任一接触节点 Q，其相关性判断的主要条件是

$$\Gamma_j(x_Q, y_Q, z_Q) \leqslant \varepsilon \quad (3.4\text{-}85)$$

且 Q 点还必须落在曲面片 Γ_j 的界限之内。

位置修正的原则是将偏离表面的接触节点，沿最短距离拉回到模具表面上。

(3) 速度约束条件的处理

整体刚度矩阵 $\boldsymbol{K}$ 集成后是一个奇异矩阵，即 $|\boldsymbol{K}| = 0$，则对应的逆矩阵 $\boldsymbol{K}^{-1}$ 不存在，此时整体刚度方程无法求解。它反映了刚度方程所定义的解中含有任意刚体运动，必须引入速度约束条件予以消除。

速度约束条件就是变形体 S_v 表面上的速度边界条件。显然 S_v 表面就是工件与模具的接触表面 S_f，这种接触状态的边界条件表现为该表面上任一点的法向运动速度给定。经离散化后 S_v 上的速度边界条件只对该表面上的节点产生约束作用，而其他点则由节点值插值自动满足。

对于 S_v 上整体节点号码为 I 的接触节点，速度约束条件表示为

$$(\dot{u}_n)_I = \boldsymbol{v}_d^T \boldsymbol{n}_I \quad (3.4\text{-}86)$$

式中，$(\dot{u}_n)_I$ 为节点 I 的法向速度分量；$\boldsymbol{n}_I$ 为节点 I 处模具表面的单位法矢量；$\boldsymbol{v}_d$ 为模具速度矢量。

由于速度场采用牛顿－拉夫森迭代法求解，则式（3.4-86）变为

$$(\Delta \dot{u}_n)_I = 0 \quad (3.4\text{-}87)$$

若接触节点 I 处所定义的局部切法向坐标系与整体坐标系重合，该速度约束条件可以直接引入到整体刚度方程中。但通常局部切法向坐标系与整体坐标系不重合，习惯上称作斜约束，此时要先对刚度方程进行旋转变换，然后才能处理约束条件。

1.5.3 网格重新划分

金属塑性成形过程中，金属材料在模具的作用下经过连续变形，从较为简单形状的毛坯变为形状复杂的工件，其间往往伴随着大变形和大位移。用刚（黏）塑性有限元模拟这种成形过程时，随着金属材料不断塑性流动，变形体内各节点的相对位置在不断改变，原本规则的有限元网格也要发生变形，某些局部区域的网格甚至严重畸变。因此，要完成变形过程模拟计算，必须进行有限元网格的重新划分，把旧网格的信息传递到新网格上。在有限元模拟中，要求单元既要有一定的抗畸变能力，又要有一定的变形能力，同时还要有较高的计算精度。

(1) 有限元网格的划分

根据网格中节点和单元的生成时序，可以将网格生成方法分成七大类：网格均匀化法、拓扑分解法、节点连元法、基于栅格法、映射单元法、保角映射法和几何分解法。

网格均匀化法首先确定网格的拓扑结构，然后确定节点和单元。这种方法往往依赖其他方法来提供网格拓扑。同时，经常在其他方法中应用均匀化技术来提高网格的质量。应用最普遍的是拉普拉斯均匀化方法，这种方法将每个内部节点调整到相邻节点组成的多边形形心处，通常通过迭代方法来完成。

拓扑分解法假设物体的顶点为初始网格的唯一节点，则可以将复杂的物体分解成三角形集合的简单拓扑。首先生成粗网格，如图3.4-6a。然后进行网格细化，生成如图3.4-6b所示的细网格。

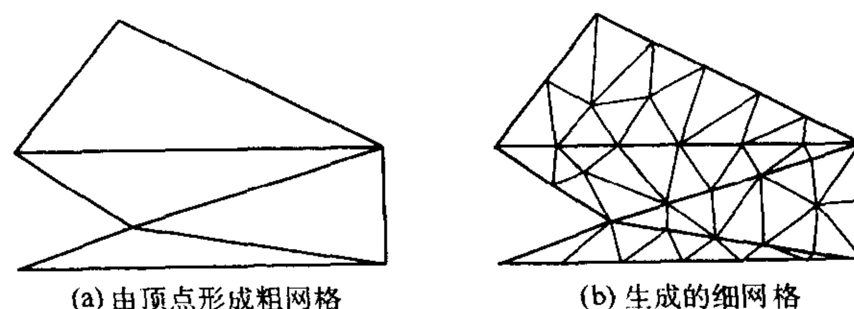

(a) 由顶点形成粗网格 (b) 生成的细网格

图3.4-6 拓扑分解法

节点连元法首先根据指定的网格密度生成区域内的所有节点，然后连接这些节点以生成形状尽可能好的单元，如图3.4-7所示。包含两个阶段：节点生成和单元生成。

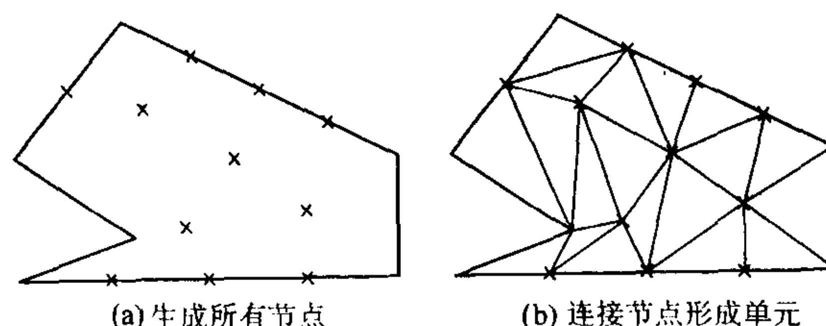

(a) 生成所有节点 (b) 连接节点形成单元

图3.4-7 节点连元法

映射单元法首先生成参数空间内的四边形或三角形栅格的网格模板，然后通过混合函数将网格模板映射到四边或三

边子区域上，在子区域内形成网格。这种方法首先需要将复杂的物体分解成四边或三边区域，如图3.4-8所示。

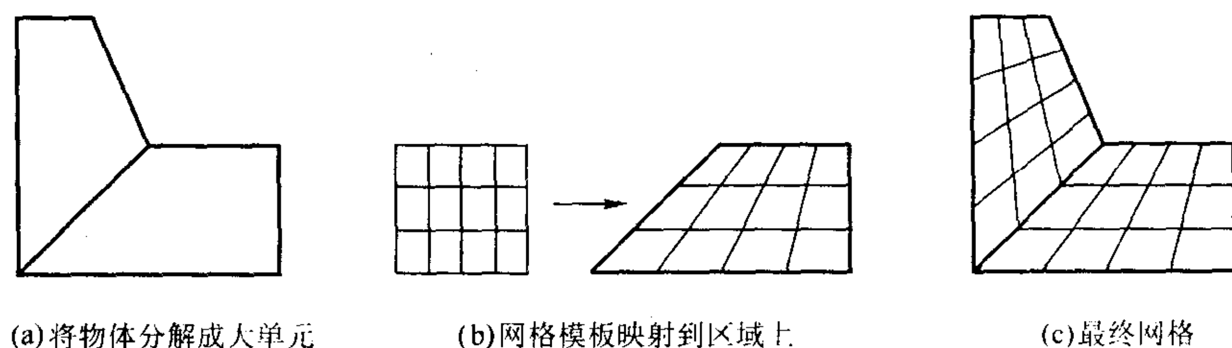


图3.4-8　映射单元法

基于栅格法首先生成栅格，然后将栅格加在物体上并删除那些落在物体之外的栅格单元，最后在区域内生成不同类型的网格。

保角映射法首先在参数空间内构造一个与物体 P 顶点数目相同的多边形 Q，然后在 Q 内生成网格模板，根据多边形 P 和 Q 的顶点之间的对应关系，建立从 Q 到 P 的保角映射函数 F，将多边形 Q 内的网格映射到物体 P 上。

几何分解法同时生成网格的节点和单元，首先将子域从物体中移出，然后再对子域进行细分形成不同类型的单元。不像拓扑分解法那样忽略几何因素，而是在分解物体时考虑了单元的形状和尺寸，试图来生成形状好的单元。

(2) 有限元网格的重划分

金属塑性成形过程中，当网格变形到一定程度时，就必须进行相应的处理。如果只在工件表层出现个别畸变单元，则可以采用局部调整的方法加以解决。如果大量的畸变网格出现在工件的内部，则局部调整已无法满足要求，此时应停止计算，在旧网格的基础上重新生成一套适合于计算的新网格。

网格重新划分技术是实现塑性成形过程有限元模拟关键技术之一。它包括网格畸变判别、重新划分网格和新旧网格之间的数据转换等三个主要步骤，其中重新划分网格是该技术的重点和难点。

1) 网格重划的判断准则　在进行网格重划时，首先要建立判断准则以确定何时进行网格重划，下面介绍采用工件与模具的干涉量和雅可比矩阵行列式的值来判断是否需要进行重划。

六面体单元的边界是线性边界，当单元边界无法很好地贴近模具表面时会发生干涉现象。随着干涉量的增加，有限元计算的精度会逐渐降低。因此，在确定网格性能是否合格之前，必须先进行工件与模具的边界干涉判断，其判断准则如下。

如图3.4-9，设 P 点是单元干涉面的中心点，Q 点是与之相对面的中心点，R 点是 PQ 连线与模具表面的交点，h_i 是 P 点与 R 点之间的距离，h_e 是 P 点与 Q 点之间的距离。则干涉判断准则为：

$$h_i / h_e \geq C_i \tag{3.4-88}$$

式中，C_i 为干涉判据常数，常取为0.2~0.3。

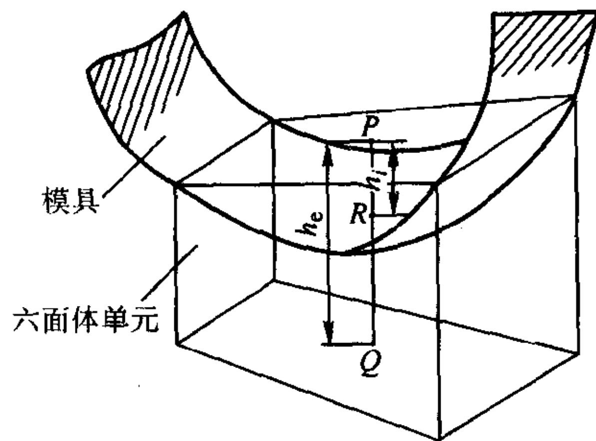


图3.4-9　边界干涉示意图

为保证有限元求解过程顺利进行，用于坐标变换和积分变换的单元 Jacobian 矩阵的行列式 $|J|$ 的值须大于零，即等参元是外凸的。$|J|$ 的表达式为：

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} & \frac{\partial z}{\partial \eta} \\ \frac{\partial x}{\partial \zeta} & \frac{\partial y}{\partial \zeta} & \frac{\partial z}{\partial \zeta} \end{vmatrix} \tag{3.4-89}$$

对于八节点六面体单元，若单元的任意内角均满足 $0 < \theta < \pi$，则其 $|J|$ 必大于零。因此，基于 Jacobian 矩阵的行列式的值的网格畸变准则可叙述如下。

设 P_1 点为单元的某一节点，P_2、P_3 和 P_4 点是与 P_1 点相邻的节点，$\boldsymbol{\lambda}_2$、$\boldsymbol{\lambda}_3$ 和 $\boldsymbol{\lambda}_4$ 分别为从 P_1 点到 P_2、P_3 和 P_4 点的单位矢量，如图3.4-10所示。如果由单位矢量 $\boldsymbol{\lambda}_2$、$\boldsymbol{\lambda}_3$ 和 $\boldsymbol{\lambda}_4$ 构成的平行六面体的体积为负值，则网格发生畸变，即：

$$\begin{cases} (\boldsymbol{\lambda}_2 \times \boldsymbol{\lambda}_3) \cdot \boldsymbol{\lambda}_4 \leq C_d \\ (\boldsymbol{\lambda}_3 \times \boldsymbol{\lambda}_4) \cdot \boldsymbol{\lambda}_2 \leq C_d \\ (\boldsymbol{\lambda}_4 \times \boldsymbol{\lambda}_2) \cdot \boldsymbol{\lambda}_3 \leq C_d \end{cases} \tag{3.4-90}$$

式中，C_d 为网格畸变判据常数，实际应用一般可取为0.01。

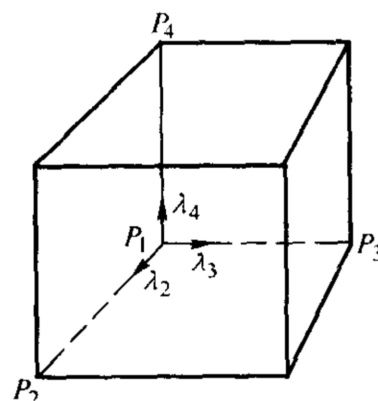


图3.4-10　六面体网格畸变示意图

当上述两个网格重划判断准则的一个满足条件时，就应该停止计算，重新生成有限元网格。

2) 新网格的自动生成　这是网格重划最重要的一步，需要在旧的有限元网格边界包围的空间区域内生成新的网格。首先，根据旧网格的节点和单元信息重构边界曲面。然后，根据网格边界的形状和材料的流动特性，采用中轴面分解技术将变形体分解为若干个子域。最后，在每个子域内生成新的六面体网格。

由于重划后的网格边界与旧网格边界存在一定的误差，因此需要对新网格的边界进行修正，以提高变形体在边界处的拟合精度，可采用投影修正法将自动生成的新网格的边界节点投影到变形体的旧边界上，这样可以使变形体边界在重划前后基本上保持一致。

3) 新节点的包含测试及局部坐标的计算　为正确实现新旧网格系统间的场变量信息传递，必须先确定新网格节点落在哪个旧单元内，这称为包含测试。本文采用体积校验法进行包含测试，这种方法不涉及正负号和方向问题，计算较为方便并且行之有效。

设 V_0 为旧六面体单元的体积，V_i 为新节点 P 与旧单元边界面的四个节点连线后形成的五面体的体积，如果 P 点

在旧单元内，则有：

$$\sum_{i=1}^{6} V_i \Big/ V_0 = 1 \qquad (3.4\text{-}91)$$

如果 P 点在旧单元外，则有：

$$\sum_{i=1}^{6} V_i \Big/ V_0 > 1 \qquad (3.4\text{-}92)$$

新网格节点 P 在旧网格系统内的自然坐标（ξ，η，ζ）可由下式求出：

$$\begin{cases} \sum_{i=1}^{8} N_i(\xi,\eta,\zeta)x_i = x \\ \sum_{i=1}^{8} N_i(\xi,\eta,\zeta)y_i = y \\ \sum_{i=1}^{8} N_i(\xi,\eta,\zeta)z_i = z \end{cases} \qquad (3.4\text{-}93)$$

式中，（x_i，y_i，z_i）为旧网格节点的全局坐标，（x，y，z）为新网格节点 P 的全局坐标，$N_i(\xi,\eta,\zeta)$为形状函数，其表达式如下：

$$N_i(\xi,\eta,\zeta) = \frac{1}{8}(1+\xi_i\xi)(1+\eta_i\eta)(1+\zeta_i\zeta) \qquad (3.4\text{-}94)$$

式中，（ξ_i，η_i，ζ_i）为旧网格节点 i 的自然坐标。

4）新旧网格之间场变量信息传递　新的有限元网格生成后，须把旧网格上的有关信息（包括与变形历史有关的等效应变、速度场及边界条件等）传递到新网格上，以保证计算的连续性。

等效应变的传递可分为两步：首先将旧网格单元的等效应变值插值到旧网格节点上；然后再将旧网格节点的等效应变值传递到新网格节点上，继而计算新网格单元的等效应变。

由 n 个单元包围的旧网格节点 j 的等效应变值 $\bar{\varepsilon}_j$ 可以体积加权平均得到：

$$\bar{\varepsilon}_j = \left(\sum_{i=1}^{n} \bar{\varepsilon}_i V_i\right) \Big/ \sum_{i=1}^{n} V_i \qquad (3.4\text{-}95)$$

式中，V_i 和 $\bar{\varepsilon}_i$ 分别为旧网格单元 i 的体积和等效应变值。

首先进行包含测试，然后可插值求出新网格节点 k 的等效应变值 $\bar{\varepsilon}_k$：

$$\bar{\varepsilon}_k = \sum_{j=1}^{8} N_j \bar{\varepsilon}_j \qquad (3.4\text{-}96)$$

新网格单元内任一点 l 的等效应变值 $\bar{\varepsilon}_l$ 可通过下式求出：

$$\bar{\varepsilon}_l(\xi,\eta,\zeta) = \sum_{k=1}^{8} N_k(\xi,\eta,\zeta)\bar{\varepsilon}_k \qquad (3.4\text{-}97)$$

式中，（ξ，η，ζ）为点 l 的自然坐标值。

单元的等效应变值可由式（3.4-94）求出高斯积分点的等效应变值后平均得到。

由于变形体的旧网格节点的速度已知，新网格节点的速度场可以直接从旧网格中插值获得。

体积成形过程有限元模拟的关键技术还包括对刚性区的处理、初始速度场生成、摩擦条件的处理等。

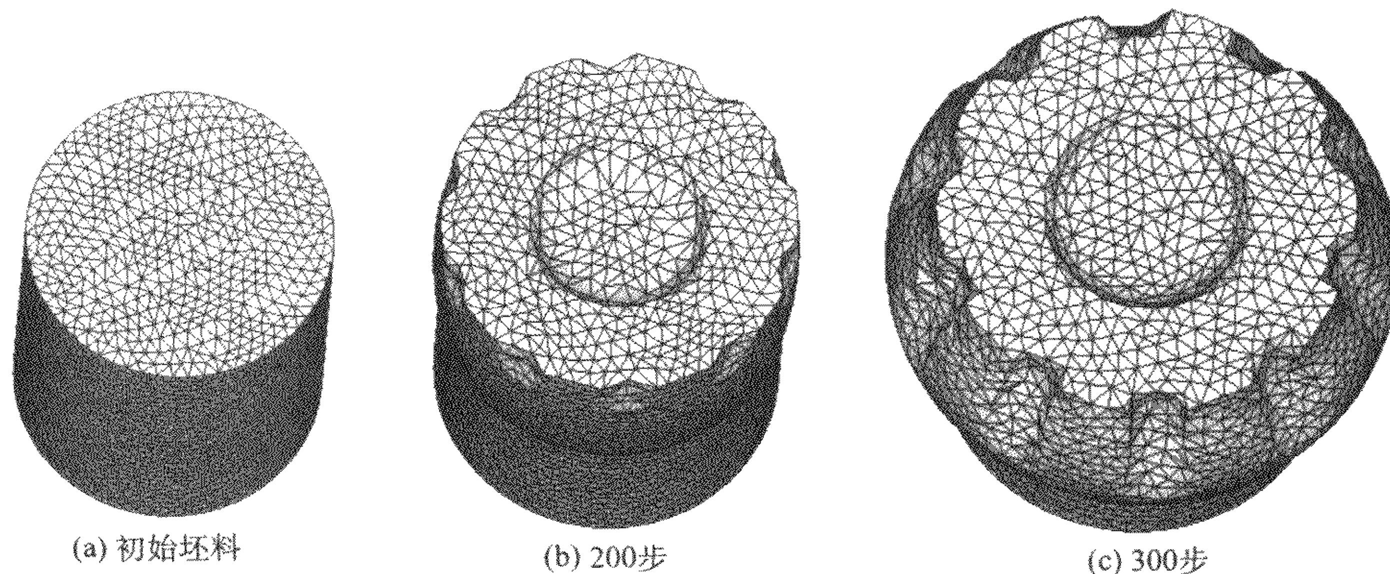
由于刚（黏）塑性有限元建立于刚（黏）塑性变分原理之上，而刚（黏）塑性变分原理只适用于塑性变形区，因此，在刚性区内应变速率接近零或等于零，在计算过程中会引起泛函变分的奇异，造成计算结果的溢出，因此有必要区分塑性区与刚性区。但在计算开始时，很难准确地确定塑性变形区与刚性区的交界面。为了解决这个问题，常采取简化的处理办法。将塑性区域假定为一个较大的范围，取得计算结果后，引入等效应变速率截断值 $\dot{\varepsilon}_0$，当某单元的等效应变速率满足 $\dot{\varepsilon} \geqslant \dot{\varepsilon}_0$ 时为塑性区，否则为刚性区。对于刚性区单元取 $\dot{\varepsilon} = \dot{\varepsilon}_0$，并认为该区域的应力与应变关系呈线性。

对于初始速度场的生成，可以设定一个初始速度场作为迭代的初始点，并从该初始点开始进行反复迭代直至收敛于真实解。当初始速度要接近于真实解时，可以取得较快的收敛速度。初始速度场的确定非常重要，当然初始速度场并不要十分精确，但必须满足边界条件，并且基本符合此时材料变形的流动规律。初始速度场的生成有两种较为通用的方法，即近似泛函法和线性本构关系法。近似泛函法生成初始速度场的具体思路是：从广义变分原理的几种泛函出发，构造一个与之近似的泛函 G，由其驻值条件得到的有限元方程是一个线性方程组，然后解此方程组求出满足边界条件的速度场，作为初始速度场。线性本构关系法实际上也是一种近似泛函法。它与前述近似泛函的不同之处在于，它是基于塑性本构关系的近似线性化假设构造泛函。这种近似泛函得到的有限元方程必然是一个线性方程组，将其解答用作初始速度场。

有限元模拟过程中，如何正确处理摩擦边界条件也是非常重要的。工件变形时在与模具接触面上受到的摩擦力作用，对金属材料流动模式、工件几何尺寸、内部缺陷、模具受力状态以及总载荷和总能量都有很大的影响作用。然而，塑性加工中的摩擦是在高压、高温条件下发生的，并且随着工件的塑性变形而变化，所以其机制十分复杂，影响因素也很多，如接触面上的润滑条件、模具表面状态、变形温度以及材料化学成分及性能等等。因此正确处理摩擦边界条件，选择合理的摩擦模型将直接影响到有限元计算结果的准确性。摩擦力的计算模型通常采用库仑模型和剪切模型，摩擦力作为表面力引入泛函中。

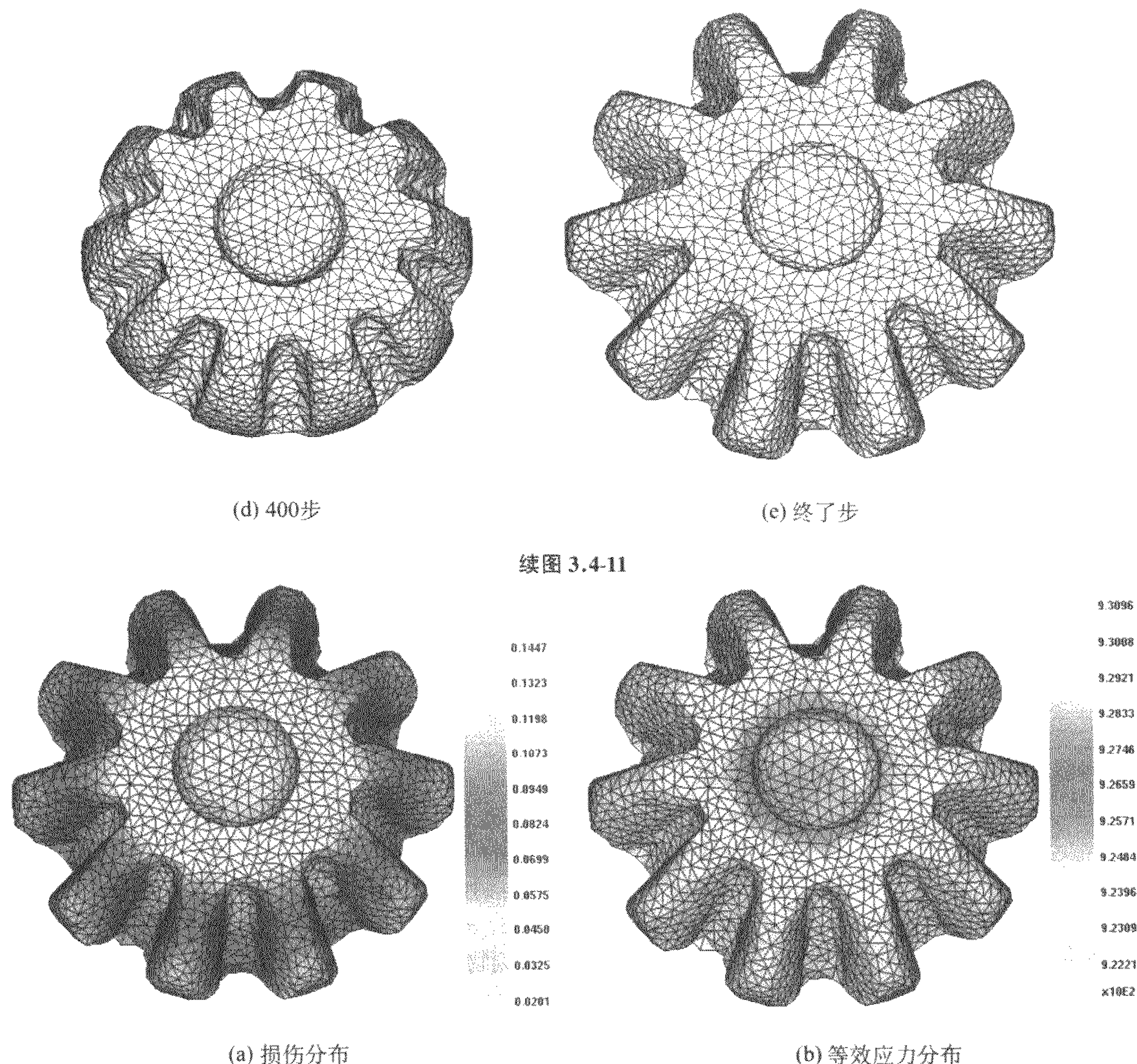
1.6 计算实例

以轿车上的关键零件之一的圆锥齿轮冷精锻过程为例，对其变形过程进行有限元数值分析。刚塑性材料的参数为：弹性模量 210 000 MPa，密度 7.8 g/cm^3，泊松比 0.29，屈服强度 840 MPa。成形时冲头的挤压速率为 5 mm/s。齿轮变形过程的金属流动变化如图 3.4-11 所示。成形后工件的损伤分布和等效应力分布见图 3.4-12 所示。



(a) 初始坯料　(b) 200步　(c) 300步

图 3.4-11　齿轮精锻过程金属流动变化图



(d) 400步　　(e) 终了步

续图 3.4-11

(a) 损伤分布　　(b) 等效应力分布

图 3.4-12　工件的损伤和等效应力分布图

2　弹塑性有限元法

1967 年，Marcal 和 King 首先提出了弹塑性有限元法。1968 年，Yamada 等推导了塑性应力 - 应变矩阵。1970 年，Hibbit 等提出了建立在有限变形理论基础上的大变形有限元列式。20 世纪 70 年代中期，Osias、McMeeking 等采用欧拉描述法建立了大变形有限元列式。此后大变形弹塑性有限元法不断完善。采用弹塑性有限元法分析金属成形问题，不仅能计算工件的变形和应力、应变分布，而且还能有效地处理卸载问题，计算金属成形过程结束后工件的回弹和残余应力、残余应变的分布。因此，它适宜于板料成形等问题的模拟。但是，弹塑性有限元法采用的增量型本构关系不允许使用大的变形增量，总的计算时间较长。

弹塑性变形分析与线弹性变形分析的基本差别在于前者的应力应变关系是非线性的、依赖于变形历史的。因此，弹塑性变形分析一般采用增量分析方法，即将加载过程分为若干个增量步，在每个增量步中使物体所受的载荷或给定的边界位移产生增量。在一小段增量的范围内对应力应变关系进行线性化处理，通过增量形式的有限元方程求得增量步中的位移增量，然后求得应变增量和应力增量，并逐步累加。

2.1　小变形弹塑性有限元法

对于小位移小变形，可采用小变形弹塑性理论对其变形进行有限元分析。对金属塑性成形过程的每一个小变形增量应用这种方法进行变形分析，通过一系列增量分析来实现对整个成形过程的分析。每一增量步中产生的应变称为应变增量，应变增量张量为小应变张量，而应力增量张量就是柯西应力增量张量。其本构关系要写成增量形式，并引入塑性流动法则和加载 - 卸载准则，虚功方程也要写成增量的形式，由此得到的有限元方程是非线性方程组，要采用迭代法求解。

2.1.1　应力应变矩阵

增量形式的小应变弹塑性本构方程的一般张量形式和矩阵形式分别为

$$\Delta\sigma_{ij} = C^{ep}_{ijkl}\Delta\varepsilon_{kl} = (C^{e}_{ijkl} - \alpha C^{p}_{ijkl})\Delta\varepsilon_{kl} \tag{3.4-98}$$

$$\Delta\boldsymbol{\sigma} = \boldsymbol{C}^{ep}\Delta\boldsymbol{\varepsilon} = (\boldsymbol{C}^{e} - \alpha\boldsymbol{C}^{p})\Delta\boldsymbol{\varepsilon} \tag{3.4-99}$$

其中 $\boldsymbol{C}^{ep}$ 称为弹塑性应力应变矩阵，简称弹塑性矩阵。小应变弹塑性有限元分析中上式各项的具体表达式阐述如下。

(1) 三维变形问题的弹塑性矩阵

对于三维变形问题，式 (3.4-99) 中各个矢量和矩阵分别是：

$$\Delta\boldsymbol{\sigma} = [\Delta\sigma_{11} \quad \Delta\sigma_{22} \quad \Delta\sigma_{33} \quad \Delta\sigma_{12} \quad \Delta\sigma_{23} \quad \Delta\sigma_{31}]^{T} \tag{3.4-100}$$

$$\begin{aligned}\Delta\boldsymbol{\varepsilon} &= [\Delta\varepsilon_{11} \quad \Delta\varepsilon_{22} \quad \Delta\varepsilon_{33} \quad \Delta\varepsilon_{12} \quad \Delta\varepsilon_{23} \quad 2\Delta\varepsilon_{31}]^{T} \\ &= \left[\frac{\partial(\Delta u_1)}{\partial x_1}\frac{\partial(\Delta u_2)}{\partial x_2}\frac{\partial(\Delta u_3)}{\partial x_3}\frac{\partial(\Delta u_1)}{\partial x_2}+\frac{\partial(\Delta u_2)}{\partial x_1}\times\right. \\ &\left.\frac{\partial(\Delta u_2)}{\partial x_3}+\frac{\partial(\Delta u_3)}{\partial x_2}\frac{\partial(\Delta u_3)}{\partial x_1}+\frac{\partial(\Delta u_1)}{\partial x_3}\right]^{T}\end{aligned} \tag{3.4-101}$$

$\boldsymbol{C}^{e}$ 和 $\boldsymbol{C}^{p}$ 分别为

$$C^{e}=\frac{E}{1+\nu}\begin{bmatrix}\frac{1-\nu}{1-2\nu} & & & & & \text{对称}\\ \frac{\nu}{1-2\nu} & \frac{1-\nu}{1-2\nu} & & & & \\ \frac{1-\nu}{1-2\nu} & \frac{\nu}{1-2\nu} & \frac{1-\nu}{1-2\nu} & & & \\ 0 & 0 & 0 & \frac{1}{2} & & \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}\end{bmatrix} \quad (3.4\text{-}102)$$

$$C^{p}=\frac{9G^2}{(E_t+3G)\bar{\sigma}^2}\times\begin{bmatrix}\sigma'_{11}\sigma'_{11} & & & & & \text{对称}\\ \sigma'_{11}\sigma'_{22} & \sigma'_{22}\sigma'_{22} & & & & \\ \sigma'_{11}\sigma'_{33} & \sigma'_{22}\sigma'_{33} & \sigma'_{33}\sigma'_{33} & & & \\ \sigma'_{11}\sigma_{12} & \sigma'_{22}\sigma_{12} & \sigma'_{33}\sigma_{12} & \sigma_{12}\sigma_{12} & & \\ \sigma'_{11}\sigma_{23} & \sigma'_{22}\sigma_{23} & \sigma'_{33}\sigma_{23} & \sigma_{12}\sigma_{23} & \sigma_{23}\sigma_{23} & \\ \sigma'_{11}\sigma_{31} & \sigma'_{22}\sigma_{31} & \sigma'_{33}\sigma_{31} & \sigma_{12}\sigma_{31} & \sigma_{23}\sigma_{31} & \sigma_{31}\sigma_{31}\end{bmatrix} \quad (3.4\text{-}103)$$

(2) 平面应变问题和轴对称问题

平面应变问题和轴对称问题有一些共同的特点，可以将这种分析用同一个程序来实现。对于平面应变问题

$$\Delta\boldsymbol{\sigma}=[\Delta\sigma_{11}\quad\Delta\sigma_{22}\quad\Delta\sigma_{33}\quad\Delta\sigma_{12}]^{T} \quad (3.4\text{-}104)$$

$$\Delta\boldsymbol{\varepsilon}=[\Delta\varepsilon_{11}\quad\Delta\varepsilon_{22}\quad 0\quad 2\Delta\varepsilon_{12}]^{T} \quad (3.4\text{-}105)$$

对于轴对称问题

$$\Delta\boldsymbol{\sigma}=[\Delta\sigma_{r}\quad\Delta\sigma_{z}\quad\Delta\sigma_{\theta}\quad\Delta\sigma_{zr}]^{T} \quad (3.4\text{-}106)$$

$$\Delta\boldsymbol{\varepsilon}=[\Delta\varepsilon_{r}\quad\Delta\varepsilon_{z}\quad\Delta\varepsilon_{\theta}\quad 2\Delta\varepsilon_{zr}]^{T}=\left[\frac{\partial(\Delta u_r)}{\partial r},\frac{\partial(\Delta u_z)}{\partial z},\Delta\frac{u_r}{r},\frac{\partial(\Delta u_z)}{\partial r}+\frac{\partial(\Delta u_r)}{\partial z}\right]^{T} \quad (3.4\text{-}107)$$

这两类问题的应力应变矩阵 C^{ep} 即为三维变形问题应力应变矩阵的四阶主子式，只是轴对称问题中，要用下标 r，z，θ 分别代替三维问题中的下标 1，2，3。

(3) 平面应力问题

对于平面应力问题，式（3.4-99）中的各个矢量和矩阵分别是：

$$\Delta\boldsymbol{\sigma}=[\Delta\sigma_{11}\quad\Delta\sigma_{22}\quad\Delta\sigma_{12}]^{T} \quad (3.4\text{-}108)$$

$$\Delta\boldsymbol{\varepsilon}=[\Delta\varepsilon_{11}\quad\Delta\varepsilon_{22}\quad 2\Delta\varepsilon_{12}]^{T} \quad (3.4\text{-}109)$$

C^{e} 和 C^{p} 分别为

$$C^{e}=\frac{E}{1-\nu^2}\begin{bmatrix}1 & & \text{对称}\\ \nu & 1 & \\ 0 & 0 & \frac{1-\nu}{2}\end{bmatrix} \quad (3.4\text{-}110)$$

$$C^{p}=\frac{E}{Q(1-\nu^2)}\begin{bmatrix}(\sigma'_{11}+\nu\sigma'_{22})^2 & & \text{对称}\\ (\sigma'_{11}+\nu\sigma'_{22})(\sigma'_{22}+\nu\sigma'_{11}) & (\sigma'_{22}+\nu\sigma'_{11})^2 & \\ (1-\nu)(\sigma'_{11}+\nu\sigma'_{22})\sigma_{12} & (1-\nu)(\sigma'_{22}+\nu\sigma'_{11})\sigma_{12} & (1-\nu)^2\sigma_{12}\sigma_{12}\end{bmatrix}$$

$$Q=\sigma'_{11}\sigma'_{11}+\sigma'_{22}+2\nu\sigma'_{11}\sigma'_{22}+2(1-\nu)\sigma_{12}\sigma_{12}+2E_t(1-\nu)\bar{\sigma}^2/(9G) \quad (3.4\text{-}111)$$

(4) 正交各向异性材料的应力应变矩阵

对于三维问题，$\Delta\boldsymbol{\sigma}$ 和 $\Delta\boldsymbol{\varepsilon}$ 仍采用式（3.4-100）和（3.4-101），满足 Hill 正交各向异性屈服准则的等向强化材料，若弹性阶段符合虎克定律，则其弹性矩阵 C^{e} 仍为式（3.4-99），而 C^{p} 为

$$C^{p}=\frac{b}{E_t+c}S \quad (3.4\text{-}112)$$

式中

$$b=\frac{9G^2}{Y^2}$$

$$c=\frac{9G\ (\sigma'_{11}\sigma'_{11}+\sigma'_{22}\sigma'_{22}+\sigma'_{33}\sigma'_{33}+2\sigma'_{12}\sigma'_{12}+2\sigma'_{23}\sigma'_{23}+2\sigma'_{31}\sigma'_{31})}{2Y^2}$$

$$S=\begin{bmatrix}\sigma'_{11}\sigma'_{11} & & & & & \text{对称}\\ \sigma'_{11}\sigma'_{22} & \sigma'_{22}\sigma'_{22} & & & & \\ \sigma'_{11}\sigma'_{33} & \sigma'_{22}\sigma'_{33} & \sigma'_{33}\sigma'_{33} & & & \\ \sigma'_{11}\sigma'_{12} & \sigma'_{22}\sigma'_{12} & \sigma'_{33}\sigma'_{12} & \sigma'_{12}\sigma'_{12} & & \\ \sigma'_{11}\sigma'_{23} & \sigma'_{22}\sigma'_{23} & \sigma'_{33}\sigma'_{23} & \sigma'_{12}\sigma'_{23} & \sigma'_{23}\sigma'_{23} & \\ \sigma'_{11}\sigma'_{31} & \sigma'_{22}\sigma'_{31} & \sigma'_{33}\sigma'_{31} & \sigma'_{12}\sigma'_{31} & \sigma'_{23}\sigma'_{31} & \sigma'_{31}\sigma'_{31}\end{bmatrix}$$

$$Y=\sqrt{\frac{3}{2\ (F+Q+H)}}$$

$$\sigma'_{ij}=\frac{2}{3}Y^2A_{ij}$$

$A_{11}=Q\ (\sigma_{11}-\sigma_{33})\ +H\ (\sigma_{11}-\sigma_{22})$，$A_{12}=A_{21}=N\sigma_{12}=N\sigma_{21}$

$A_{22}=F\ (\sigma_{22}-\sigma_{33})\ +H\ (\sigma_{22}-\sigma_{11})$，$A_{23}=A_{32}=L\sigma_{23}=L\sigma_{32}$

$A_{33}=Q\ (\sigma_{33}-\sigma_{11})\ +F\ (\sigma_{33}-\sigma_{22})$，$A_{31}=A_{13}=M\sigma_{31}=N\sigma_{13}$

F、Q、H、N、L 和 M 是材料的各向异性参数。

2.1.2 虚功方程

对塑性成形过程进行小变形弹塑性有限元分析时，通常将载荷分解为若干各增量，按增量法求解。设已求得了 t 时刻的解，此时物体在体积力 ${}^{t}b_i$、表面力 ${}^{t}p_i$ 的作用下处于平衡状态，物体 V 中应力场和应变场分别为 ${}^{t}\sigma_{ij}$ 和 ${}^{t}\varepsilon_{ij}$。在 $t+\Delta t$ 时刻物体应该满足的平衡方程为

$${}^{t}\sigma_{ij,j}+\Delta\sigma_{ij,j}+\ ({}^{t}b_i+\Delta b_i)\ =0 \quad (3.4\text{-}113)$$

在 V 内边界条件为

$$({}^{t}\sigma_{ij}+\Delta\sigma_{ij})n_j={}^{t}p_i+\Delta p_i \quad \text{在 } S_p \text{ 上}$$

$${}^{t}u_i+\Delta u_i={}^{t}\bar{u}_i+\Delta u \quad \text{在 } S_u \text{ 上}$$

式中，Δu_i，Δp_i，Δb_i 和 $\Delta\sigma_{ij}$ 为从 t 时刻到 $t+\Delta t$ 时刻这一增量步中的位移增量、表面力增量、体积增量和应力增量。

在 $t+\Delta t$ 时刻，增量形式的虚功方程为

$$\int_V({}^{t}\sigma_{ij}+\Delta\sigma_{ij})\delta(\Delta\varepsilon_{ij})\mathrm{d}V=\int_{S_p}({}^{t}p_i+\Delta p_i)\delta(\Delta u_i)\mathrm{d}S+\int_V({}^{t}b_i+\Delta b_i)\delta(\Delta u_i)\mathrm{d}V \quad (3.4\text{-}114)$$

其中

$$\Delta\varepsilon_{ij}=\frac{1}{2}\ (\Delta u_{i,j}+\Delta u_{j,i})$$

将增量形式的本构方程式（3.4-98）代入式（3.4-114）得

$$\int_V C^{ep}_{ijkl}\Delta\varepsilon_{kl}\delta(\Delta\varepsilon_{ij})\mathrm{d}V=\int_{S_p}^{t+\Delta t}p_i\delta(\Delta u_i)\mathrm{d}S+\int_V^{t+\Delta t}b_i\delta(\Delta u_i)\mathrm{d}V-\int_V^{t}\sigma_{ij}\delta(\Delta\varepsilon_{ij})\mathrm{d}V \quad (3.4\text{-}115)$$

其中

$${}^{t+\Delta t}p_i={}^{t}p_i+\Delta p_i$$

$${}^{t+\Delta t}b_i={}^{t}b_i+\Delta b_i$$

2.1.3 有限元方程

将所分析的弹塑性材料离散化，对于任一单元采用与弹性变形有限元分析相同的位移（位移增量）插值函数和应变矩阵，并采用矩阵形式的弹塑性本构方程式（3.4-99）。

$$\left.\begin{aligned}{}^{t}u&=N^{t}u^{e}\\ \Delta u&=N\Delta u^{e}\end{aligned}\right\} \quad (3.4\text{-}116)$$

$$\Delta\varepsilon=B\Delta u^{e} \quad (3.4\text{-}117)$$

$$\Delta \boldsymbol{\sigma} = \boldsymbol{C}^{ep} \boldsymbol{B} \Delta \boldsymbol{u}^e \quad (3.4\text{-}118)$$

则将以上各式代入式（3.4-115）得到任一单元的虚功方程

$$\int_{V^e} \delta(\Delta \boldsymbol{u}^e)^{\mathrm{T}} \boldsymbol{B}^{\mathrm{T}} \boldsymbol{C}^{ep} \boldsymbol{B} \Delta \boldsymbol{u}^e \mathrm{d}V$$

$$= \int_{S_p^e} \delta(\Delta \boldsymbol{u}^e)^{\mathrm{T}} \boldsymbol{N}^{\mathrm{T}\ t+\Delta t} \boldsymbol{p} \mathrm{d}S + \int_{V^e} \delta(\Delta \boldsymbol{u}^e)^{\mathrm{T}} \boldsymbol{N}^{\mathrm{T}\ t+\Delta t} \boldsymbol{b} \mathrm{d}V -$$

$$\int_{V^e} \delta(\Delta \boldsymbol{u}^e)^{\mathrm{T}} \boldsymbol{B}^{\mathrm{T} t} \boldsymbol{\sigma} \mathrm{d}V \quad (3.4\text{-}119)$$

由于节点位移增量 Δu^e 和节点虚位移增量 $\delta(\Delta u^e)$ 与单元体积及面积分无关，由上式可得

$$\int_{V^e} \boldsymbol{B}^{\mathrm{T}} \boldsymbol{C}^{ep} \boldsymbol{B} \Delta \boldsymbol{u}^e \mathrm{d}V$$

$$= \int_{S_p^e} \boldsymbol{N}^{\mathrm{T}\ t+\Delta t} p \mathrm{d}S + \int_{V^e} \boldsymbol{N}^{\mathrm{T}\ t+\Delta t} b \mathrm{d}V - \int_{V^e} \boldsymbol{B}^{\mathrm{T}\ t} \sigma \mathrm{d}V$$

$$(3.4\text{-}120)$$

令

$$K^e = \int_{V^e} \boldsymbol{B}^{\mathrm{T}} \boldsymbol{C}^{ep} \boldsymbol{B} \mathrm{d}V$$

$$P^e = \int_{S_p^e} \boldsymbol{N}^{\mathrm{T}\ t+\Delta t} p \mathrm{d}S + \int_{V^e} \boldsymbol{N}^{\mathrm{T}\ t+\Delta t} b \mathrm{d}V$$

$${}^t F^e = \int_{V^e} \boldsymbol{B}^{\mathrm{T}\ t} \sigma \mathrm{d}V$$

$$\Delta R^e = P^e - {}^t F^e$$

式（3.4-120）可写成

$$K^e \Delta u^e = \Delta R^e \quad (3.4\text{-}121)$$

上式即单元刚度方程。把所有单元刚度方程进行集合，即可获得整体有限元方程

$$\boldsymbol{K} \Delta \boldsymbol{U} = \Delta \boldsymbol{R} \quad (3.4\text{-}122)$$

其中

$$\boldsymbol{K} = \sum_e K^e$$

$$\Delta \boldsymbol{R} = \sum_e \Delta R^e$$

式中，$\boldsymbol{K}$ 为整体刚度矩阵；$\Delta \boldsymbol{U}$ 为整体节点位移增量列阵；$\Delta \boldsymbol{R}$ 为整体节点载荷增量列阵。

实际物体弹塑性变形过程中，内部可能存在弹性区、过渡区、弹塑性区和塑性卸载区四种不同状态的区域。对于每个积分点，都要判断它属于哪个区域，以便采用适当的弹塑性矩阵的相应形式。物体中的四种区域，可以用屈服函数区分如下。

（1）弹性区

$$F\ ({}^t\sigma_{ij}, {}^tY)\ < 0$$

$$F\ ({}^{t+\Delta t}\sigma_{ij}, {}^{t+\Delta t}Y)\ < 0$$

即在 t 时刻和 $t+\Delta t$ 时刻都处于弹性变形状态，$\alpha = 0$，$\boldsymbol{C}^{ep} = \boldsymbol{C}^e$。

（2）弹塑性区

$$F\ ({}^t\sigma_{ij}, {}^tY)\ \geqslant 0$$

$$F\ ({}^{t+\Delta t}\sigma_{ij}, {}^{t+\Delta t}Y)\ \geqslant 0$$

即在 t 时刻和 $t+\Delta t$ 时刻都处于弹塑性变形状态，$\alpha = 1$，$\boldsymbol{C}^{ep} = \boldsymbol{C}^e - \boldsymbol{C}^p$。

（3）塑性卸载区

$$F\ ({}^t\sigma_{ij}, {}^tY)\ \geqslant 0$$

$$F\ ({}^{t+\Delta t}\sigma_{ij}, {}^{t+\Delta t}Y)\ < 0$$

即在 t 时刻处于弹塑性状态，在增量步中发生卸载，$\alpha = 0$，$\boldsymbol{C}^{ep} = \boldsymbol{C}^e$。

（4）过渡区

$$F\ ({}^t\sigma_{ij}, {}^tY)\ < 0$$

$$F\ ({}^{t+\Delta t}\sigma_{ij}, {}^{t+\Delta t}Y)\ \geqslant 0$$

即在 t 时刻处于弹性状态，在增量中进入弹塑性状态。可将此增量步分为两个阶段：第一阶段为弹性变形，应变增量为 $m\Delta\varepsilon_{ij}$（$0 < m < 1$），正好使材料屈服，$\alpha = 0$，$\boldsymbol{C}^{ep} = \boldsymbol{C}^e$；第二阶段为弹塑性变形，应变增量为 $(1-m)\Delta\varepsilon_{ij}$，$\alpha = 1$，$C^{ep} = \boldsymbol{C}^e - \boldsymbol{C}^p$。于是整个增量步的弹塑性矩阵为 $m\boldsymbol{C}^e +（1-m）（\boldsymbol{C}^e - \boldsymbol{C}^p）= \boldsymbol{C}^e -（1-m）\boldsymbol{C}^p$。上式中可令 $\alpha = 1-m$，于是确定上述四种区域的弹塑性矩阵的问题归结为确定 α 值的问题（$0 \leqslant \alpha \leqslant 1$）。

在增量步计算中，首先按弹性区计算，算得的应力值为 ${}^{t+\Delta t}\sigma_{ij}^*$。设此时的流动应力为 σ_s，若 $F({}^{t+\Delta t}\sigma_{ij}^*, \sigma_s) > 0$，则该积分点处于过渡区，可取

$$m = \frac{\sigma_s - {}^t\bar{\sigma}}{{}^{t+\Delta t}\bar{\sigma} - {}^t\bar{\sigma}}$$

式中　${}^t\bar{\sigma}$ 和 ${}^{t+\Delta t}\bar{\sigma}$——$t$ 时刻和 $t+\Delta t$ 时刻的等效应力。

2.2　有限应变弹塑性有限元分析

非线性问题包括材料非线性和几何非线性，对于工程应用中具有较强非线性特点的问题，需要采用非线性理论来研究大位移和大变形。因此采用有限应变弹塑性有限元法可以较为准确的对这类问题进行模拟分析。在有限应变弹塑性有限元分析中，可以按欧拉描述，也可以按拉格朗日描述。一般地说，流体力学问题常采用欧拉描述，固体力学问题常采用拉格朗日描述。

2.2.1　应力应变矩阵

对于有限变形问题，在变形态构形上的应力应变关系的速率型方程形式为

$$\hat{\sigma}_{ij} = C_{ijkl}^{ep} d_{kl} \quad (3.4\text{-}123)$$

其中，C_{ijkl}^{ep} 与小变形弹塑性本构方程中的对应项 C_{ijkl}^{ep} 相同，其中的应力分量和应力偏量采用柯西应力分量和柯西应力偏量。

$$\dot{E}_{ij} = \frac{1}{2}\left(\frac{\partial v_k}{\partial X_i}\frac{\partial x_k}{\partial X_j} + \frac{\partial x_k}{\partial X_i}\frac{\partial v_k}{\partial X_j}\right) = \frac{\partial x_k}{\partial X_i}\frac{\partial x_l}{\partial X_j} d_{kl} \quad (3.4\text{-}124)$$

又

$$\sigma_{ij} = \frac{\rho}{\rho_0}\frac{\partial x_i}{\partial X_\alpha}\frac{\partial x_j}{\partial X_\beta} S_{\alpha\beta}$$

对上式求物质导数，并把式（3.4-123）和式（3.4-124）代入，认为弹塑性变形中材料体积不变，可得拉格朗日描述的速率型本构方程

$$\dot{S}_{ij} = \frac{\partial X_i}{\partial x_\alpha}\frac{\partial X_i}{\partial X_\alpha} C_{\alpha\beta kl}^{ep} \frac{\partial X_m}{\partial X_k}\frac{\partial X_n}{\partial x_l}\dot{E}_{mn} - \frac{\partial X_j}{\partial x_p}\frac{\partial X_n}{\partial x_p} S_{im}\dot{E}_{mn} - \frac{\partial X_i}{\partial x_p}\frac{\partial X_n}{\partial x_p} S_{jm}\dot{E}_{mn}$$

$$(3.4\text{-}125)$$

在计算中可先令 $x_\alpha = X_\alpha$，求出近似解，然后用迭代法作进一步的修正。这样，式（3.4-125）可写成

$$\dot{S}_{ij} = C_{ijmn}^{ep} - \sigma_{im}\dot{E}_{mj} - \sigma_{jm}\dot{E}_{mi} \quad (3.4\text{-}126)$$

其矩阵形式为

$$\Delta \boldsymbol{S} = (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\ \Delta \boldsymbol{E} \quad (3.4\text{-}127)$$

这就是修正拉格朗日法的克希荷夫应力张量增量与格林应变张量增量关系的矩阵表达式。对于三维空间问题，符合 Mises 屈服准则和等向强化的弹塑性材料，式（3.4-127）中各项写出如下：

a)　$\Delta \boldsymbol{S} = [\Delta S_{11}\quad \Delta S_{22}\quad \Delta S_{33}\quad \Delta S_{12}\quad \Delta S_{23}\quad \Delta S_{31}]^{\mathrm{T}}$　(3.4-128)

b)　$\Delta \boldsymbol{E} = [\Delta E_{11}\quad \Delta E_{22}\quad \Delta E_{33}\quad 2\Delta E_{12}\quad 2\Delta E_{23}\quad 2\Delta E_{31}]$　(3.4-129)

其中，式（3.4-129）中诸分量与位移增量分量 Δu_i 的关系为

$$\Delta E_{ij} = \frac{1}{2}\left[\frac{\partial(\Delta u_i)}{\partial X_j} + \frac{\partial(\Delta u_j)}{\partial X_i} + \frac{\partial(\Delta u_k)}{\partial X_i}\frac{\partial(\Delta u_k)}{\partial X_j}\right]$$

$$(3.4\text{-}130)$$

c)　$\boldsymbol{C}^{ep} = \boldsymbol{C}^e - \alpha \boldsymbol{C}^p$

d) $\tau_\sigma =$

$$\begin{bmatrix} 2\sigma_{11} & & & & & \\ 0 & 2\sigma_{22} & & & \text{对称} & \\ 0 & 0 & 2\sigma_{33} & & & \\ \sigma_{12} & \sigma_{12} & 0 & \frac{1}{2}(\sigma_{11}+\sigma_{22}) & & \\ 0 & \sigma_{23} & \sigma_{23} & \frac{1}{2}\sigma_{31} & \frac{1}{2}(\sigma_{22}+\sigma_{23}) & \\ \sigma_{31} & 0 & \sigma_{31} & \frac{1}{2}\sigma_{23} & \frac{1}{2}\sigma_{12} & \frac{1}{2}\sigma_{11}+\sigma_{33} \end{bmatrix} \quad (3.4\text{-}131)$$

2.2.2 虚功率方程

在变形态物体上用欧拉变量描述的虚功率方程为

$$\int_V \sigma_{ij}\delta d_{ij}\mathrm{d}V = \int_{S_p} p_i\delta v_i\mathrm{d}S + \int_V b_i\delta v_i\mathrm{d}V \quad (3.4\text{-}132)$$

又

$$\dot{E}_{ij} = \frac{\partial x_k}{\partial x_i}\frac{\partial X_l}{\partial X_j}d_{kl}$$

$$S_{ij} = J\frac{\partial X_i}{\partial x_k}\frac{\partial X_j}{\partial x_l}\sigma_{kl}$$

$$\mathrm{d}V = J\mathrm{d}V_0$$

由以上三式可将式（3.4-132）左边改写为

$$\int_V \sigma_{ij}\delta d_{ij}\mathrm{d}V = \int_{V_0} S_{ij}\delta\dot{E}_{ij}\mathrm{d}V_0$$

式中，V_0 为与 V 相对应的物体在初始参考构形时的体积。

对于保守载荷的情况，有

$$b_i^0\mathrm{d}V_0 = b_i\mathrm{d}V$$

$$p_i^0\mathrm{d}S_0 = p_i\mathrm{d}S$$

式中，b_i^0、p_i^0 为与 b_i^0 和 p_i^0 相对应的在参考态构形中的单位体积的体力和受载表面 S_{0P} 上的表面力分量。

因此从式（3.4-132）得到用克希荷夫应力张量和格林应变张量变化率表示的，拉格朗日描述的虚功率方程，即

$$\int_{V_0} S_{ij}\delta\dot{E}_{ij}\mathrm{d}V_0 = \int_S p_i^0\delta v_i\mathrm{d}S_0 + \int_{V_0} b_i^0\delta v_i\mathrm{d}V_0 \quad (3.4\text{-}133)$$

若在位移约束表面 S_{0u} 上已知其质点的速度为 $\bar{v}_i$，则在 S_{0u} 边界面上的速度边界条件为

$$v_i = \bar{v}_i$$

2.2.3 有限元方程

（1）应变增量及其分解

格林应变为

$$E = {}^tE + \Delta E$$

用 $\Delta E_{ij}^{\mathrm{L}}$ 和 $\Delta E_{ij}^{\mathrm{N}}$ 分别表示应变增量的线性部分和非线性部分，则格林应变增量 ΔE_{ij} 可写为

$$\Delta E_{ij} = \Delta E_{ij}^{\mathrm{L}} + \Delta E_{ij}^{\mathrm{N}} \quad (i,\ j=1,\ 2,\ 3) \quad (3.4\text{-}134)$$

式中

$$2\Delta E_{ij}^{\mathrm{L}} = \frac{\partial\Delta u_i}{\partial X_j} + \frac{\partial\Delta u_j}{\partial X_i} \quad (3.4\text{-}135)$$

$$2\Delta E_{ij}^{\mathrm{N}} = \frac{\partial\Delta u_k}{\partial X_i}\frac{\partial\Delta u_k}{\partial X_j} \quad (3.4\text{-}136)$$

以上三式的矩阵形式为

$$\Delta\boldsymbol{E} = \Delta\boldsymbol{E}^{\mathrm{L}} + \Delta\boldsymbol{E}^{\mathrm{N}} \quad (3.4\text{-}137)$$

$$\Delta\boldsymbol{E}^{\mathrm{L}} = \boldsymbol{L}\Delta\boldsymbol{u} \quad (3.4\text{-}138)$$

$$\Delta\boldsymbol{E}^{\mathrm{N}} = \frac{1}{2}\Delta\boldsymbol{\theta}\Delta\boldsymbol{\beta} \quad (3.4\text{-}139)$$

其中

$$\boldsymbol{L} = \begin{bmatrix} \frac{\partial}{\partial X_1} & 0 & 0 \\ 0 & \frac{\partial}{\partial X_2} & 0 \\ 0 & 0 & \frac{\partial}{\partial X_3} \\ \frac{\partial}{\partial X_2} & \frac{\partial}{\partial X_1} & 0 \\ 0 & \frac{\partial}{\partial X_3} & \frac{\partial}{\partial X_2} \\ \frac{\partial}{\partial X_3} & 0 & \frac{\partial}{\partial X_1} \end{bmatrix} \quad (3.4\text{-}140)$$

$$\Delta\boldsymbol{u} = [\Delta u_1\quad \Delta u_2\quad \Delta u_3]^{\mathrm{T}} \quad (3.4\text{-}141)$$

$$\Delta\boldsymbol{\beta}_I = \left[\frac{\partial\Delta u_1}{\partial X_I}\quad \frac{\partial\Delta u_2}{\partial X_I}\quad \frac{\partial\Delta u_3}{\partial X_I}\right]^{\mathrm{T}} \quad (I=1,\ 2,\ 3) \quad (3.4\text{-}142)$$

$$\Delta\boldsymbol{\beta} = [\Delta\boldsymbol{\beta}_1^{\mathrm{T}}\quad \Delta\boldsymbol{\beta}_2^{\mathrm{T}}\quad \Delta\boldsymbol{\beta}_3^{\mathrm{T}}]^{\mathrm{T}} \quad (3.4\text{-}143)$$

$$\Delta\boldsymbol{\theta} = \begin{bmatrix} \Delta\boldsymbol{\beta}_1^{\mathrm{T}} & \boldsymbol{0}^{\mathrm{T}} & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{0}^{\mathrm{T}} & \Delta\boldsymbol{\beta}_2^{\mathrm{T}} & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{0}^{\mathrm{T}} & \boldsymbol{0}^{\mathrm{T}} & \Delta\boldsymbol{\beta}_3^{\mathrm{T}} \\ \Delta\boldsymbol{\beta}_2^{\mathrm{T}} & \Delta\boldsymbol{\beta}_1^{\mathrm{T}} & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{0}^{\mathrm{T}} & \Delta\boldsymbol{\beta}_3^{\mathrm{T}} & \Delta\boldsymbol{\beta}_2^{\mathrm{T}} \\ \Delta\boldsymbol{\beta}_3^{\mathrm{T}} & \boldsymbol{0}^{\mathrm{T}} & \Delta\boldsymbol{\beta}_1^{\mathrm{T}} \end{bmatrix} \quad (3.4\text{-}144)$$

（2）单元增量刚度方程和整体增量刚度方程

1）单元应变增量　三维连续体离散化，并选取单元位移插值函数后，单元内任一点的位移增量可用节点位移增量表示为

$$\Delta\boldsymbol{u} = \boldsymbol{N}\Delta\boldsymbol{u}^{\mathrm{e}} \quad (3.4\text{-}145)$$

式中　$\boldsymbol{N}$——形函数矩阵；

$\Delta\boldsymbol{u}^{\mathrm{e}}$——单元节点位移增量列阵。

将式（3.4-145）代入式（3.4-138）得

$$\Delta\boldsymbol{E}^{\mathrm{L}} = \boldsymbol{LN}\Delta\boldsymbol{u}^{\mathrm{e}} = \boldsymbol{B}_{\mathrm{L}}\Delta\boldsymbol{u}^{\mathrm{e}} \quad (3.4\text{-}146)$$

式中的 $\boldsymbol{B}_{\mathrm{L}}$ 与线弹性或小变形弹塑性应变矩阵是相同的，称为线性应变矩阵。将式（3.4-145）代入式（3.4-142）有

$$\Delta\boldsymbol{\beta}_I = \boldsymbol{G}_I\Delta\boldsymbol{u}^{\mathrm{e}} \quad (I=1,\ 2,\ 3) \quad (3.4\text{-}147)$$

式中

$$\boldsymbol{G}_I = [G_I^{(1)}\quad G_I^{(2)}\quad \cdots\quad G_I^{(n)}] \quad (3.4\text{-}148)$$

n 为单元节点数，而其中

$$G_I^{(k)} = \begin{bmatrix} \frac{\partial N^{(k)}}{\partial X_I} & 0 & 0 \\ 0 & \frac{\partial N^{(k)}}{\partial X_I} & 0 \\ 0 & 0 & \frac{\partial N^{(k)}}{\partial X_I} \end{bmatrix} \quad (k=1,\ 2,\ \cdots,\ n) \quad (3.4\text{-}149)$$

把 I 作行号，并按1，2，3排列后得

$$\Delta\boldsymbol{\beta} = \boldsymbol{G}\Delta\boldsymbol{u}^{\mathrm{e}} \quad (3.4\text{-}150)$$

式中

$$\boldsymbol{G} = \begin{bmatrix} G_1^{(1)} & G_1^{(2)} & \cdots & G_1^{(n)} \\ G_2^{(1)} & G_2^{(2)} & \cdots & G_2^{(n)} \\ G_3^{(1)} & G_3^{(2)} & \cdots & G_3^{(n)} \end{bmatrix} \quad (3.4\text{-}151)$$

把式（3.4-150）代入式（3.4-139）得

$$\Delta\boldsymbol{E}^{\mathrm{N}} = \frac{1}{2}\Delta\boldsymbol{\theta}\boldsymbol{G}\Delta\boldsymbol{u}^{\mathrm{e}} = \boldsymbol{B}_{\mathrm{N}}^{*}\Delta\boldsymbol{u}^{\mathrm{e}} \quad (3.4\text{-}152)$$

式中，$\boldsymbol{B}_{\mathrm{N}}^{*}$ 称为非线性应变矩阵。

将式（3.4-139）和式（3.4-152）代入式（3.4-137），得

$$\Delta \boldsymbol{E} = \boldsymbol{B}\Delta \boldsymbol{u}^e \tag{3.4-153}$$

其中

$$\boldsymbol{B} = \boldsymbol{B}_L + \boldsymbol{B}_N^* \tag{3.4-154}$$

2）单元应力及其增量　按UL法求解时，以 t 时刻构形为参考构形，t 时刻构形中的克希荷夫应力即为柯西应力，而 $t+\Delta t$ 时刻的克希荷夫应力列阵为

$$\boldsymbol{S} = \boldsymbol{\tau} + \Delta \boldsymbol{S} \tag{3.4-155}$$

其中，τ 是 t 时刻以 X_i 为拉格朗日坐标的参考构形的单元中任一点的柯西应力。把式（3.4-153）代入式（3.4-127）可得式中的 $\Delta \boldsymbol{S}$ 为

$$\Delta \boldsymbol{S} = (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\boldsymbol{B}\Delta \boldsymbol{u}^e = (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)(\boldsymbol{B}_L + \boldsymbol{B}_N^*)\Delta \boldsymbol{u}^e \tag{3.4-156}$$

3）单元增量刚度方程和整体增量刚度方程　将虚功率方程式（3.4-133）应用于 t 时刻参考构形中的任一单元 e，写成增量形式，得

$$\int_{V^e} \delta(\Delta \boldsymbol{E})^T S \mathrm{d}V = \int_{S_p^e} \delta(\Delta \boldsymbol{u})^T \boldsymbol{p} \mathrm{d}S + \int_{V^e} \delta(\Delta \boldsymbol{u})^T \boldsymbol{b}\rho \mathrm{d}V \tag{3.4-157}$$

把式（3.4-155）和式（3.4-156）代入上式，得

$$\int_{V^e} \delta(\Delta \boldsymbol{E})^T [\boldsymbol{\tau} + (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\Delta \boldsymbol{E}] \mathrm{d}V = \int_{S_p^e} \delta(\Delta \boldsymbol{u})^T \boldsymbol{p} \mathrm{d}S + \int_{V^e} \delta(\Delta \boldsymbol{u})^T \boldsymbol{b}\rho \mathrm{d}V \tag{3.4-158}$$

式中

$$\begin{aligned}\delta \boldsymbol{E} &= \delta \boldsymbol{E}^L + \delta \boldsymbol{E}^N = \boldsymbol{B}_L \delta \Delta \boldsymbol{u}^e + \frac{1}{2}(\boldsymbol{\delta}\Delta\boldsymbol{\theta}\Delta\boldsymbol{\beta} + \Delta\boldsymbol{\theta}\delta\Delta\boldsymbol{\beta}) \\ &= \boldsymbol{B}_L \delta \Delta \boldsymbol{u}^e + \Delta\boldsymbol{\theta}\delta\Delta\boldsymbol{\beta} = (\boldsymbol{B}_L + \boldsymbol{B}_N)\delta\Delta \boldsymbol{u}^e\end{aligned} \tag{3.4-159}$$

其中 $\boldsymbol{B}_N = 2\boldsymbol{B}_N^*$。

于是式（3.4-158）左边可以写成

$$\begin{aligned}&\int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T (\boldsymbol{B}_L^T + \boldsymbol{B}_N^T)\boldsymbol{\tau} \mathrm{d}V + \int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T (\boldsymbol{B}_L^T + \boldsymbol{B}_N^T) \\ &(\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)(\boldsymbol{B}_L^T + \boldsymbol{B}_N^*)\Delta \boldsymbol{u}^e \mathrm{d}V = \int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T \boldsymbol{B}_L^T \boldsymbol{\tau} \mathrm{d}V + \\ &\int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T \boldsymbol{G}^T \Delta\boldsymbol{\theta}^T \boldsymbol{\tau} \mathrm{d}V + \int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T \boldsymbol{B}_L^T \boldsymbol{C}^{ep} \boldsymbol{B}_L \Delta \boldsymbol{u}^e \mathrm{d}V - \\ &\int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T \boldsymbol{B}_L^T \boldsymbol{\tau}_\sigma \boldsymbol{B}_L \Delta \boldsymbol{u}^e \mathrm{d}V + \int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T [\boldsymbol{B}_L^T (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\boldsymbol{B}_N^* + \\ &\boldsymbol{B}_L^T (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\boldsymbol{B}_L + \boldsymbol{B}_N^T (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\boldsymbol{B}_N^*]\Delta \boldsymbol{u}^e \mathrm{d}V\end{aligned} \tag{3.4-160}$$

其中

$$\int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T \boldsymbol{G}^T \Delta\boldsymbol{\theta}^T \boldsymbol{\tau} \mathrm{d}V = \int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T \boldsymbol{G}^T \boldsymbol{T}\boldsymbol{G}\Delta \boldsymbol{u}^e \mathrm{d}V$$

$$\boldsymbol{T} = \begin{bmatrix} \sigma_{11} I_{3\times3} & & \text{对称} \\ \sigma_{12} I_{3\times3} & \sigma_{22} I_{3\times3} & \\ \sigma_{13} I_{3\times3} & \sigma_{23} I_{3\times3} & \sigma_{33} I_{3\times3} \end{bmatrix}$$

而式（3.4-158）右边可写成

$$\int_{S_p^e} \delta(\Delta \boldsymbol{u}^e)^T \boldsymbol{N}^T \boldsymbol{p}^T \mathrm{d}S + \int_{V^e} \delta(\Delta \boldsymbol{u}^e)^T \boldsymbol{N}^T \boldsymbol{b}\rho \mathrm{d}V$$

综合以上各式可得到单元增量刚度方程

$$(\boldsymbol{k}_0^e + \boldsymbol{k}_\sigma^e + \boldsymbol{k}_1^e)\Delta \boldsymbol{u}^e = \boldsymbol{P}^e - \boldsymbol{F}^e - \boldsymbol{Q}^e \tag{3.4-161}$$

式中

$$\boldsymbol{k}_0^e = \int_{V^e} \boldsymbol{B}_L^T \boldsymbol{C}^{ep} \boldsymbol{B}_L \mathrm{d}V \tag{3.4-162}$$

$$\boldsymbol{k}_\sigma^e = \int_{V^e} \boldsymbol{G}^T \boldsymbol{T}\boldsymbol{G} \mathrm{d}V \tag{3.4-163}$$

$$\boldsymbol{k}_1^e = \int_{V^e} \boldsymbol{B}_N^T (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\boldsymbol{B}_N^* \mathrm{d}V \tag{3.4-164}$$

$$\boldsymbol{P}^e = \int_{S_p^e} \boldsymbol{N}^T \boldsymbol{p} \mathrm{d}S + \int_{V^e} \boldsymbol{N}^T \boldsymbol{b}\rho \mathrm{d}V \tag{3.4-165}$$

$$\boldsymbol{F}^e = \int_{V^e} \boldsymbol{B}_L^T \boldsymbol{\tau} \mathrm{d}V \tag{3.4-166}$$

$$\boldsymbol{Q}^e = \int_{V^e} [\boldsymbol{B}_L^T (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\boldsymbol{B}_N + \boldsymbol{B}_N^T (\boldsymbol{C}^{ep} - \boldsymbol{\tau}_\sigma)\boldsymbol{E}_L] \mathrm{d}V \tag{3.4-167}$$

以上几式中，$\boldsymbol{k}_0^e$ 与位移增量 $\Delta \boldsymbol{u}^e$ 无关，称为小位移刚度矩阵；$\boldsymbol{k}_\sigma^e$ 与已知应力 $\boldsymbol{\tau}$ 有关，称为初应力刚度矩阵；$\boldsymbol{k}_1^e$ 与位移增量有关，称为大位移刚度矩阵；为形成对称刚度矩阵；将产生不对称矩阵的项移至等式右边作为虚拟载荷 $\boldsymbol{Q}^e$；$\boldsymbol{F}^e$ 称为初应力节点内力；$\boldsymbol{P}^e$ 是外力（面力和体力）节点载荷。

式（3.4-161）可写为

$$\boldsymbol{K}^e \Delta \boldsymbol{u}^e = \boldsymbol{P}^e - \boldsymbol{F}^e - \boldsymbol{Q}^e \tag{3.4-168}$$

根据将单元刚度方程集合成整体刚度方程的方法，把所有单元增量刚度方程进行集合，就可得到整体刚度方程

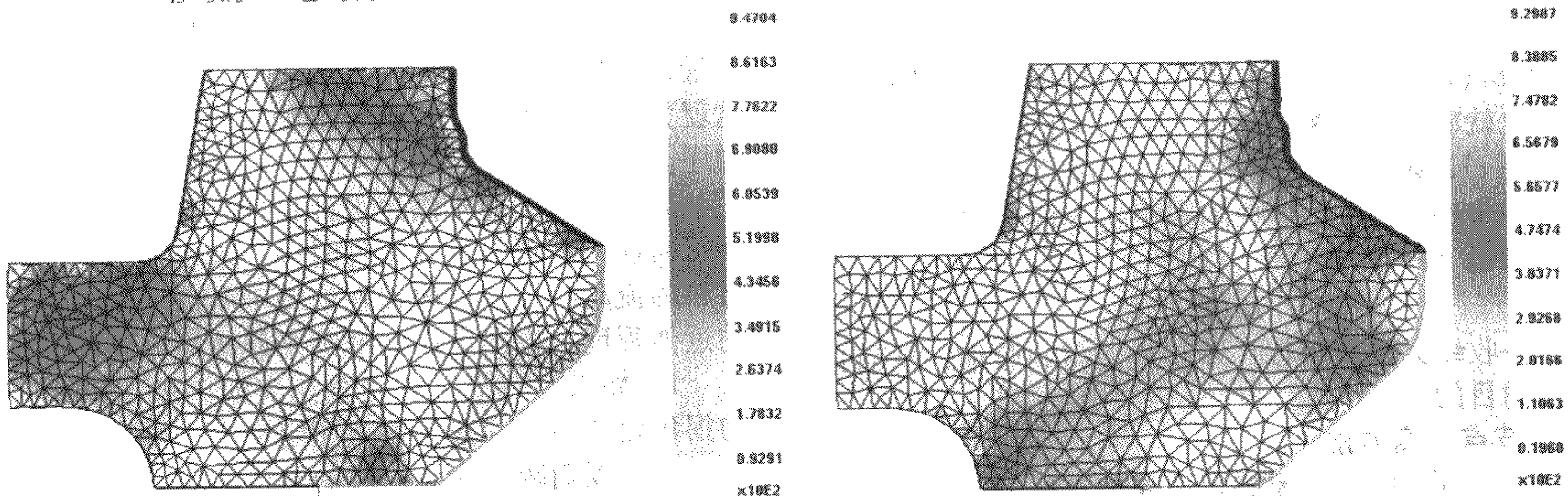
$$\boldsymbol{K}\Delta \boldsymbol{U} = \boldsymbol{P} - \boldsymbol{F} - \boldsymbol{Q} = \boldsymbol{R} \tag{3.4-169}$$

式中，$\boldsymbol{K}$ 为整体切线刚度矩阵；$\Delta \boldsymbol{U}$ 为整体节点位移增量列阵；$\boldsymbol{P}$ 为整体外力节点载荷列阵；$\boldsymbol{F}$ 为整体初始应力内力节点力列阵；$\boldsymbol{Q}$ 为整体虚拟载荷列阵；$\boldsymbol{R}$ 为整体节点不平衡力列阵。

求解此非线性方程组，即可得到节点位移增量。

2.2.4 计算实例

仍以上一章实例介绍的轿车圆锥齿轮冷精锻过程为例，对其变形过程进行弹塑性分析。材料的参数与前例相同，只是材料的类型要从刚塑性改为弹塑性，冲头的挤压速率不变，但考虑到收敛问题，将计算步长减小。在锻压过程结束后，改变冲头运动方向并开启上模进行卸载回弹分析，得到的齿轮锻压终了和回弹后的等效应力分布见图3.4-13，回弹前后锻件尺寸变化如图3.4-14所示。由图可见，回弹后工件的应力得到释缓，而尺寸也发生了相应的变化。



(a)回弹前　　(b)回弹后

图3.4-13　齿轮成形回弹前后的等效应力分布比较

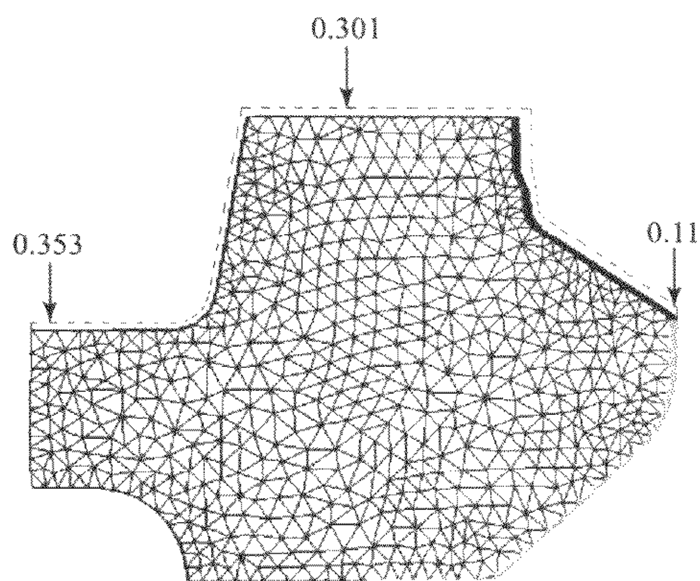


图3.4-14　齿轮成形回弹前后的尺寸变化

2.3　有限元反向模拟

2.3.1　概述

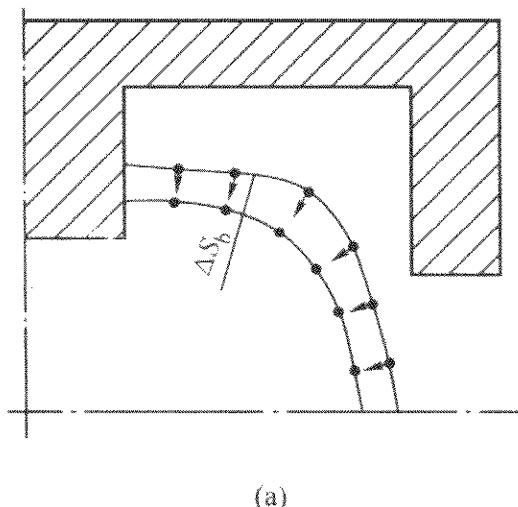
现在，计算机模拟在产品设计和材料成形过程中发挥越来越重要的作用，通过数值模拟计算，可以全面了解材料在变形过程中的应力应变场、温度场并预测成形缺陷的产生。正向模拟是从模具接触坯料开始，分析坯料在模具型腔内的变形填充过程，直到材料完全充满型腔。而反向模拟则相反，上模和下模最初是闭合的，并且冲头也达到要求的行程，开始模拟时，模具分开同时冲头回程，坯料做收缩运动，最后得到几何形状简单的坯料或预锻件。

作为塑性加工领域的前沿课题之一，有限元反向模拟技术使得技术人员更为合理地预测与优化坯料和预锻件的形状，更为有效和合理地控制终锻件的成形及其内部应力应变的分布。

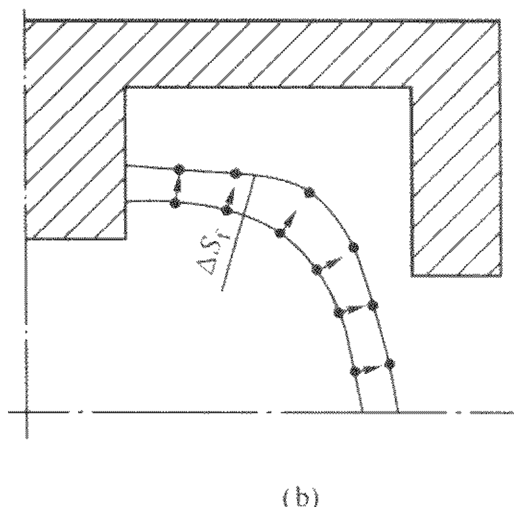
2.3.2　自由边的收缩

所谓坯料的自由边，是指材料在模具型腔中不与模腔内壁接触的边，进行反向模拟就是从贴合状态收缩到正向模拟的起始状态，所以从一定意义上讲，反向模拟就是材料自由边的收缩过程。

收缩时，每一步反向模拟到达的新位置都要用正向模拟来进行检验，以保证在相同的时间增量内，材料的自由边能够回到该步反向模拟前的位置，也即正反向模拟的位置统一性。由于自由边的形状是由边界节点位置确定的，所以自由边的收缩运动也即是边界节点的方向运动。可以设定边界节点的运动方向为该条边界的法线方向（如图3.4-15所示）。



(a)



(b)

图3.4-15　自由边的收缩原理

如果以 B 表示材料在模具型腔内的边界几何条件（包括自由边和非自由边），用 $B_b(i)$ 表示表示反向模拟进行到第 i 步（$t=t(i)$）时材料边界的几何位置，用 $B_f(i)$ 表示在第 i 步开始进行正向模拟时的边界，以 $\Delta S_b(i)$ 表示在反向模拟进行下一步计算时的边界收缩量，用 $\Delta S_f(i)$ 表示在正向模拟第 i 步上进行下一步计算时的边界扩张量，以 Q 来表示对处于某一几何边界条件下的速度场进行优化。则具体的计算步骤如下：

1）$\Delta S_b(i)=\Delta S$

2）$B_b(i+1)=B_b(i)+\Delta S(i)$

3）$B_f(i+1)=B_b(i+1)$

4）$Q(B_f(i+1))\rightarrow\Delta S_f(i+1)$

5）$B_f(i)=B_f(i+1)-\Delta S_f(i+1)$

6）如果 $B_b(i)=B_f(i)$，如果 $\Delta S_b(i)$ 是理想的，则 $B_b(i+1)$ 的位置也就确定了，模拟步数（$i=i+1$），返回1）；如果 $B_b(i)\neq B_f(i)$，则令 $\Delta S_b(i)=\Delta S_f(i+1)$，返回2）。

上述第一步中的 ΔS 是任选的一个边界收缩量，对应于任选的时间增量 Δt，即有

$$\Delta S=Q[B_f(i)]\Delta t$$

采用这一收缩方法就可以一步步地使材料收缩，从而达到反向模拟的目的。

2.3.3　边界节点分离方法

在反向模拟过程中，需要确定边界节点的分离方法。一般有两种主要方法，即模具接触时序的反向跟踪法和Teterin形状复杂因子法。下面主要介绍方向跟踪法。

反向模拟过程中，边界条件时序能通过试制预锻件的正向模拟而获得，虽然此预锻件还不完全满足设计目标。需要开发一种记录独立于网格划分所需要的边界条件时序的方法，来解决边界节点的接触时序问题。一旦此时序不知道，必须确定一种适合反向模拟的修改时序方法。

反向跟踪法把模具表面分成一些直线或者圆弧片段，在应用试制坯的正向模拟过程中，计算并记录下每一个模具片段与变形体接触的时刻。应用这种方法，就不需要获得初始网格划分或相同数量的边界节点。如果某个模具片段反复与工件接触分离，那么仅最后接触时刻被记录下来。在应用模具的接触时间时，为了充分反映边界条件的变化，每个模具片段的长度应该小于工件边界单元的最小边长。对于这样的模具片段，可能有某些片段在工件表面节点之间，这样永不和边界节点接触，因此，关于它们的接触时序，通过应用两个最近的模具片段的时序线性插值来确定为最终的数据点，这就为在两节点间的模具片段能接触工件提供了一个连续的时序。

如果此试制坯不满足设计目标，那必须修改直接从试制坯正向模拟获得的边界条件时序，来实现模具的充满状态和材料的流动特性。修改程度依赖于所选的试制坯与希望形状的近似程度。提出下面的准则来修改边界条件。

1）如果有一个没有完全充满的区域，则此区域未充满的拐角处的模具片段接触时序定义为上模到达所需最终位置的时间，逐渐减少在此片段的两边的时间来匹配最近已被定义时序了的片段。

2）对于有多于一个深腔的锻模，在最早被充满的型腔中，模具片段的接触时序应该被减少，使得所有的深腔在锻

造过程结束时同时被充满。如果一个或多个深腔没被充满，按准则1）修改时序。此准则能确保最终的加工中合理的材料体积分配。

3）如果在试制坯的正向模拟最后存在一个大的飞边，调整在飞边区域附近的模具片段的接触时间，来满足所需要的飞边尺寸。定义最外面在飞边槽内的模具片段的接触时间为上模运动到期望最终位置的时间。

由于对已记录的边界条件有所修改，某些工件边界节点有可能在方向模拟时穿透模具。为防止这种现象发生，对于这些节点需要加以额外的接触限制，并且这种限制一直保持直到这些节点没有穿透模具的趋势。对于某一模拟时间步长，在工件几何形状根据已收敛的速度场更新后，通过检查模具的坐标来找到已穿透模具的任何节点，这些节点是典型的自由表面节点，其在前一步已经被分离了。对于这些穿透点重新应用仅允许节点沿着模具表面运动的接触限制。通过重返前一步的开始状态，再次确定工件中的已收敛的速度场，然后更新工件几何形状和模具位置。在当前模拟步完成时，对穿透节点的额外限制立即解除，这种检查过程连续执行，并贯穿整个模拟过程的始终。

在反向模拟中如果根据已修改的记录，一个模具片段规定在当前步与工件分离，那么位于此片段边界节点将与模具分离。在试制坯正向模拟过程中，与变形工件接触的最后模具片段成为反向模拟中首先与工件分离的片段。除了最初的模具片段，所有的片段在方向模拟结束时都成为自由边，其间的工件形状就是预锻件形状。此预锻件形状被定义为一系列连接两个相邻节点的直线。由于黏塑性有限元模拟的逐步运动和逐渐加入的边界条件经常导致预锻件有崎岖的表面，这不是实际的模具设计面，因此，为了模具制造应该从直线和圆弧片段生成一个拟合的光滑曲面，平滑的目标是尽可能使成形的模具边界与初始预锻件形状接近。一般地，平滑量是很小的，这种平滑量或者调整量可以很容易通过手工或利用曲线拟合方法来完成。

2.3.4　计算实例

叶片是火箭、飞机喷气发动机及汽轮发动机中的关键零件，叶片预制坯的形状和尺寸直接影响了叶片质量。采用反向模拟来确定预制坯的形状尺寸，对于叶片的研制生产具有重要意义。叶片属于长轴类锻件，其变形过程可近似看作为平面变形。反向模拟时的摩擦因数取0.3，上模的速度为0.42 mm/s。在反向模拟结束后，根据得到的预制坯料形状做一定的优化修正，然后对其进行正向模拟，以验证反向模拟结果的正确性。

叶片反向模拟过程的形状变化如图3.4-16所示，正向模拟验证的结果如图3.4-17所示。比较形状变化图可见，将反向模拟得到的预制坯代入正向模拟中进行计算，坯料恰好同时充满模具两侧，没有飞边产生，可见反向模拟得到的预制坯形状是很合理的。

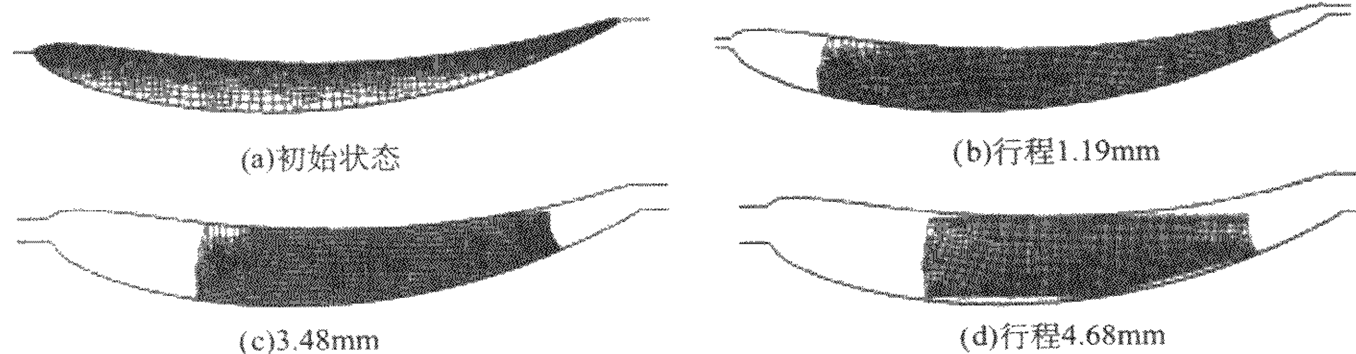


图3.4-16　叶片反向模拟过程工件的形状变化

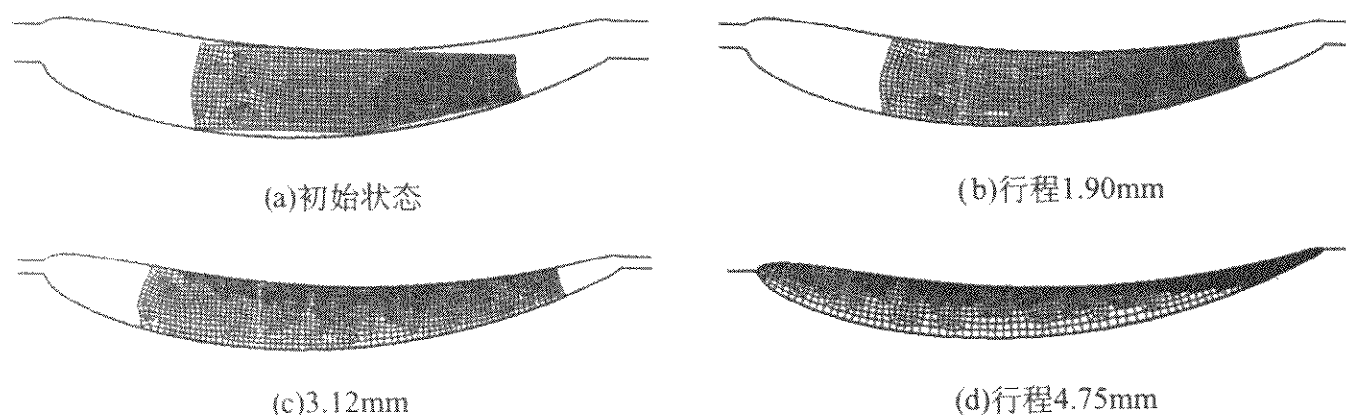


图3.4-17　叶片正向模拟验证过程工件的形状变化

3　无网格法

3.1　概述

自从有限元方法提出至今，国内外对其的研究已经相当成熟，应用范围也越来越广泛，如正反向模拟、热力耦合分析以及微观组织的计算等。但是，应用有限元方法对材料成形过程进行数值模拟时，网格的畸变导致的网格重划分问题是无法回避的（尤其在模拟裂纹的扩展等问题时）。由于网格的重划分会影响到计算精度并浪费计算时间，尤其是三维体积成形时网格重划分技术的不完善，使得人们开始转向另一种数值计算方法，即无网格方法。

自从1977年Lucy L. A. 和Monaghan J. J. 首次提出基于拉格朗日公式的光滑质点流体力学无网格数值计算方法以来，无网格方法已经被广泛用于解决二维和三维问题、静力和动力问题、线性和非线性问题等。

无网格法将连续体离散为独立的节点，无需将节点连接成为单元，这样可以完全避免网格的生成和重化，位移场函数通过节点的离散几何插值得到，可以保证基本场变量在连续体内的连续性，因而提高了计算精度和解题效率。在处理裂纹扩展、界面移动及其他大变形问题上具有有限元法无法比拟的优越性。

3.2　无网格法基本原理

无网格法的基本点是离散化材料为节点，并消除网格单元。这一方法又可以归纳为以下5大类。①光滑质点流体动力学方法（smoothed particle hydrodynamics，SPH），最初这一方法主要应用于计算大量粒子的相互作用，后来这一方法也开始成功地应用到连续体中。②广义有限差分法（generalized finite difference method），这一方法是Liszka和Orkisz于1980年提出的，主要用来处理任意排列的节点问题。③整体剖分方法（partition of unity method），Melenk于1996年提出了整体有

限元剖分方法（partition of unity finite element method，PUFEM），同年 Duarte 和 Oden 介绍了一种更通用的 hp - cloud 方法，其根本特点就是节点形函数与增强函数求乘积。④再生核质点方法（reproducing kernel particle method，RKPM），该方法是 Liu Jun 与 Zhang 于 1995 年根据小波原理提出的，采用窗口函数和傅立叶变换建立了新的形函数，可应用于弹性、塑性和动力学问题。⑤移动最小二乘近似（moving least squares approximations，MLS），这一方法最早由 Lancaster 和 Salkaukas 于 1981 年提出，直到 1992 年 Nayroles、Touzot 和 Villon 将 MLS 应用于伽辽金方法，并称之为扩散单元法（diffuseelement method）。随后，Belytschko、Lu 和 Gu 于 1994 年对该方法进行了改进，并称为无网格伽辽金方法（element - free Galerkin，EFG）。通过应用再生核公式，Liu、Li 和 Belytschko 于 1996 年提出了移动最小二乘再生方法。

3.2.1 移动最小二乘近似

在空间域 Ω 中相对于空间坐标 x 的独立变量 $u(x)$，其近似量为 $\tilde{u}(X)$。假定域 Ω 中任意点 x 可以用参数 u_1，u_2，…，u_n 来表示，相应的节点位置为 x_1，x_2，…，x_n，则移动最小二乘近似方法包含四个关键问题，即权函数 $w_t(x)$、基函数 $p^T(x)$、系数 $a^T(x)$ 和形函数 $\phi_t(x)$。

1）权函数 $w_t(x)$，与每个节点 x_t 相关，且仅在节点 x_t 的一个很小的邻域（或称影响域）内为非零值。一般地，权函数在域 Ω 内的积分为 1，且应为单调递减函数。为构造权函数，对于每个节点 x_t 的影响域大小 φ_{mt} 可以定义为

$$\varphi_{mt} = \varphi_{max}\alpha_t \tag{3.4-170}$$

式中，φ_{max} 为比例参数，对于静态分析可以取值 2.0 ~ 4.0，对于动态分析可以取值 2.0 ~ 2.5。α_t 足够大，这样每个节点的影响域都大到至少可以影响两个相邻的节点。

另外为方便后面分析，定义 $r = \dfrac{\varphi_t}{\varphi_{mt}}$，而 $\varphi_t = |x - x_t|$。

常用的权函数有指数函数、高斯函数、三次样条和四次样条函数。下面给出在一维问题中常用的权函数。

① 指数函数

$$w_t(x) = w(r) = \begin{cases} e^{-(r/\beta)^2} & r \leqslant 1 \\ 0 & r > 1 \end{cases} \tag{3.4-171}$$

② 三次样条函数

$$w_t(x) = w(r) = \begin{cases} \dfrac{2}{3} - 4r^2 + 4r^3 & r \leqslant \dfrac{1}{2} \\ \dfrac{4}{3} - 4r + 4r^2 - \dfrac{4}{3}r^3 & \dfrac{1}{2} < r \leqslant 1 \\ 0 & r > 1 \end{cases} \tag{3.4-172}$$

③ 四次样条函数

$$w_t(x) = w(r) = \begin{cases} 1 - 6r^2 + 8r^3 - 3r^4 & r \leqslant 1 \\ 0 & r > 1 \end{cases} \tag{3.4-173}$$

对于二维问题，影响域的形式有两种表述方式，即圆形和方形区域。对于圆形域，权函数的形式不变，但此时的 r 定义为

$$r = \sqrt{(x - x_t)^2 + (y - y_t)^2} \tag{3.4-174}$$

对于方形域，权函数采用如下形式

$$w_t(x) = w(r_x)w(r_y) \tag{3.4-175}$$

式中

$$r_x = \frac{|x - x_t|}{\varphi_{max}\alpha_{tx}} \qquad r_y = \frac{|y - y_t|}{\varphi_{max}\alpha_{ty}} \tag{3.4-176}$$

2）基函数 $p^T(x) = [p_1(x), p_2(x), \cdots, p_k(x)]$，其中 $p_i(x)$ 为单项式函数，k 是项数。通常使用的基函数可以参见表 3.4-1。

表 3.4-1 移动最小二乘近似中通常使用的基函数

基函数	一维	二维
一次	$p^T(x) = [1 \quad x]$	$p^T(x) = [1 \quad x \quad y]$
二次	$p^T(x) = [1 \quad x \quad x^2]$	$p^T(x) = [1 \quad x \quad y \quad x^2 \quad xy \quad y^2]$
三次	$p^T(x) = [1 \quad x \quad x^2 \quad x^3]$	$p^T(x) = [1 \quad x \quad y \quad x^2 \quad xy \quad y^2 \quad x^3 \quad x^2y \quad xy^2 \quad y^3]$

3）系数 $a^T(x) = [a_1(x), a_2(x), \cdots, a_k(x)]$，由加权最小二乘拟和得到。

函数 $u(x)$ 的移动最小二乘近似函数 $\tilde{u}(x)$ 可写为

$$\tilde{u}(x) = \sum_{i=1}^{k} p_i(x)a_i(x) = p^T(x)a(x) \tag{3.4-177}$$

设定

$$I = \sum_{t=1}^{n} w_t(x)[\tilde{u}(x, x_t) - u_t]^2 \tag{3.4-178}$$

式中，n 为节点数，u_t 为节点 x_t 处的自由度。

式（3.4-178）也可以表示为

$$I = [Pa(x) - u]^T W(x)[Pa(x) - u] \tag{3.4-179}$$

式中

$$u = [u_1 \quad u_2 \quad \cdots \quad u_n]^T \tag{3.4-180a}$$

$$P = \begin{bmatrix} p_1(x_1) & p_2(x_1) & \cdots & p_k(x_1) \\ p_1(x_2) & p_2(x_2) & \cdots & p_k(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ p_1(x_n) & p_2(x_n) & & p_k(x_n) \end{bmatrix} \tag{3.4-180b}$$

$$W(x) = \begin{bmatrix} w(x - x_1) & 0 & \cdots & 0 \\ 0 & w(x - x_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & w(x - x_n) \end{bmatrix} \tag{3.4-180c}$$

那么系数 a 可以通过求 I 的最小值得到，即

$$\frac{\partial I}{\partial a} = A(x)a(x) - B(x)u = 0 \tag{3.4-181}$$

式中

$$A(x) = P^T W(x) P = \sum_{t=1}^{n} w(x - x_t)p(x_t)p^T(x_t) \tag{3.4-182}$$

$$B(x) = P^T W(x) = [w(x - x_1)p(x_1)w(x - x_2)p(x_2)\cdots w(x - x_n)p(x_n)] \tag{3.4-183}$$

根据式（3.4-181），得到

$$a(x) = A^{-1}(x)B(x)u \tag{3.4-184}$$

4）形函数 $\phi_t(x)$

根据式（3.4-177）和式（3.4-184），得到

$$\tilde{u}(x) = \sum_{t=1}^{n} \phi_t(x)u_t = \Phi(x)u \tag{3.4-185}$$

式中，形函数 $\phi_t(x)$ 定义为

$$\phi_t(x) = [p^T(x)A(x)B(x)]_t \tag{3.4-186}$$

则形函数的偏导数可写为

$$\phi_{t,i} = [p^T_{,i}A^{-1}B + p^T(A^{-1})_{,i}B + p^T A^{-1}(B)_{,i}]_t \tag{3.4-187}$$

3.2.2 离散化

1）伽辽金方法 根据区域 Ω 内的偏微分方程和边界条件

$$Lu(x)=f \quad (3.4\text{-}188)$$

$$\nabla^2 u=f \quad (3.4\text{-}189)$$

$$在\ \Gamma_u\ 上，\ u=\bar{u} \quad (3.4\text{-}190)$$

$$在\ \Gamma_q\ 上，\ u_{,n}=\bar{u}_{,u} \quad (3.4\text{-}191)$$

式中，$\Gamma=\Gamma_u\cup\Gamma_q$ 为区域 Ω 的边界，$q=\frac{\partial u}{\partial n}$为函数 u 的法向导数，n 为边界上任一点法向向量。

根据变分原理，与微分方程（3.4-188）相应的变分弱形式为

$$a[\delta u(x),u(x)]=[\delta u(x),f] \quad (3.4\text{-}192)$$

式中，$\delta u(x)$为测试函数；u 为试函数。对于方程（3.4-188）～式（3.4-191），

$$a(\delta u,u)=\int_\Omega \delta u_{,i}u_{,i}\mathrm{d}\Omega \quad (3.4\text{-}193a)$$

$$(\delta u,f)=\int_\Omega \delta uf\mathrm{d}\Omega-\int_{\Gamma_q}\delta u\bar{u}_{,n}\mathrm{d}\Gamma \quad (3.4\text{-}193b)$$

式中，在 Γ_u 上，$\delta u=0$；$u(x)$和$\delta u(x)$应满足 C^0 的要求。

将式（3.4-185）所示的移动最小二乘近似函数代入式（3.4-193），得离散方程为

$$\sum_J\int_\Omega \Phi_{J,i}\Phi_{J,i}\mathrm{d}\Omega u_J=f_I \quad (3.4\text{-}194)$$

式中

$$f_J=\int_\Omega \Phi_J f\mathrm{d}\Omega-\int_{\Gamma_q}\Phi_J u_{,n}f\mathrm{d}\Gamma \quad (3.4\text{-}195)$$

2）积分实现　在无网格伽辽金方法中，计算式（3.4-194）和式（3.4-195）的积分通常采用下列三种方式。

① 点积分，也就是通过节点值实现积分。

$$\int_\Omega F(x)\mathrm{d}\Omega=\sum_{I=1}^{N_P}F(X_I)\Delta V \quad (3.4\text{-}196)$$

② 网格积分，即按照规则分割出的独立于结构的矩形背景网格，形成积分单元。背景网格中的每一个单元，称为网格结构。

③ 采用有限元网格作为背景积分网格。

第①种积分方案类似于配点法，速度最快，但是有可能不稳定。通过在能量泛函中增加稳定项，选择稳定系数，可以得到相应的控制方程，但是缺点是计算过程较为复杂。第②、③两种积分方案则没有彻底地抛弃网格，都要利用背景网格来完成积分。不过背景网格的利用仅仅是为了完成区域和边界积分，并不影响无网格方法的本质。实际计算表明，第②种积分方案积分网格内部包含的不连续性对结果的影响是很小的。第③种积分方案适用于有限元与无网格方法的耦合计算。

3.2.3 边界条件的处理

移动最小二乘形函数不具有插值特征，目前本质边界条件的处理方法主要有：拉格朗日乘子法、修正的变分原理、耦合的有限元-无网格伽辽金法、修正的配点法、奇异权函数法、罚函数法和位移约束方程方法等。下面介绍拉格朗日乘子法和罚函数法。

1）拉格朗日乘子法　分析二维弹性力学边值问题，控制方程为

$$\nabla\cdot\boldsymbol{\sigma}+\boldsymbol{b}=0\quad 在\ \Omega\ 内 \quad (3.4\text{-}197a)$$

边界条件为

$$\boldsymbol{\sigma}\cdot\boldsymbol{n}=\bar{\boldsymbol{t}}\quad 在边界\ \Gamma_t\ 上 \quad (3.4\text{-}197b)$$

$$\boldsymbol{u}=\bar{\boldsymbol{u}}\quad 在边界\ \Gamma_u\ 上 \quad (3.4\text{-}197c)$$

式中，$\boldsymbol{\sigma}$ 为应力张量，$\boldsymbol{b}$ 为体力向量，$\boldsymbol{u}$ 为位移场向量，$\boldsymbol{n}$ 为边界上的单位法向向量，Γ_t 和 Γ_u 分别为区域 Ω 的应力边界和位移边界。

设位移场向量 $\boldsymbol{u}\in H^1$，拉格朗日乘子 $\boldsymbol{\lambda}\in H^0$，试函数 $\delta\boldsymbol{v}\in H^1$，$\delta\boldsymbol{\lambda}\in H^0$，则与方程（3.4-197）相对应的变分弱形式为

$$\int_\Omega \nabla_s\delta\boldsymbol{v}\cdot\boldsymbol{\sigma}\mathrm{d}\Omega-\int_\Omega\delta\boldsymbol{v}\cdot\boldsymbol{b}\mathrm{d}\Omega-\int_{\Gamma_t}\delta\boldsymbol{v}\cdot\bar{\boldsymbol{t}}\mathrm{d}\Gamma-\int_{\Gamma_t}[\delta\boldsymbol{\lambda}\cdot(\boldsymbol{u}-\bar{\boldsymbol{u}})+\delta\boldsymbol{v}\cdot\boldsymbol{\lambda}]\mathrm{d}\Gamma=0 \quad (3.4\text{-}198)$$

式中，$\nabla_s\boldsymbol{v}$ 为$\nabla\boldsymbol{v}$ 的对称部分，H^0 和 H^1 分别表示零阶和一阶索伯烈夫（Sobolev）空间。这里是通过拉格朗日乘子施加本质边界条件。拉格朗日乘子为

$$\lambda(X)=N_I(s)\lambda_I\quad (X\in\Gamma_u) \quad (3.4\text{-}199)$$

式中，$N_I(s)$为拉格朗日插值，s 为沿边界的弧长。

采用式（3.4-185）的形式构造近似函数 u 和试函数 δu，代入方程（3.4-198），并采用伽辽金法，经整理，得离散方程

$$\begin{bmatrix}K & G\\ G^{\mathrm{T}} & 0\end{bmatrix}\begin{Bmatrix}u\\ \lambda\end{Bmatrix}=\begin{Bmatrix}f\\ q\end{Bmatrix} \quad (3.4\text{-}200)$$

式中

$$K_{IJ}=\int_\Omega \boldsymbol{B}_I^{\mathrm{T}}\boldsymbol{D}\boldsymbol{B}_J\mathrm{d}\Omega$$

$$G_{IK}=-\int_{\Gamma_u}\boldsymbol{\Phi}_I\overline{\boldsymbol{N}}_K\mathrm{d}\Gamma$$

$$f_I=\int_{\Gamma_t}\boldsymbol{\Phi}_I\bar{\boldsymbol{t}}\mathrm{d}\Gamma+\int_\Omega\boldsymbol{\Phi}_I\boldsymbol{b}\mathrm{d}\Omega$$

$$q_K=-\int_{\Gamma_u}\overline{N}_K\bar{u}$$

$$\boldsymbol{B}_I(X)=\begin{bmatrix}\Phi_{I,x}(X) & 0\\ 0 & \Phi_{I,y}(X)\\ \Phi_{I,y}(X) & \Phi_{I,x}(X)\end{bmatrix}$$

$$\overline{\boldsymbol{N}}_K=\begin{bmatrix}N_K & 0\\ 0 & N_K\end{bmatrix}$$

$$\boldsymbol{D}=\frac{E_1}{1-\mu_1^2}\begin{bmatrix}1 & \mu_1 & 0\\ \mu_1 & 1 & 0\\ 0 & 0 & \frac{1-\mu_1}{2}\end{bmatrix}$$

对于平面应力问题，$E_1=E$，$\mu_1=\mu$，对于平面应变问题 $E_1=\frac{E}{1-\mu^2}$，$\mu_1=\frac{\mu}{1-\mu}$，E 和 μ 分别为弹性模量和泊松比。

2）罚函数法　罚函数法是通过引入罚参数 α，已满足本质边界条件。与方程（3.4-197）相对应的泛函为

$$\Pi=\int_\Omega\frac{1}{2}\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\sigma}\mathrm{d}\Omega-\int_\Omega\boldsymbol{u}^{\mathrm{T}}\boldsymbol{b}\mathrm{d}\Omega-\int_{\Gamma_t}\boldsymbol{u}^{\mathrm{T}}\bar{\boldsymbol{t}}\mathrm{d}\Gamma+\int_{\Gamma_u}\frac{1}{2}\alpha(\boldsymbol{u}-\bar{\boldsymbol{u}})^2\mathrm{d}\Gamma \quad (3.4\text{-}201)$$

式中，$\boldsymbol{\varepsilon}$ 为应变张量，$\boldsymbol{\sigma}$ 为应力张量。与方程（3.4-201）相对应的离散形式为

$$\boldsymbol{K}^*\boldsymbol{u}=(\boldsymbol{K}_1+\boldsymbol{K}_2)\boldsymbol{u}^*=\boldsymbol{f} \quad (3.4\text{-}202)$$

式中，$\boldsymbol{K}$，$\boldsymbol{f}$ 由下列子矩阵 $\boldsymbol{K}_{IJ}$和子向量$\boldsymbol{f}_I$ 组成

$$\boldsymbol{K}_{IJ}=\int_\Omega\boldsymbol{B}_I^{\mathrm{T}}\boldsymbol{D}\boldsymbol{B}_J\mathrm{d}\Omega+\alpha\int_{\Gamma_u}\boldsymbol{\Phi}_I\boldsymbol{S}\boldsymbol{\Phi}_J\mathrm{d}\Gamma \quad (3.4\text{-}203)$$

$$\boldsymbol{f}_I=\int_\Omega\boldsymbol{\Phi}_I\boldsymbol{b}\mathrm{d}\Omega+\int_{\Gamma_t}\boldsymbol{\Phi}_I\bar{\boldsymbol{t}}\mathrm{d}\Gamma+\alpha\int_{\Gamma_u}\bar{\boldsymbol{u}}\boldsymbol{\Phi}_I\boldsymbol{S}\bar{\boldsymbol{u}}\mathrm{d}\Gamma \quad (3.4\text{-}204)$$

式中，$\boldsymbol{S}=\begin{bmatrix}s_x & 0\\ 0 & s_y\end{bmatrix}$；$s_i=\begin{cases}1 & 当\ \Gamma_u\ 上有\ u_i\ 位移约束\\ 0 & 当\ \Gamma_u\ 上无\ u_i\ 位移约束\end{cases}$。

3.3 有限元与无网格耦合

有限元与无网格耦合方法提出的目的是尽可能地减小固体

和结构有限元分析过程中的网格畸变问题，并减少计算时间。

有限元与无网格耦合分析模型见图3.4-18，在材料变形复杂的区域运用无网格方法，而在其他区域则采用有限元方法计算。下面简要介绍有限元与无网格耦合的方法。

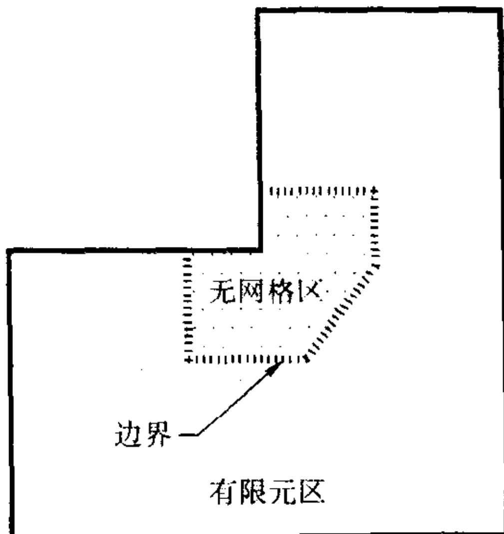


图3.4-18　有限元与无网格耦合分析模型示意图

首先，构造了基于移动最小二乘的无网格近似方法。无网格 Galerkin 方法和再生核质方法是两种有代表性的移动最小二乘方法。下面按照再生核质方法推导无网格近似方法。

在 x 点的离散近似解可以写为

$$u_i(x) \approx u_i^h(x) = \begin{cases} \sum\limits_{\substack{L \\ x_L \in \Omega_1}}^{KP} \Phi_L^{[m]}(x; x - x_L) d_{iL} & \forall x \in \Omega_1 \\ \sum\limits_{\substack{I \\ x_L \in \Omega_2}}^{NP} \overline{w}_a^{[n]}(x; x - x_I) d_{iI} + \sum\limits_{\substack{L \\ x_L \in \Gamma_3}}^{MP} \Phi_L^{[m]}(x; x - x_L) d_{iL} & \forall x \in \Omega_2 \end{cases} \tag{3.4-205}$$

式中，Ω_1 为有限元计算子域；Ω_2 为无网格计算子域；全部区域 $\Omega = \Omega_1 \cup \Omega_2$；有限元域与无网格域之间的接触边界 $\Gamma_3 = \Omega_1 \cap \Omega_2$；$\overline{w}_a^{[n]}$ 为 $[n]$ 阶再生核函数；a 是核尺寸；$\Phi_L^{[m]}$ 为带插值 $[m]$ 阶的有限元形函数；NP 是无网格质点总数；KP 是有限元中每个单元的节点数；d_{iL} 是近似系数；MP 是接触界面上总的有限元节点数。

再生条件为

$$\sum_{\substack{I \\ x_I \in \Omega_{\text{Meshfree}}}}^{NP} \overline{w}_a^{[n]}(x; x - x_I) x_{1I}^i x_{2I}^j + \sum_{\substack{J \\ x_J \in \Gamma_{\text{Meshfree}}}}^{MP} \Phi_J^{[m]}(x; x - x_J) x_{1J}^i x_{2J}^j = x_1^i x_2^j \ (i, j = 0, 1, \cdots, n) \tag{3.4-206}$$

通常，有限元插值 $[m]$ 阶数选择与再生函数阶数 $[n]$ 相同，近似函数在 Ω_2 域内连续。如果再生函数阶数 $[n]$ 大于有限元插值阶数 $[m]$，即 $n > m$，则近似函数仅在 Ω_2/Γ_3 上连续。

采用上述的再生条件，得到有限元与无网格方法耦合的近似解。

$$u_i^h(x) = \sum_{\substack{I \\ x_I \in \Omega_2}}^{NP} \tilde{\Psi}_I(x) d_{iI} + \sum_{\substack{J \\ x_J \in \Gamma_3}}^{MP} \Phi_J^{[m]}(x) d_{iJ} = \sum_{\substack{I \\ x_I \in \Omega_2}}^{NP} \hat{\Psi}_I(x) d_{iI} \tag{3.4-207}$$

式中

$$\tilde{\Psi}_I(x) = \sum_{\substack{I \\ x_I \in \Omega_2}}^{NP} \{ \boldsymbol{H}^{[n]^{\mathrm{T}}}(0) \boldsymbol{M}^{[n]^{-1}}(x) \boldsymbol{H}^{[n]}(x - x_I) - \sum_{\substack{J \\ x_J \in \Gamma_3}}^{MP} \boldsymbol{H}^{[n]^{\mathrm{T}}}(x - x_J) \boldsymbol{M}^{[n]^{-1}}(x) \boldsymbol{H}^{[n]}(x - x_I) \Phi_J^{[m]}(x; x - x_J) \} w_a(x - x_I) d_{iI} \tag{3.4-208}$$

式中，$\tilde{\Psi}_I$ 是相应于无网格节点 I 的再生核函数；$\Phi_J^{[m]}$ 是相应于有限元节点 J 的标准有限元形函数。$\hat{\Psi}_I(x)$ 是节点 I 的修正的有限元与无网格耦合形函数。

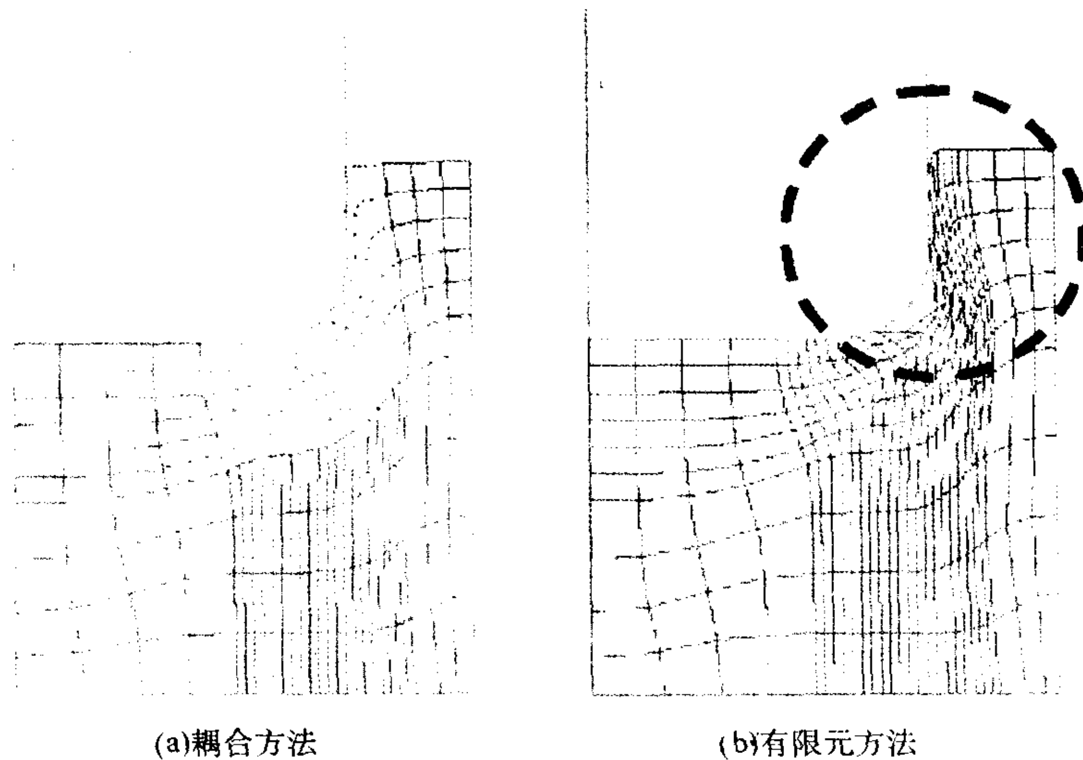
另外，对于式（3.4-208），还应满足界面约束条件

$$\tilde{\Psi}_I(x) = 0，对所有节点 \{I：假定(\Psi_I) \cap \Gamma_3 \neq 0\}，且 x \in \Gamma_3 \tag{3.4-209}$$

3.4　计算实例

本例给出一个采用有限元与无网格方法耦合的计算实例，同时给出了仅采用有限元方法计算的对比结果。

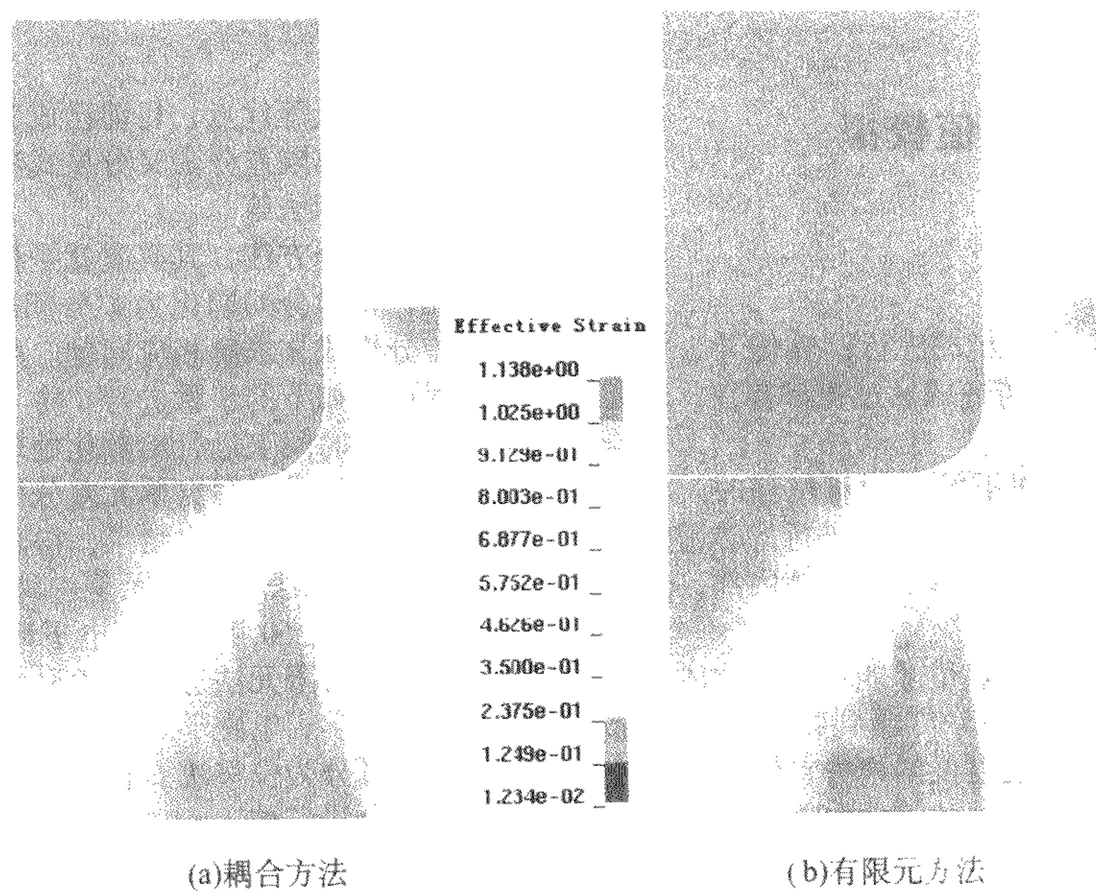
实例模拟了采用圆柱形刚冲头挤压杯形零件的过程，模拟模型中工件与刚冲头之间的摩擦因数取0.2，冲头挤压速率为0.25 mm/s。在耦合分析模型中，材料变化剧烈处，即冲头圆角附近区域采用无网格计算，其余部分采用有限元网格计算。计算结果如图3.4-19和3.4-20所示。



(a)耦合方法　　(b)有限元方法

图3.4-19　变形后的网格变化比较

在采用有限元方法计算时，由于在凸模圆角区域产生了严重的网格畸变而导致计算不能收敛而结束（见图 3.4-19b）。而采用耦合的方法则有效地避免了网格的畸变，很好地完成变形过程的计算。比较两种方法计算的结果，可以发现有限元与无网格耦合方法具有较低且更为均匀的应变分布。



(a)耦合方法　(b)有限元方法

图 3.4-20　变形后的等效应变比较

编写：王新云（华中科技大学）

第5章 铸造成形数值模拟方法

1 铸造凝固过程温度场数值模拟

1.1 概述

1.1.1 传热的基本方式

热量传递方式可以通过三种方式进行：热传导，热对流，热辐射。这三种传热方式在铸造成形过程都存在，下面先简要介绍一些基本概念。

1）热传导　物体各部分之间不发生相对位移时，依靠分子、原子及自由电子等微观粒子的热运动进行的热量传递称为热传导，简称导热。在紧密的不透明的物体内部，热量只能依靠导热方式传递。

只有在物体处于不同温度时，热量才能从一个物体传递到另一个物体，或从物体的某一部分传递到物体的另一部分。热总是从温度高的地方流向温度低的地方，铸件凝固冷却时，铸件内部的温度高于外界，因此铸件内部向其外侧以及铸型传递热量。

在三维迪卡尔坐标系统，连续介质各点在同一时刻的温度分布叫做温度场，温度场一般可表达为$T=f(x,y,z,t)$。若温度场不随时间变化，则称做稳定温度场，由此产生的导热为稳定导热；若温度场随时间改变，则称做不稳定温度场，不稳定温度场的导热为不稳定导热。

导热的基本定律是Fourier定律，Fourier定律的具体内容我们在后面再阐述。

2）热对流　热对流是指流体中温度不同的各部分相互混合的宏观运动引起热量传递的现象。热对流总与流体的导热同时发生，可以看作是流体流动时的导热。对流换热的情况比只有热传导的情况复杂。对流换热可以用Newton冷却定律来描述，即

$$q=\alpha(T_f-T_w) \tag{3.5-1}$$

式中，q为热流密度；α为表面传热系数；T_f为流体的特征温度；T_w为固体边界温度。

对流换热按引起流动运动的不同原因可分为自然对流和强制对流两大类。自然对流是由于流体冷、热部分的密度不同而引起的，如暖气片表面附近热空气向上流动就是自然对流。如果流体的流动是由于水泵或其他压差所造成的，则称为强制对流。

3）热辐射　物体通过电磁波传递能量的方式称为辐射。物体会因各种原因发出辐射能，其中因热的原因发出辐射能的现象称为热辐射。自然界中各个物体都不停地向空间发出热辐射，同时又不断地吸收其他物体发出的热辐射。发出与吸收过程的综合效果造成了物体间以辐射方式进行了热量传递。辐射换热可以用Stefen－Boltsman定律来描述，即

$$q=\varepsilon\varepsilon_0 T_s^4 \tag{3.5-2}$$

式中，q为热流密度；T_s为表面的绝对温度；ε为辐射黑度；ε_0为Stefen－Boltsman常数。

1.1.2 传热分析的常用数值方法

采用什么样的数值方法是进行传热分析的一个重要问题。目前比较成熟的数值方法有：有限差分法（FDM）、有限元法（FEM）、直接差分法（DFM）、边界元法（BEM）。下面分别加以介绍。

1）有限差分法　有限差分法，又称泰勒展开差分法，是最早用于传热的计算方法。该方法具有差分公式导出简单和计算成本低等优点，目前已成为应用最为广泛的一种数值分析方法。有限差分法又分显式法、隐式法、交替方向隐式法、控制体积法等。

有限差分方法，其实质就是将求解区域划分为有限个网格单元，将微分问题化为差分问题，离散化得到差分格式，利用差分格式来求解相应问题。用有限差分来求解不稳定导热过程可按如下的步骤进行：将不稳定导热过程所涉及的区域在空间和时间上进行离散化处理；物性条件、初始条件和边界条件的设定；写出单元差分格式；将求解的过程编成计算程序，由计算机算出结果，得到温度场相关结果。

2）有限元法　有限元是一种被广泛使用的数值计算方法，从20世纪60年代起与电子计算机的应用一起获得了飞速的发展。有限元法适用的学科和领域十分广泛，如固体力学、流体力学、电磁学、传热学等都广泛使用有限元法。有限元法又可分为位移法、利用余位进行变化的方法和用混合积分的混合法三种。

有限元法的位移法，其实质就是将求解区域划分为有限个单元，通过构造插值函数，把问题化为一个变分问题（即求泛函数值的问题），经过离散化得到计算格式，利用计算格式来求解相应问题。变分法证明：求解某些微分方程的问题等效于将泛函数的相关量进行最小化。如果相关于因变量的节点值使泛函数最小，那么所得到的条件表达式就是所需要的离散化方程。也就是说，求解一个微分方程边值问题就可以通过寻找某一变分问题的极值函数来解决。有限元法求解导热问题正是利用微分方程边值问题等价于相应变分问题这一特点的。

用有限元法求解不稳定导热过程可归纳为如下的步骤：将不稳定导热过程所涉及的区域在空间和时间上进行离散化处理；物性条件、初始条件和边界条件的设定；写出单元泛函数表达式；构造每个单元的插值函数；求得泛函数极值条件的代数方程表达式；构造代数方程组；将求解的过程编成计算程序，由计算机算出结果，得到温度场相关结果。

有限元法具有很大的灵活性和适应性，并且便于与应力场进行耦合计算，因而受到人们的重视。然而有限元法存在着数据准备复杂、对硬件设备要求高等缺点，这就影响了它的应用，尤其是在微机上的应用。

3）直接差分法　历来采用的数值解析法几乎都是以微分方程式为基础。而直接差分法则不然，直接差分法是将欲求解的系统分割为许多微小单元，各单元的物理现象不是通过微分方程式表示，而是直接表示为计算机可以进行计算的差分方程式，之后求解差分方程式，得出数值解。直接差分法中，节点以及节点所代表的节点领域的概念很重要，因为温度和固相率是因节点的值而离散的，而且节点领域内的各种守恒定律也是用数学式来表示的缘故。根据节点和节点领域的定义方法，直接差分法大致可以分为内节点法和外节点法两大类。

① 内节点法　把分割系统后所得到的微小单元本身定义为节点领域，而把单元的外心定义为节点。该方法中要避免单元的形状不同，外心有可能不在节点领域内的情况。

② 外节点法　把分割单元的顶点定义为节点，由各单元中各边的垂直平分线（三维空间是垂直平分面）构成的新领域被定义为节点领域。

将直接差分法和有限差分法进行比较，DFM优点为：a）能够使用三角形和四边形（在三维场合为四、五、六面体等

各种单元，所以能够解复杂形状的问题；b）对每个单元能够指定其物性值，所以能够容易地解由多种物质组成的系统的问题。DFM缺点为：a）输入数据量大；b）程序复杂；c）计算时间稍长。

4）边界元法 边界元法在原理上不同于上述几种方法。用边界元法处理恒定问题时只需对边界进行分割即可，非常方便。即使对于非稳定问题，也具有受分割单元的形状和大小限制小的优点。但是采用边界元法，其计算公式的推导和程序实现很复杂，而且必须解联立方程式；另外，如果物性变化的话，则难以考虑对流。

1.1.3 本节的思路与主要内容

本节的思路是采用传热分析中最常用的有限差分法对铸造凝固过程的数学模型进行离散。将重点介绍以下内容：铸造凝固过程的数学模型、基于有限差分的离散格式、温度场的初始条件、边界条件、潜热处理方法以及温度场数值模拟流程图等。

1.2 数学模型

1.2.1 傅立叶定律

在大量实验基础上，J.B.Fourier于1882年指出，单位时间内由热传导而通过单位面积的热量（比热流量）与温度梯度成正比。

在一维空间，Fourier定律表示成下式

$$\dot{q} = -\lambda \frac{\partial T}{\partial x} \tag{3.5-3}$$

式中，$\dot{q}$ 为比热流量，W/m²；T 为温度，K；x 为坐标值，m；$\frac{\partial T}{\partial x}$为温度梯度，K/m；$\lambda$ 为导热系数，W/（m·K）。

负号表明导热的方向永远沿着温度降低的方向，即导热热流从高温区流向低温区。

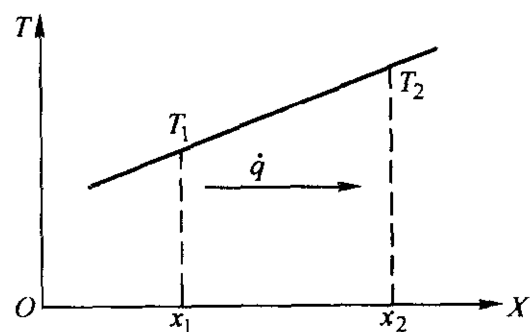


图3.5-1 一维空间的热传导传热

如图3.5-1，当 x 方向的温度分布呈线性时，温度梯度 $\partial T/\partial x$ 为：

$$\frac{\partial T}{\partial x} = \frac{T_2 - T_1}{x_2 - x_1}$$

根据式（3.5-3），沿方向 x 产生的比热流量为：

$$\dot{q} = -\lambda \frac{T_2 - T_1}{x_2 - x_1} \tag{3.5-4}$$

1.2.2 热传导微分方程（直角坐标系）

如图3.5-2所示，假设微元体 x、y、z 三个方向上的尺寸为 dx、dy、dz，显然微元体的体积为 dxdydz。

根据Fourier定律，流入此微元体的热量为：

$$\begin{aligned} Q_x &= -\lambda \frac{\partial T}{\partial x} \mathrm{d}y\mathrm{d}z \\ Q_y &= -\lambda \frac{\partial T}{\partial y} \mathrm{d}x\mathrm{d}z \\ Q_z &= -\lambda \frac{\partial T}{\partial z} \mathrm{d}x\mathrm{d}y \end{aligned} \tag{3.5-5}$$

而流出此微元体的热量为：

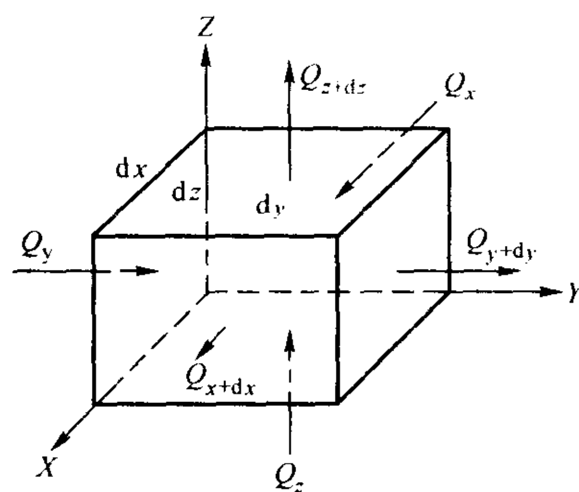


图3.5-2 三维空间微元体热平衡图

$$\begin{aligned} Q_{x+\mathrm{d}x} &= -\lambda \frac{\partial}{\partial x}\left(T + \frac{\partial}{\partial x}\mathrm{d}x\right)\mathrm{d}y\mathrm{d}z \\ Q_{y+\mathrm{d}y} &= -\lambda \frac{\partial}{\partial y}\left(T + \frac{\partial}{\partial y}\mathrm{d}y\right)\mathrm{d}x\mathrm{d}z \\ Q_{z+\mathrm{d}z} &= -\lambda \frac{\partial}{\partial z}\left(T + \frac{\partial}{\partial z}\mathrm{d}z\right)\mathrm{d}x\mathrm{d}y \end{aligned} \tag{3.5-6}$$

如物体中无内热源，根据能量守恒定律：流入热量－流出热量＝微元体内蓄热量的增加，即：

$$Q_{入} - Q_{出} = \Delta Q \tag{3.5-7}$$

而单位时间内微元体蓄热量增量：

$$\Delta Q = \rho c_p \frac{\partial T}{\partial t}\mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{3.5-8}$$

将上述式（3.5-5）、式（3.5-6）、式（3.5-7）代入式（3.5-8），整理得：

$$\rho c_p \frac{\partial T}{\partial t} = \lambda\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) \tag{3.5-9}$$

式中，ρ，c_p，λ 为常数。

令$\frac{\lambda}{\rho c_p} = \alpha$，则式（3.5-9）变为：

$$\begin{aligned} \frac{\partial T}{\partial t} &= \alpha\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) \\ &= \alpha \nabla^2 T \end{aligned} \tag{3.5-10}$$

式中，∇^2为拉普拉斯运算符号（算子）；α 为导温系数，m²/s。

该方程的物理意义：

1）当$\nabla^2 T > 0$时，$\frac{\partial T}{\partial t} > 0$，物体被加热；

2）当$\nabla^2 T = 0$时，$\frac{\partial T}{\partial t} = 0$，稳定温度场；

3）当$\nabla^2 T < 0$时，$\frac{\partial T}{\partial t} < 0$，物体被冷却。

式（3.5-10）即为三维热传导微分方程，亦即温度场数值模拟的数学模型，式（3.5-11）、式（3.5-12）分别是一维、二维场合下温度场的热传导微分方程。

一维场合：
$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2} \tag{3.5-11}$$

二维场合：
$$\frac{\partial T}{\partial t} = \alpha\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) \tag{3.5-12}$$

1.3 基于有限差分方法的离散

下面将采用有限差分方法来对上述温度场数学模型在时间上和空间上进行离散。首先介绍二维场合下的离散格式，在此基础上介绍三维离散格式。

1.3.1 二维场合的离散格式

在二维情况下，对傅立叶热传导微分方程式（3.5-12）进行基于有限差分法的离散。如图3.5-3所示，单元 i 是一边长为 Δx 的正四边形单元，它与相邻的四个单元进行热量交换。在微小的时间 Δt 内，单元 i 吸收的热量 Q 为：

$$Q=\rho_i c_{pi}(\Delta x)^2(T_i^{t+\Delta t}-T_i^t) \tag{3.5-13}$$

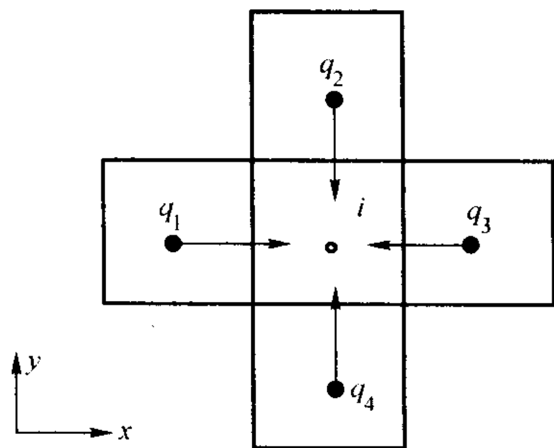


图 3.5-3　二维差分单元 i 的热平衡关系图

相邻的单元 1、2、3、4 单元 i 的热量总和 Q_{sum} 为：

$$Q_{sum}=\sum_{j=1}^{4}\frac{\Delta x}{\dfrac{\Delta x/2}{\lambda_i}+\dfrac{\Delta x/2}{\lambda_j}}(T_j^t-T_i^t)\Delta t \tag{3.5-14}$$

根据能量守恒定律，由式（3.5-13）、式（3.5-14）得：

$$\rho_i c_{pi}(\Delta x)^2(T_i^{t+\Delta t}-T_i^t)=\sum_{j=1}^{4}\frac{\Delta x(T_j^t-T_i^t)\Delta t}{\dfrac{\Delta x}{2\lambda_i}+\dfrac{\Delta x}{2\lambda_j}} \tag{3.5-15}$$

整理式（3.5-15）得：

$$T_i^{t+\Delta t}=T_i^t+\frac{\Delta t}{\rho_i c_{pi}\Delta x}\sum_{j=1}^{4}\frac{T_j^t-T_i^t}{\dfrac{\Delta x}{2\lambda_i}+\dfrac{\Delta x}{2\lambda_j}} \tag{3.5-16}$$

将式（3.5-16）变形得：

$$T_i^{t+\Delta t}=\left[1-\frac{\Delta t}{\rho_i c_{pi}\Delta x}\sum_{j=1}^{4}\frac{1}{\dfrac{\Delta x}{2\lambda_i}+\dfrac{\Delta x}{2\lambda_j}}\right]T_i^t+\frac{\Delta t}{\rho_i c_{pi}\Delta x}\sum_{j=1}^{4}\frac{T_j^t}{\dfrac{\Delta x}{2\lambda_i}+\dfrac{\Delta x}{2\lambda_j}} \tag{3.5-17}$$

由式（3.5-17）知，单元 i 在 $t+\Delta t$ 时刻的温度等于 t 时刻自身温度以及相邻 4 个单元温度的线性组合。显而易见，如果相邻单元温度高或低，单元 i 的温度也相应地大或小；另外从物理含义来说，单元 i 在 t 时刻温度高，则其在 $t+\Delta t$ 时刻的温度也应该高，即等式右边第一项系数必须不小于零，即

$$1-\frac{\Delta t}{\rho_i c_{pi}\Delta x}\times a_i\geqslant 0 \tag{3.5-18}$$

式中

$$a_i=\sum_{j=1}^{4}\frac{1}{\dfrac{\Delta x}{2\lambda_i}+\dfrac{\Delta x}{2\lambda_j}}$$

整理得

$$\Delta t\leqslant(\rho_i c_{pi}\Delta x)/a_i，且\ \Delta t>0 \tag{3.5-19}$$

1.3.2　三维场合的离散格式

在三维场合下，对傅立叶热传导微分方程式（3.5-10）进行基于有限差分法的离散。如图 3.5-4 所示，单元 i 是一边长为 Δx 的正六面体单元，它与相邻的六个单元进行热量交换。在微小的时间 Δt 内，单元 i 吸收的热量 Q 为：

$$Q=\rho_i c_{pi}(\Delta x)^3(T_i^{t+\Delta t}-T_i^t) \tag{3.5-20}$$

从相邻的单元 1、2、3、4、5、6 流入网格 i 的热量总和 Q_{sum} 为：

$$Q_{sum}=\sum_{j=1}^{6}\frac{\Delta x\Delta x}{\dfrac{\Delta x/2}{\lambda_i}+\dfrac{\Delta x/2}{\lambda_j}}(T_j^t-T_i^t)\Delta t \tag{3.5-21}$$

根据能量守恒定律，由式（3.5-20）、式（3.5-21）得：

$$\rho_i c_{pi}(\Delta x)^3(T_i^{t+\Delta t}-T_i^t)=\sum_{j=1}^{6}\frac{\Delta x\Delta x(T_j^t-T_i^t)\Delta t}{\dfrac{\Delta x}{2\lambda_i}+\dfrac{\Delta x}{2\lambda_j}} \tag{3.5-22}$$

整理得单元 i 在 $t+\Delta t$ 时刻的温度计算公式为：

$$T_i^{t+\Delta t}=T_i^t+\frac{\Delta t}{\rho_i c_{pi}\Delta x}\sum_{j=1}^{6}\frac{(T_j^t-T_i^t)}{\dfrac{\Delta x}{2\lambda_i}+\dfrac{\Delta x}{2\lambda_j}} \tag{3.5-23}$$

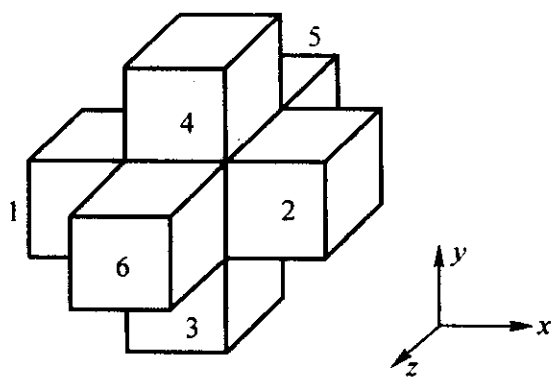


图 3.5-4　三维差分单元 i 的热平衡关系图

与二维情况一样，Δt 必须满足一定条件才能保证数值解的稳定。由式（3.5-23）知，单元 i 在 $t+\Delta t$ 时刻的温度等于 t 时刻自身温度以及相邻 6 个单元温度的线性组合。显而易见，相邻 6 个单元温度的高低，直接影响了单元 i 在 $t+\Delta t$ 时刻温度的大小；同样，单元 i 在 t 时刻温度高，则其在 $t+\Delta t$ 时刻的温度也应该高，即等式右边第一项系数必须不小于零，即

$$1-\frac{\Delta t}{\rho_i c_{pi}\Delta x}\times a_i\geqslant 0 \tag{3.5-24}$$

式中

$$a_i=\sum_{j=1}^{6}\frac{1}{\dfrac{\Delta x}{2\lambda_i}+\dfrac{\Delta x}{2\lambda_j}}$$

整理得

$$\Delta t\leqslant(\rho_i c_{pi}\Delta x)/a_i，且\ \Delta t>0 \tag{3.5-25}$$

式（3.5-25）即为数值解收敛性条件，在实际的程序应用中，对于立方体单元 i 来说时间步长 Δt 满足下式即可。

$$\Delta t\leqslant\rho c_p(\Delta x)^2/(6\lambda)，且\ \Delta t>0 \tag{3.5-26}$$

1.4　初始条件与边界条件

1.4.1　初始条件

从差分方程［参见式（3.5-17）、式（3.5-23）］可以看出要确定各单元在新时刻（$t+\Delta t$）时的温度值，必须首先知道前一时刻（t）时的温度值。因此，初始条件就是要确定 $t=0$ 时刻（开始计算时刻）各单元的温度值。

对于三维温度场 $T^t=f(x,y,z,t)$，初始时刻（$t=0$）的温度场为：

$$T^0=f(x,y,z,0) \tag{3.5-27}$$

在进行初始温度的设置时，可以假设铸件瞬间充型、初温均布，即可以用如下方程来表示。

铸件部分：$T^0_{cast} = f_c(x,y,z,0)$ (3.5-28)

铸型部分：$T^0_{mold} = f_m(x,y,z,0)$ (3.5-29)

1.4.2 边界条件

如图 3.5-5 所示，与边界相接的微元体的热量守恒公式为 $\Delta Q_{\Delta t} = Q_{in\Delta t} - Q_{out\Delta t}$，即：

$$\rho c_p V\ (T^{t+\Delta t} - T^t)\ = \dot{q}_s A\Delta t - \left(-\lambda A\Delta t \left.\frac{\partial T}{\partial x}\right|_{x=\Delta x}\right) \tag{3.5-30}$$

其中，A 为断面积。如果 A 为单位断面积，即 $A=1$，则体积 $V=1\times\Delta x$，上式变形为：

$$\rho c_p(T^{\Delta+\Delta t} - T^t)\Delta x = \dot{q}_s\Delta t + \lambda\Delta t\frac{\partial T}{\partial x} \tag{3.5-31}$$

如 Δx 趋于无限小（$\Delta x\to 0$），则可得边界上的传热方程：

$$\dot{q}_s + \lambda\frac{\partial T}{\partial x} = 0 \tag{3.5-32}$$

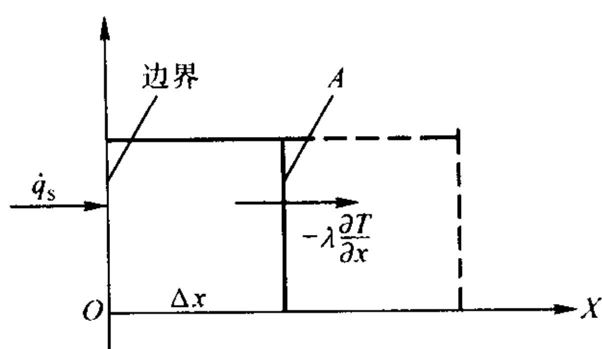


图 3.5-5 边界微元体热量流入与流出

1) 热传达边界条件　在流体（液体、气体）和固体相接触的场合，即使是流体一侧，在边界面上也仍然会因热传导而引起热的流动，即

$$\dot{q}_s = -\lambda_f\frac{\partial T_f}{\partial x} \tag{3.5-33}$$

式中文字下标 f 表示为流体的值。如果流体一侧的温度分布 T_f 为已知的话，代入式（3.5-32）以后就能导出边界条件式。可是在流体的场合，热不仅会因热传导，而且也会由于流体的流动而引起热的流动（对流传热），所以流体一侧的温度分布是不容易知道的，因此引入如下的传热系数 h [$W/(m^2\cdot K)$]。

$$q_u = -\lambda_f\frac{\partial T_f}{\partial x} \equiv h(T_a - T_s) \tag{3.5-34}$$

式中，T_s 为固体表面（边界）的温度；T_a 为流体的代表温度。

因此由式（3.5-32）、式（3.5-34）得出边界条件式为

$$x = x_s;\ h(T_a - T_s) + \lambda\frac{\partial T}{\partial x} = 0 \tag{3.5-35}$$

从传热系数的定义式（3.5-34）就可了解，传热系数 h 将随着流体的导热系数 λ_f 和流体一侧的温度分布（此随流动状态而变化）而变化。因此如果能解流动场和温度场，就能计算出传热系数。

2) 热辐射边界条件　热也可由于热辐射而传导，特别是在液体金属等高温物体的表面，传热的主要方式是热辐射。

上面提过热辐射是由于电磁波引起的热流动现象，它以光速传播。根据布朗克定律，温度为 $T(K)$ 的黑体（不反射电磁波的理想物质）所辐射波长从 λ_m 到 $\lambda_m + d\lambda_m$ 的比辐射能 $E_{b\lambda}d\lambda_m$（$J\cdot s^{-1}\cdot m^{-2}$），可以用下式表示。

$$E_{b\lambda}d\lambda_m = \frac{2\pi c_1}{n^2\lambda^5\ [\exp\ (c_2/n\lambda T)\ -1]} \tag{3.5-36}$$

式中，λ_m 为光速为 c 的介质中的波长；λ 为真空中的波长，$\lambda = n\lambda_m$；n 为折射率，$n=\lambda/\lambda_m = c_0/c$；$c_0$ 为真空中的光速（3×10^8 m/s）；$c_1 = hc_0^2$；$c_2 = hc_0/k$；h 是布朗克常数；k 是玻尔兹曼常数。

折射率 n，除玻璃等以外（石英为 1.5），一般可以当作 1，即 $n=1$，$\lambda_m = \lambda$。

对于一般的传热问题，很少讨论各种波长的辐射能量问题，而是讨论整个波长的辐射能。因此，如果式（3.5-36）在全波长条件下积分，则能得到以下的斯蒂芬－波尔兹曼定律。

$$E_b = \int_0^\infty E_{b\lambda}d\lambda = \Gamma T^4 \tag{3.5-37}$$

式中，Γ 为斯蒂芬－波尔兹曼常数，等于 5.67×10^{-8} W·m^{-2}。

式（3.5-37）中的 E_b 是温度长 $T(K)$ 的黑体的比辐射能（比热流量），而实际物体的辐射能 E 要比此值小，为

$$E = \varepsilon\Gamma T^4 \tag{3.5-38}$$

式中，ε 为（全）辐射系数。

热辐射能是以光速传播的。由于物质表面的反射，从某个面 S_1 实际流出的热量不仅决定于 S_1，而且受周围的面的影响。例如求解浇包中的液体金属表面流出的热辐射能时，必须考虑浇包壁和盖的反射与热辐射。

如果周围的影响不大的话，将式（3.5-38）代入式（3.5-32）的 $\dot{q}_s$ 中 [$T^4 \gg T_a^4$，假定 T_a 是周围温度（K）。另外因为式（3.5-38）是流出的比热流量，所以加上负号]，则边界条件式为

$$-\varepsilon\Gamma T^4 + \lambda\frac{\partial T}{\partial x} = 0,\ x = x_s\ (\text{边界}) \tag{3.5-39}$$

3) 热触热阻边界条件　在固体相互接触的场合，如铸型和砂箱，砂型和冷铁，或者轧辊和铸锭等等，由于实际接触面积比名义接触面积要小，所以在接触面之间产生了温度差（$T_1 - T_2$）。这种情况下的界面比热流量，在引入热阻 R 之后用下式表示。

$$\dot{q}_s = \frac{T_1 - T_2}{R} \tag{3.5-40}$$

此处，如果假定 $h_R = 1/R$，则式（3.5-40）和式（3.5-34）为同一形式，即与热传达边界的条件相同。h_R 称为传热系数。

4) 完全接触边界条件　实际上这种边界条件是很少见的。这种场合的边界条件式可用式（3.5-41）来表达。

$$\lambda_1\frac{\partial T_1}{\partial x} = \lambda_2\frac{\partial T_2}{\partial x} \tag{3.5-41}$$

5) 绝热边界条件　这种情况下边界条件的可用式（3.5-42）来表达。

$$\lambda\frac{\partial T}{\partial x} = 0 \tag{3.5-42}$$

6) 温度为定值的边界条件　这种情况下边界条件的可用式（3.5-43）来表达。

$$T = \text{定值} \tag{3.5-43}$$

7) 比热流量为定值的边界条件　这种情况下边界条件的可用式（3.5-44）来表达。

$$\dot{q}_s = \text{定值} \tag{3.5-44}$$

1.5 潜热处理

1.5.1 定义

液相的内能 E_L 大于固相的内能 E_S，因此，当合金凝固由液相变为固相时，必须产生 $\Delta E = E_L - E_S$ 的内能变化。这个内能变化 ΔE（通常用 L 表示）称为凝固潜热，或称为熔化潜热（latent heat of fusion）。

1.5.2　考虑了析出潜热的热能守恒式

假定单位体积单位时间内固相率的增加率为$\partial g_s/\partial t$，潜热放出的热量为 $\rho L\frac{\partial g_s}{\partial t}$

考虑潜热后的热能守恒式，一维问题为：

$$\rho c_p\frac{\partial T}{\partial t}=\lambda\frac{\partial^2 T}{\partial x^2}+\rho L\frac{\partial g_s}{\partial t}\tag{3.5-45}$$

而 $\rho L\frac{\partial g_s}{\partial t}=\rho L\frac{\partial g_s}{\partial T}\frac{\partial T}{\partial t}$，则一维热能守恒式变为：

$$\rho(c_p-L\cdot\frac{\partial g_s}{\partial T})\frac{\partial T}{\partial t}=\lambda\frac{\partial^2 T}{\partial x^2}\tag{3.5-46}$$

同理，在二维场合下考虑潜热后的热能守恒式为：

$$\rho(c_p-L\frac{\partial g_s}{\partial T})\frac{\partial T}{\partial t}=\lambda(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2})\tag{3.5-47}$$

三维场合下考虑潜热后的热能守恒式为：

$$\rho(c_p-L\frac{\partial g_s}{\partial T})\frac{\partial T}{\partial t}=\lambda(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2})\tag{3.5-48}$$

关键求固相率 g_s 和温度 T 的关系。严格讲，质量相率 f_s 和体积固相率 g_s 是不同的，但以下近似认为 $f_s=g_s$。

1.5.3　固相率和温度的关系

一般从状态图可知道 f_s 与 T 的关系，但对恒温下凝固的纯金属，共晶凝固和包晶凝固，其固相率不能根据温度来确定。对于具有一定结晶温度范围的合金，固相结晶析出的固－液共存区中，液相线温度是与液相浓度相对应的。

（1）已知平衡分配系数 k

$$k_0=C_s/C_L\tag{3.5-49}$$

式中，C_s 为固相浓度；C_L 为液相浓度。

图3.5-6是杠杆定律的示意图，图3.5-7是二元合金相图一角。假定液相线为直线，k 为常数，对于 $k<1$ 的合金，T_L 与 C_K 呈线性关系，则有：

$$T_L=T_f-\sum_i a_iC_i$$

式中，T_f 为作为熔剂的纯金属的熔点；a_i 为液相线温度随成分 i 的浓度 C_i（%）变化的下降系数。

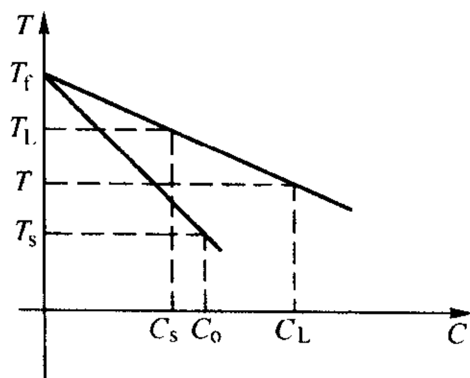


图3.5-6　杠杆定律

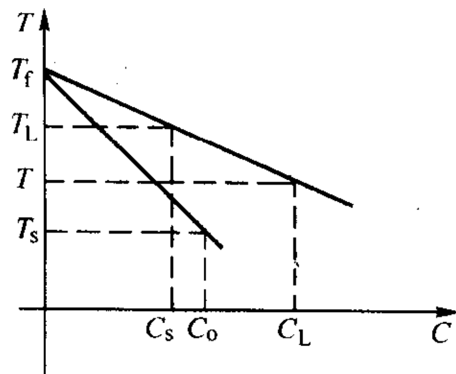


图3.5-7　二元合金相图一角

因此，为了求 T_L，必须知道固－液共存区的溶质浓度，而溶质浓度又随固相率而变化。即要知道 f_s 与 T 的关系，就是要了解 f_s 与溶质浓度的关系。

杠杆定律：

$$f_s=\frac{C_L-C_0}{C_L-C_s}\tag{3.5-50}$$

变形后得：

$$C_L=\frac{C_0}{1+f_s\ (k_0-1)}$$

代入 T_L 与 C_i 关系式：

$$T=T_f-\sum a_i\frac{C_0^i}{1+f_s(k_i-1)}$$

式中，C_0^i 为成分 i 的初始浓度；k_i 为成分 i 的平衡分配系数。

对于二元合金：

$$T=T_f-\frac{T_f-T_L}{C_0}\times\frac{C_0}{1+f_s\ (k_0-1)}$$
$$=T_f-\frac{T_f-T_L}{1+f_s\ (k_0-1)}$$

所以，

$$f_s=\frac{1}{1-k_0}\times\frac{T_L-T}{T_f-T}\tag{3.5-51}$$

（2）平衡分配系数 k 未知

先采用热分析法求出凝固开始度 T_L 和结束温度 T_s，之后，假定

1）T 与 f_s 呈线性分布，即 $T=T_L-(T_L-T_s)f_s$，所以

$$\frac{\partial f_s}{\partial T}=\frac{1}{T_L-T_s}\tag{3.5-52}$$

2）T 与 f_s 呈二次分布，即 $T=T_L-(T_L-T_s)f_s^2$，所以

$$\frac{\partial f_s}{\partial T}=-\frac{1}{2}\times\frac{1}{(T_L-T_s)^{1/2}\ (T_L-T)^{1/2}}\tag{3.5-53}$$

1.5.4　潜热的实际处理方法

（1）等价比热容法

比热容是指单位质量物体降低单位温度所释放的热量，单位质量金属在凝固温度范围内降低单位温度时释放的热量也可以理解成比热容。实际上这个比热容包括两部分，即物体的真正比热容和凝固潜热引起的比热容的增加，从而称此比热容为等价比热容或者有效比热容（亦称当量比热容），记为 C_e。那么考虑到潜热的三维能量守恒式可以写为：

$$\frac{\partial T}{\partial t}=\frac{\lambda}{\rho C_e}\ (\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2})\tag{3.5-54}$$

由式（3.5-54）及式（3.5-48）得：

$$C_e=c_p+L\frac{\partial f_s}{\partial T}\tag{3.5-55}$$

若 f_s 与 T 呈线性关系，则式（3.5-55）可写为：

$$C_e=c_p+\frac{L}{T_L-T_s}\tag{3.5-56}$$

若 f_s 与 T 呈二次分布，则式（3.5-55）可写为：

$$C_e=c_p+\frac{L}{1-k_0}\times\frac{T_L-T_f}{(T_f-T)^2}\tag{3.5-57}$$

等价比热容法适合凝固区间比较大的合金，对凝固区间较小的合金，当温度通过液相线温度和固相线温度时产生显著误差，所以采用等价比热容法来处理潜热问题时要进行温度修正，这点我们在后面要详细说明。

（2）热焓法

热焓法是基于热焓的计算公式，对于凝固过程的金属，其热焓 H 可定义为：

$$H=\int_0^T c_p\mathrm{d}T+(1-f_s)L\tag{3.5-58}$$

式（3.5-58）对温度求导，可得

$$\frac{\partial H}{\partial T}=c_p-L\frac{\partial f_s}{\partial T}\tag{3.5-59}$$

将式（3.5-59）代入式（3.5-48）即得：

$$\rho\frac{\partial H}{\partial t}=\lambda\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}\right) \quad (3.5\text{-}60)$$

这种方法与等价比热容法类似，适用于有一定结晶温度范围的合金。

(3) 温度回复法

对于共晶合金来说，凝固开始的一段时间内，固相不断增多，但温度基本上始终保持在熔点附近。这是由于释放的潜热补偿了传导带走的热量，亦即补偿了传热所引起的温度的下降。由于热量的多少常以单元体的温度变化来表示，因此可将这部分热量折算成所能补偿的温度降落，加入温度计算中去。这就是温度回升法或温度补偿法。

假定某个领域（体积 V）中固相率增加 Δg_s，其放出的潜热（被夺走的热量）用下式表示：

$$Q_S=\rho V\Delta g_S L \quad (3.5\text{-}61)$$

处理时，先不考虑潜热放出，求出微小时间 Δt 内以 T_L 线开始的温度降低：

$$\Delta T=T_L-T \quad (3.5\text{-}62)$$

如果 $\Delta t>0$，就产生凝固，由于放出潜热，温度回升到 T_L（假定无过冷），下式成立：

$$Q_s=\rho c_p V\Delta T \quad (3.5\text{-}63)$$

联立求解：

$$\Delta g_s=c_p\Delta T/L \quad (3.5\text{-}64)$$

此法采用 g_s 的增加来代替潜热的放出。若固相率为1（$\sum\Delta g_s=1$），则表明领域 V 凝固结束。

温度回升法适用于共晶合金以及结晶温度范围小的合金。

(4) 采用改良的等价比热容法的温度场有限差分格式

1) 假想凝固区间　在数值模拟中，如果遇到纯金属或共晶成分合金，此时可以假设该合金存在一定范围的凝固区间 ΔT（比如说取 $\Delta T=0.1℃$），称该凝固区间为假想凝固区间，该假想凝固区间的液相线温度 T'_L 和固相线温度 T'_s 可以按式（3.5-64）、式（3.5-64）求出。其中 T_0 是熔点或者共晶点温度，这样就可以用等价比热容法来处理了。

$$T'_L=T_0+\Delta T/2 \quad (3.5\text{-}65)$$

$$T'_s=T_0-\Delta T/2 \quad (3.5\text{-}66)$$

另外对于凝固区间太小的合金，扩大其凝固区间，此时该假想凝固区间的液相线温度 T_L 和固相线温度 T_s 可以按式（3.5-67)、式（3.5-68）求出。其中 T_L、T_s 是实际的液相线温度和固相相线温度，ΔT 是假想凝固区间的大小。

$$T'_L=\frac{T_L+T_s}{2}+\frac{\Delta T}{2} \quad (3.5\text{-}67)$$

$$T'_s=\frac{T_L+T_s}{2}-\frac{\Delta T}{2} \quad (3.5\text{-}68)$$

当然，假想凝固区间会导致模拟和实际的差异，但只要假想温度区间足够小，这个差异可以忽略不计。

2) 采用改良的等价比热容法的温度场差分公式　由于提出了假想凝固区间的概念，所以我们可以采用等价比热容法来处理共晶成分合金的潜热问题，我们将基于假想凝固区间的等价比热容法称为改良的等价比热容法。现在我们来讨论采用改良的等价比热容法的温度场差分格式。

图3.5-8是某一差分单元（$\Delta z=\Delta y=\Delta x$）及其6个邻接单元，采用有限差分法将方程式（3.5-54）离散得：

$$\frac{T^{t+\Delta t}-T^t}{\Delta t}=\frac{1}{\rho C_e\Delta x}\sum_{j=1}^{6}\frac{(T_j^t-T^t)}{\frac{\Delta x}{2\lambda}+\frac{\Delta x}{2\lambda_j}} \quad (3.5\text{-}69)$$

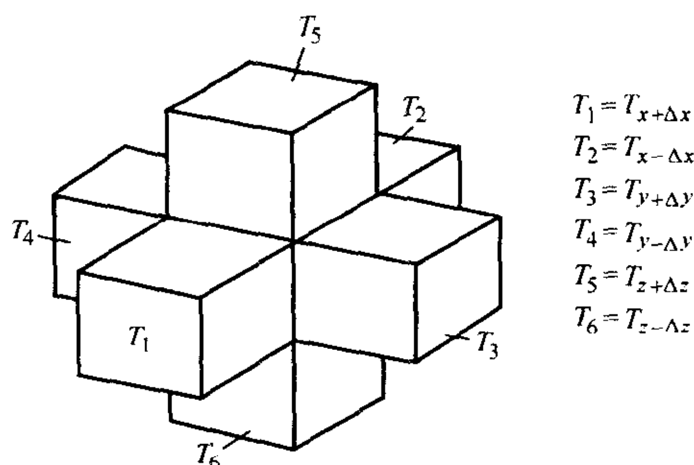


图3.5-8 某一差分单元及其邻接单元

若认为 f_S 与 T 呈线性关系，则 $C_e=c_p+\frac{L}{T'_L-T'_s}$，则上式可整理为：

$$T^{t+\Delta t}=T^t+\frac{\Delta t}{\rho C_e\Delta x}\sum_{j=1}^{6}\frac{(T_j^t-T^t)}{\frac{\Delta x}{2\lambda}+\frac{\Delta x}{2\lambda_j}} \quad (3.5\text{-}70)$$

所以，采用改良的等价比热法的温度场迭代公式总结为如下两种情况：

① 当 $T>T'_L$ 或 $T<T'_s$ 时，迭代公式为

$$T^{t+\Delta t}=T^t+\frac{\Delta t}{\rho c_p\Delta x}\sum_{j=1}^{6}\frac{(T_j^t-T^t)}{\frac{\Delta x}{2\lambda}+\frac{\Delta x}{2\lambda_j}} \quad (3.5\text{-}71)$$

其中，$\Delta t<\frac{\rho C}{6\lambda}\Delta x\Delta x$。

② 当 $T'_L>T>T'_s$ 时，迭代公式为

$$T^{t+\Delta t}=T^t+\frac{\Delta t}{\rho C_e\Delta x}\sum_{j=1}^{6}\frac{(T_j^t-T^t)}{\frac{\Delta x}{2\lambda}+\frac{\Delta x}{2\lambda_j}} \quad (3.5\text{-}72)$$

式中，$C_e=C_p+\frac{L}{T'_L-T'_s}$，$\Delta t<\frac{\rho C_e}{6\lambda}\Delta x\Delta x$。

3) 跨越凝固区间或假想凝固区间时的温度校正　图3.5-9是改良的等价比热容法的示意图，从图中可以看出在应用改良的等价比热容法时，当温度跨越 T'_L、T'_s 时计算所得的温度会有一定偏差，所以必须对温度进行适当的校正。这包括两个方面的内容，一方面是降温过程的校正；另一方面是重熔过程的校正。下面分不同情况进行讨论。

① 当温度从液相线以上降至液固线之间时，应将按式（3.5-71）得到的 $T^{t+\Delta t}$ 作如下校正：

$$T'=T'_L+(T^{t+\Delta t}-T'_L)\,C_p/C_e \quad (3.5\text{-}73)$$

② 当温度从凝固区间降至固相线以下时，应将按式（3.5-72）得到的 $T^{t+\Delta t}$ 作如下校正：

$$T'=T'_s+(T^{t+\Delta t}-T'_s)\,C_e/C_p \quad (3.5\text{-}74)$$

③ 当温度从 T_L 以上降至固相线以下时，应将按式（3.5-71）得到的 $T^{t+\Delta t}$ 作先如下校正：

$$T'=T'_L+(T^{t+\Delta t}-T'_L)\,C_p/C_e \quad (3.5\text{-}75)$$

如果 $T'<T'_s$，还要在此基础上作二次校正：

$$T''=T'_s+(T'-T'_s)\,C_e/C_p \quad (3.5\text{-}76)$$

④ 当温度上升到液固线之间时，应将按式（3.5-71）得到的 $T^{t+\Delta t}$ 作如下校正：

$$T'=T'_s+(T^{t+\Delta t}-T'_s)\,C_p/C_e \quad (3.5\text{-}77)$$

⑤ 当温度从凝固区间上升到液相线以上时，应将按式（3.5-72）得到的 $T^{t+\Delta t}$ 作如下校正：

$$T'=T'_L+(T^{t+\Delta t}-T'_L)\,C_e/C_p \quad (3.5\text{-}78)$$

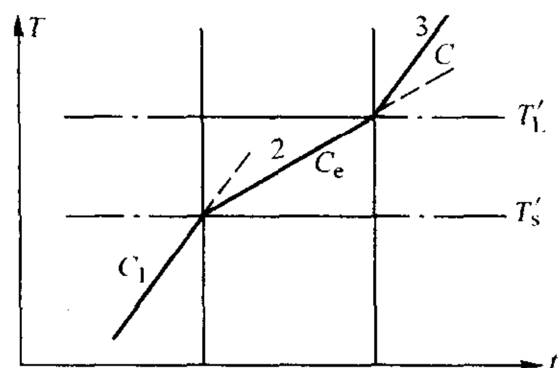
⑥ 当温度从固相线以下上升到液相线以上时，应将按式（3.5-71）得到的 $T^{t+\Delta t}$ 作如下校正：

$$T'=T'_s+(T^{t+\Delta t}-T'_s)\,C_p/C_e \quad (3.5\text{-}79)$$

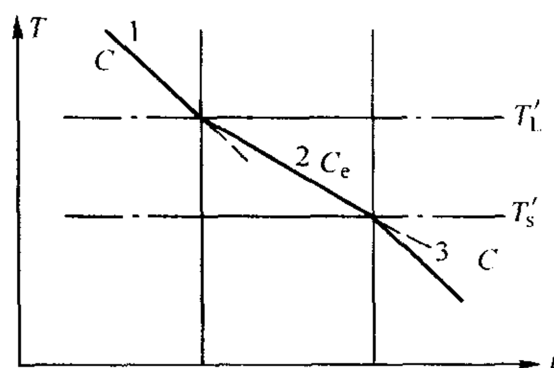
如果 $T'>T'_L$，还要在此基础上作二次校正：

$$T'' = T'_L + (T' - T'_L)\ C_e/C_p \qquad (3.5\text{-}80)$$

特别要注意的是如果是对凝固区间或假想凝固区间 ΔT 比较小或者接近0时，上述第3、6种情况将比较普遍，如处理不当则会带来很大的误差。



(a) 重熔过程



(b) 降温过程

图 3.5-9　改良的等价比热容法示意图

1.6　温度场数值模拟流程图

温度场数值模拟系统包括三大部分：前处理、计算分析、后处理。前处理包括含三维造型及网格剖分两大部分。三维造型主要是将要进行分析的对象其三维实体输入计算机；网格剖分则是将已输入的三维实体剖分成计算所需的网格单元。计算分析模块先要进行初始条件和边界条件的设置，在此基础上对前处理所得差分网格系统进行温度场分析。后处理主要任务是数据的可视化，将计算分析所得的温度场结果，真实、生动、形象地显示出来。此部分需要采用计算机图形学、多媒体技术、图形处理技术等科学的理论与方法。

图3.5-10是温度场数值模拟流程图。

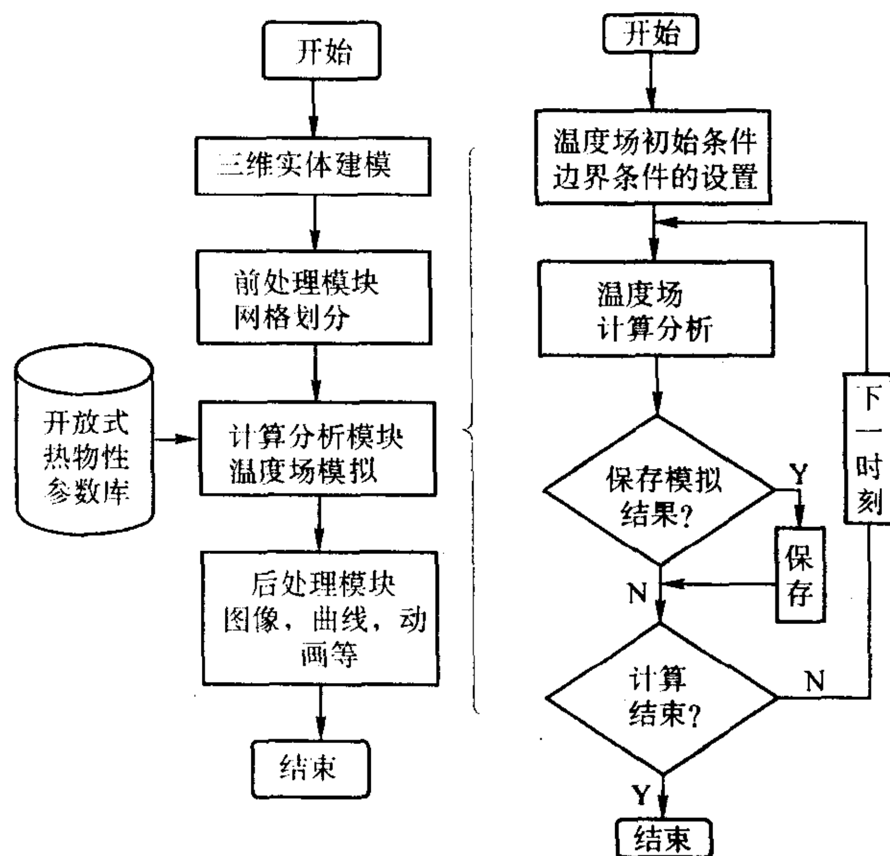


图 3.5-10　温度场数值模拟流程图

2　铸造充型过程数值模拟

2.1　概述

通常情况下，气体和液体称为流体。流体更严格的定义可述为流体是一种在微小切应力作用下会发生连续变形的物质。一切流体都有如下基础属性：在不受外力作用的情况下，流体没有它自己的形状。铸造充型过程的液态金属就是满足上述条件的牛顿流体。此外应该注意，存在一种特殊流体，即所谓非牛顿流体，上述定义对它是不适用的。

2.1.1　流体的基本概念

1）恒定流动和非恒定流动　随时间而变化的流动称为非恒定流动，不随时间变化的称为恒定流动。某种流动是当作恒定流动考虑还是当作非恒定流动考虑要根据具体场合而定。如紊流问题，从微观讲是非恒定流动，但如果考虑某段时间的平均值的话，则可以当作恒定流动。

2）层流和紊流　所谓层流是指流体的流线不相互掺混、井然不乱的一种流动方式。紊流是指流线相互掺混的一种流动方式。流动是层流还是紊流，可以根据由经验得出的某个临界雷诺数来确定。雷诺数是惯性力和黏性力之比，它是一个量纲为1的数。一般地认为雷诺数大于某一临界值时即为紊流。

3）黏性流与牛顿流体　具有黏性的黏性液体的流动称为黏性流，所谓黏性是指流体变形时产生的阻力或能量损失。如果阻力与变形速度成正比，这样的黏性流体称为牛顿流体，这个比例系数就是黏度。一般情况下，液体黏度随温度上升而降低，气体黏度随温度上升而增加。

2.1.2　流动分析的主要方法

流场计算方法有SIMPLE方法、MAC及SMAC方法、SOLA-VOF方法、SOLAMAC方法、FAN方法、Finite Volume方法、格子气模型等。下面分别作简要介绍。

1）SIMPLE方法　SIMPLE（Semi-Implicit Method for Pressure Linked Equations）法是由美国明尼苏达州大学S.V.Patankar教授总结出来的，又称压力连接方程的半隐式方法。可以用来计算非定域、不稳定速度场，计算结果能够满足连续性方程、动量方程的要求。但是该方法采用压力场和速度场双场同时迭代，计算处理速度较慢，另外对带有自由表面的流动则处理不太方便。

2）MAC及SMAC方法　MAC（Maker And Cell）方法是由美国Los Alamos国家实验室于1965年提出来的。MAC的诞生使求解类似于铸件充型过程这种黏性不可压缩非稳态带有自由表面的流动成为可能。

MAC方法主要求解的思路是：基于有限差分网格将动量守恒方程（N-S方程）和连续性方程（质量守恒方程）进行离散，并将动量守恒方程与作为约束条件的连续性方程合并变成一个与压力相关的泊松方程，通过动量守恒方程和泊松方程的迭代，求解出流动的速度场和压力场。MAC方法另一个主要特征，就是设置随流体流动的标识粒子（Marker Particles），它并不参与计算，只是作为一种描述手段，以跟踪、描述任意时刻流体自由边界的移动。

可见，MAC法在求解动量守恒和连续性方程时采用的是速度和压力两场迭代，另一方面由于采用粒子跟踪法，在处理自由表面时需要大量的内存和时间，因此，MAC法求解流动问题速度太慢以致影响其广泛应用。于是就产生了简化的MAC方法，即SMAC法（Simplified MAC），该方法在处理速度场时，在离散后的差分方程的迭代中没有压力项计算，通常校正压力项由校正势函数来取代，并用来校正速度场，校正后的速度场如不能满足质量守恒方程，则反复迭代势函数，修改速度场，直至满足质量守恒方程。该方法的特

点在于势函数一场迭代，计算速度得到很大程度的提高。

3) SOLA - VOF 方法　为了克服 MAC 及 SMAC 方法中因采用粒子跟踪来描述自由表面既费时又费内存的缺点，Los Alamos 科学实验室在 MAC 方法的基础上发展了新的计算技术，即 SOLA - VOF 方法。SOLA - VOF 是英文 Solution Algorithm - Volume of Fluid 的简称。

SOLA - VOF 方法求解速度场和压力场时，每个计算单元的校正压力直接由连续性方程算出的速度求出，然后校正速度场。整个计算过程中由速度初值及猜测压力值试算速度场的过程并不参与迭代，因而也是一场迭代，计算速度快。

与 MAC 方法不同的是 SOLA - VOF 技术采用体积函数 F 代替标识粒子来描述自由表面的位置，正是由于该方法的采用使自由表面处理速度大大加快，对计算机内存的需求显著降低，因此目前处理铸件流动问题大多采用此方法。

4) SOLAMAC 方法　SOLAMAC 方法吸取了 MAC 方法和 SOLA - VOF 方法的优点，即在求解流动问题时，利用 SOLA 方法计算速度场和压力场，用 MAC 方法中的标识粒子显示流动范围的变化，跟踪自由表面的位置。

该方法可以处理形态较为复杂的流动问题，并且速度较快。SOLA - MAC 方法的另一个特点是后置处理即流动过程模拟结果的表达比较丰富、生动，可以得到速度分布图、流线图、环流的位置和剧烈流动的范围等结果。

5) FAN 方法　FAN 技术（Flow Analysis Network Method）是由 Z.Tadmoret 研究提出的流动过程分析法。FAN 方法忽略惯性和重力作用对流动过程的影响，考虑流动过程仅由速度控制，并研究一维不可压缩牛顿流体的稳定流动。该方法求解动量方程时，不计其中的非稳态项、对流项及外力项，使动量偏微分方程得以简化。根据流动过程的条件，将结点分为四种类型，即入口结点、全充满结点、边界结点以及未充型结点。假定边界结点和未充型结点压力为零，如果给出入口结点流速，通过解联立方程即可得到每个单元的压力和流速，以及边界结点的流速和变化情况。经过进一步研究和改进流场模拟 FAN 技术，可以处理比较复杂的流动情况，在处理过程中考虑了流动过程中的凝固过程和腔内的气体反应，并在高速流动中有具体应用。

6) Finite Volume 方法　有限体积法（Finite Volume Method）的控制方程为非稳态动量方程，它能够处理高速黏性流体流动的数值计算。有限体积法采用上风差分格式解动量方程，相比之下，这种方法由于添加了隐式速度项使单步计算时间延长，迭代次数增加，因而计算效率不高。为此，目前已研究出针对有限体积法的改进松弛解法，它可以保证在提高计算效率的同时获得稳定的数值解。这种松弛解法简单分为两个基本步骤，即首先在椭圆子域内直接求解，然后将计算结果在全部计算域内沿流体运动方向进行线性高斯 - 塞得尔超松弛迭代。新算法计算效率可以比原来提高 20%。

7) 格子气模型　格子气模型是法国和美国的科学家在 20 世纪 80 年代提出的一种全新的计算流体力学方法，这种方法的提出基于以下理论：许多行为简单的微观个体组成的宏观物理系统具有很复杂的物理性质，大量个体的集行为可以表现出高度的有序性。在这个基础上提出了格子气模型（lattice - gas model）。即把流体看成是由大量的微观粒子组成的，这些微观粒子在规则或不规则的空间内按一定的法则相互左右和移动，在宏观看到的就是流体的流动。这些微观粒子的运动在热力学极限下，用粗略平均的方法可以逼近动量守恒方程，故而被认为可以代替动量守恒方程来解决流动问题。目前，这种方法正在研究之中，初期的研究结果表明，这种格子气模型比传统的解动量守恒方程的方法快1 000倍，这也是这种方法最吸引人的地方。目前已有用格子气模型来模拟带有自由表面的流体流动的研究报告。这是一种需要完善但很有前途的流动过程数值模拟方法。

2.1.3　本节的思路和主要内容

同上节铸造凝固过程数值模拟一样，本节的思路还是采用有限差分法对铸件充型过程的数学模型进行离散，并采用 SOLA - VOF 方法来进行流动分析。着重介绍如下内容：流动分析的数学模型、基于有限差分的离散格式、流动分析的 SOLA - VOF 方法、流动场的初始条件和边界条件、流动分析的数值稳定性条件以及流动场数值模拟流程图等。

2.2　数学模型

2.2.1　从 Eular 方程到 Navier - Stoks 方程

流体力学最基本的依据是牛顿第二定律 $F = ma$。在理想流体中，流体受力包括引力场的重力 G 和液体单元间的压力 P。重力 $G = mg$ 作用在质量上，而压力 P 作用在相关质量的外围表面上。由于理想流体无黏滞力，一个立方体微元的全部受力由该微元质量所受的重力与其 6 个表面所受的压力（每个方向两个面压力作用方向相反，其效果应该相减）组成。流体的加速度等于流体速度的变化率 $a = \frac{dv}{dt}$。按牛顿第二定律，微元体在 x，y，z 三方向上的速度 u，v，w 应分别满足关系

重力分量 + 表面力分量 = 质量 × 加速度分量

即

$$\mathrm{d}x\mathrm{d}y\mathrm{d}z\,\rho g_x - \mathrm{d}P\mathrm{d}y\mathrm{d}z = \mathrm{d}x\mathrm{d}y\mathrm{d}z\,\rho a_x$$
$$\mathrm{d}x\mathrm{d}y\mathrm{d}z\,\rho g_y - \mathrm{d}P\mathrm{d}z\mathrm{d}x = \mathrm{d}x\mathrm{d}y\mathrm{d}z\,\rho a_y$$
$$\mathrm{d}x\mathrm{d}y\mathrm{d}z\,\rho g_z - \mathrm{d}P\mathrm{d}x\mathrm{d}y = \mathrm{d}x\mathrm{d}y\mathrm{d}z\,\rho a_z$$

式中，$\mathrm{d}x\mathrm{d}y\mathrm{d}z$ 为微元体积，ρ 为流体密度，二者之积为微元质量，a_x，a_y，a_z 为微元流体加速度三分量，分别等于相应的速度分量对时间求导。代入求导式，并将等式两边同除以 $\mathrm{d}x\mathrm{d}y\mathrm{d}z$，得

$$\rho g_x - \frac{\partial P}{\partial x} = \rho\frac{\mathrm{d}u}{\mathrm{d}t} \tag{3.5-81}$$

$$\rho g_y - \frac{\partial P}{\partial y} = \rho\frac{\mathrm{d}v}{\mathrm{d}t} \tag{3.5-82}$$

$$\rho g_z - \frac{\partial P}{\partial z} = \rho\frac{\mathrm{d}w}{\mathrm{d}t} \tag{3.5-83}$$

三式右边各加速度项用速度分量的全导数而不是偏导数来表示，其含义是，这个加速度是同一流体微元在位置移动中速度的变化，而不是流场中同一位置流过的不同流体间的速度变化，在数学形式上，后者是流场中该点速度对时间的偏导数，它只是全导数四项中的一项，它们的关系如下：

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial t} \tag{3.5-84}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial v}{\partial z}\frac{\partial z}{\partial t} \tag{3.5-85}$$

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial t} + \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t} \tag{3.5-86}$$

将上面各式代入式（3.5-81）、式（3.5-82）以及式（3.5-83），动力学方程可写为

$$g_x - \frac{1}{\rho}\frac{\partial P}{\partial x} = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} \tag{3.5-87}$$

$$g_y - \frac{1}{\rho}\frac{\partial P}{\partial y} = \frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} \tag{3.5-88}$$

$$g_z - \frac{1}{\rho}\frac{\partial P}{\partial z} = \frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z} \tag{3.5-89}$$

这就是不可压缩理想流体的 Eular（欧拉）方程，它是理想流体的动力学方程，是通常用到的积分形式的 Bernoulli（伯努利）方程的微分形式。

对于实际流体，动力黏度 $\mu \neq 0$，存在黏性力。黏性力既存在于流体侧面运动的切向，也存在于流动的正面微元的

法向。切向力是内摩擦性质的力，与侧向速度梯度成正比：

$$\tau_{yx} = \mu \frac{\partial u}{\partial y}$$

$$\tau_{zx} = \mu \frac{\partial u}{\partial z}$$

其中 τ 表示流体侧面的切向力，其第一下标表示该力所在侧面的法向，第二下标表示该力的方向。如第一式的含义是流体朝向 y 轴的侧面所受 x 方向的切向力正比于流体 x 方向流速 u 沿 y 轴方向的速度梯度。如此类推。

法向黏力是牵连性质的力，对于不可压缩流体，它与速度方向上的速度梯度成正比：

$$n_{xx} = \mu \frac{\partial u}{\partial x}$$

其中 n 表示流动方向正面法向黏力，下标含义同上。以上三个力都发生在流体的一个表面上，而一个受力的微元体在每个方向都有两个面，作用于微元的是这两个面同方向力之差。因此，各面同一个方向的黏性力之合力应为：

$$s_x = \mu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right)$$

注意到运动黏度 $\gamma = \frac{\mu}{\rho}$，微元体各面黏性力的向量和叠加在重力和压力上，三个方向上的动力学方程式（3.5-87）～式（3.5-89）变成

$$g_x - \frac{1}{\rho}\frac{\partial P}{\partial x} + \gamma\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} \tag{3.5-90}$$

$$g_y - \frac{1}{\rho}\frac{\partial P}{\partial y} + \gamma\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right) = \frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} \tag{3.5-91}$$

$$g_z - \frac{1}{\rho}\frac{\partial P}{\partial z} + \gamma\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right) = \frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z} \tag{3.5-92}$$

这就是实际流体的 Navier－Stoks 方程。

2.2.2 分离时间变量

求解数理方程，特别是用数值方法求解数理方程，常常需要分离出时间变量，以降低求解难度。对于式（3.5-90）～式（3.5-92），将它们移项，得到：

$$\frac{\partial u}{\partial t} = g_x - \frac{1}{\rho}\frac{\partial P}{\partial x} - \left(u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}\right) + r\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) \tag{3.5-93}$$

$$\frac{\partial v}{\partial t} = g_y - \frac{1}{\rho}\frac{\partial P}{\partial y} - \left(u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z}\right) + r\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right) \tag{3.5-94}$$

$$\frac{\partial w}{\partial t} = g_z - \frac{1}{\rho}\frac{\partial P}{\partial z} - \left(u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z}\right) + r\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right) \tag{3.5-95}$$

类似温度场方程的离散处理，将式（3.5-93）～式（3.5-95）差分化，就可从当前时刻的速度值 u，v，w 分别求出下一时刻的速度值 u'，v'，w'。这是流动场速度计算的基本迭代公式。

2.2.3 方程的矢量形式

在一些文字叙述中，为求表达的简练，常常借用有关的算符，将式（3.5-90）～式（3.5-92）三式合成为矢量形式（分别记速度向量为 $\boldsymbol{V} = u\cdot\bar{i} + v\cdot\bar{j} + w\cdot\bar{k}$，重力向量为 $\boldsymbol{G} = g_x\bar{i} + g_y\bar{j} + g_z\bar{k}$）：

$$\boldsymbol{G} - \frac{1}{\rho}\nabla \boldsymbol{P} + r\nabla^2\boldsymbol{V} = \frac{\partial}{\partial t}\boldsymbol{V} + \boldsymbol{V}\nabla\boldsymbol{V} \tag{3.5-96}$$

其中，∇ 为一阶微分算子（或算符）

$$\nabla = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z} \tag{3.5-97}$$

注意它和 Δ 算符的区别：

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{3.5-98}$$

Δ 为二阶微分算子，也叫 Laplace（拉普拉斯）微分算子。这两种算符作用于一个场函数变量（既可作用于矢量也可作用于标量）时，都可按四则运算分配法则将其中的三项分别作用于该函数变量，但注意一阶微分算子本身具有矢量属性，经其作用后变量性质发生变化。作用于标量，其结果变成矢量，如温度梯度 $\nabla T = \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right)T = \frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \frac{\partial T}{\partial z}$ 是一有方向量；作用于矢量，其结果变成标量，如速度散度 $\nabla V = \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right)V = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}$ 是一无方向纯数量，其中 u，v，w 是速度矢量 V 的三个分量。

2.2.4 连续性方程

在式（3.5-93）～式（3.5-95）中，压力 P 是方程组的第四个求解变量，也是时空四维空间的函数。为能求得确定解，必须在三个方程之外补入一个约束方程，这就是连续性方程。对于不可压缩流体无源流动场而言，在充满流体的流动域中任何一点，流体速度的散度应该等于 0，也就是无源无漏，质量守恒。其数学形式为

$$\mathrm{div}\,V = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{3.5-99}$$

矢量的散度可写为 $\mathrm{div}V$，也可写为 ∇V。在下文的叙述中，也用 D 来表示散度，即 $D = \mathrm{div}V$。

2.3 数学模型的离散

欲准确求出 Navier－Stoks 方程和连续性方程的数学解析解是非常困难的，因此需要采用数值求解方法。数值求解的实质就是将连续的求解空间离散成有限个相对独立的微元体，然后基于这些微元体进行求解计算，最后将所有微元体求解结果在时间上联系起来作为整个求解目标的结果。所以，数值求解的前提就是要对上述偏微分方程组在空间上和时间上进行离散。

2.3.1 离散格式的选择

离散格式是指采用何种方式将连续的场变量（速度、压力）进行离散。目前铸造数值模拟一般采用了 S.V. 帕坦卡教授提出的交错网格进行离散，各变量在三维交错网格中的位置如图 3.5-11 所示。这种网格形式的采用与非交错网格相比，有两点好处：①避免了不合乎实际的速度场却能满足连续性方程的问题；②两个相邻网格点之间的压力差成了位于这两个网格点之间速度分量的自然驱动力，更具有明确的物理意义。

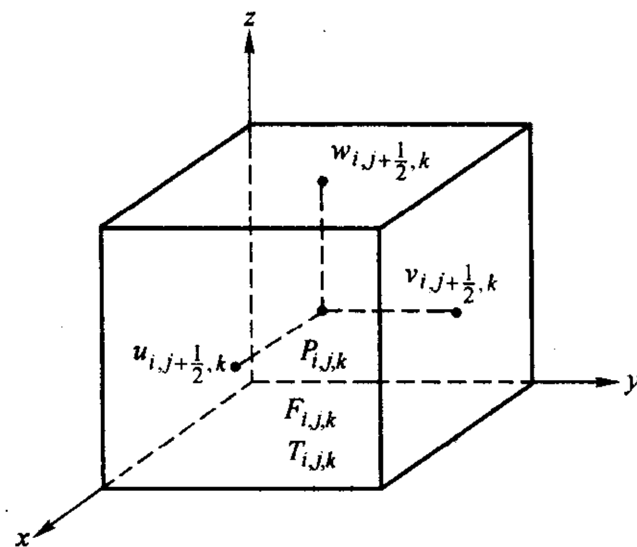


图 3.5-11 交错网格示意图

2.3.2 动量方程（Navier－Stoks 方程）的离散

利用上述交错网格离散格式，对动量守恒方程

(Navier - Stoks 方程）进行离散，可以得到如下形式的离散化方程。

$$u^{n+1}_{i+\frac{1}{2},j,k} = u^{n}_{i+\frac{1}{2},j,k} + \delta t \left[\frac{P^{n+1}_{i,j,k} - P^{n+1}_{i+1,j,k}}{\rho \delta x_{i+\frac{1}{2}}} + g_x - FUX - FUY - FUZ + VISX \right] \tag{3.5-100}$$

$$v^{n+1}_{i,j+\frac{1}{2},k} = v^{n}_{i,j+\frac{1}{2},k} + \delta t \left[\frac{P^{n+1}_{i,j,k} - P^{n+1}_{i,j+1,k}}{\rho \delta y_{j+\frac{1}{2}}} + g_y - FVX - FVY - FVZ + VISY \right] \tag{3.5-101}$$

$$w^{n+1}_{i,j,k+\frac{1}{2}} = w^{n}_{i,j,k+\frac{1}{2}} + \delta t \left[\frac{P^{n+1}_{i,j,k} - P^{n+1}_{i,j,k+1}}{\rho \delta z_{k+\frac{1}{2}}} + g_z - FWX - FWY - FWZ + VISZ \right] \tag{3.5-102}$$

其中

$$FUX = \frac{u_{i+\frac{1}{2},j,k}}{\delta x_{au}} [\delta x_{i+1} DUB + \delta x_i DUT + \alpha \mathrm{sgn}(u)(\delta x_{i+1} DUB - \delta x_i DUT)]$$

$$FUY = \frac{v_{i+\frac{1}{2},j,k}}{\delta y_{au}} [\delta y_{j+\frac{1}{2}} DUL + \delta y_{j-\frac{1}{2}} DUR + \alpha \mathrm{sgn}(v)(\delta y_{j+\frac{1}{2}} DUL - \delta y_{j-\frac{1}{2}} DUR)]$$

$$FUZ = \frac{w_{i+\frac{1}{2},j,k}}{\delta z_{au}} [\delta z_{k+\frac{1}{2}} DUQ + \delta z_{k-\frac{1}{2}} DUH + \alpha \mathrm{sgn}(w)(\delta z_{k+\frac{1}{2}} DUQ - \delta z_{k-\frac{1}{2}} DUH)]$$

$$FVX = \frac{u_{i,j+\frac{1}{2},k}}{\delta x_{av}} [\delta x_{i+\frac{1}{2}} DVB + \delta x_{i-\frac{1}{2}} DVT + \alpha \mathrm{sgn}(u)(\delta x_{i+\frac{1}{2}} DVB - \delta x_{i-\frac{1}{2}} DVT)]$$

$$FVY = \frac{v_{i,j+\frac{1}{2},k}}{\delta y_{av}} [\delta y_{j+1} DVL + \delta y_j DVR + \alpha \mathrm{sgn}(v)(\delta y_{j+1} DVL - \delta y_j DVR)]$$

$$FVZ = \frac{w_{i,j+\frac{1}{2},k}}{\delta z_{av}} [\delta z_{k+\frac{1}{2}} DVQ + \delta z_{k-\frac{1}{2}} DVH + \alpha \mathrm{sgn}(w)(\delta z_{k+\frac{1}{2}} DVQ - \delta z_{k-\frac{1}{2}} DVH)]$$

$$FWX = \frac{u_{i,j,k+\frac{1}{2}}}{\delta x_{aw}} [\delta x_{i+\frac{1}{2}} DWB + \delta x_{i-\frac{1}{2}} DWT + \alpha \mathrm{sgn}(u)(\delta x_{i+\frac{1}{2}} DWB - \delta x_{i-\frac{1}{2}} DWT)]$$

$$FWY = \frac{v_{i,j,k+\frac{1}{2}}}{\delta y_{aw}} [\delta y_{j+\frac{1}{2}} DWL + \delta y_{j-\frac{1}{2}} DWR + \alpha \mathrm{sgn}(v)(\delta y_{j+\frac{1}{2}} DWL - \delta y_{j-\frac{1}{2}} DWR)]$$

$$FWZ = \frac{w_{i,j,k+\frac{1}{2}}}{\delta z_{aw}} [\delta z_{k+1} DWQ + \delta z_k \cdot DWH + \alpha \mathrm{sgn}(w)(\delta z_{k+1} DWQ - \delta z_k DWH)]$$

$$VISX = \gamma \left[\frac{DUT - DUB}{\delta x_{i+\frac{1}{2}}} + \frac{DUR - DUL}{\delta y_j} + \frac{DUH - DUQ}{\delta z_k} \right]$$

$$VISY = \gamma \left[\frac{DVT - DVB}{\delta x_i} + \frac{DVR - DVL}{\delta y_{j+\frac{1}{2}}} + \frac{DVH - DVQ}{\delta z_k} \right]$$

$$VISZ = \gamma \left[\frac{DWT - DWB}{\delta x_i} + \frac{DWR - DWL}{\delta y_j} + \frac{DWH - DWQ}{\delta z_{k+\frac{1}{2}}} \right]$$

$$DUB = \frac{u_{i+\frac{1}{2},j,k} - u_{i-\frac{1}{2},j,k}}{\delta x_i} \qquad DUT = \frac{u_{i+\frac{3}{2},j,k} - u_{i+\frac{1}{2},j,k}}{\delta x_{i+1}}$$

$$DUL = \frac{u_{i+\frac{1}{2},j,k} - u_{i+\frac{1}{2},j-1,k}}{\delta y_{j-\frac{1}{2}}} \qquad DUR = \frac{u_{i+\frac{1}{2},j,k} - u_{i+\frac{1}{2},j,k}}{\delta y_{j+\frac{1}{2}}}$$

$$DUQ = \frac{u_{i+\frac{1}{2},j,k} - u_{i+\frac{1}{2},j,k-1}}{\delta z_{k-\frac{1}{2}}} \qquad DUH = \frac{u_{i+\frac{1}{2},j,k+1} - u_{i+\frac{1}{2},j,k}}{\delta z_{k+\frac{1}{2}}}$$

$$DVB = \frac{v_{i,j+\frac{1}{2},k} - v_{i-1,i+\frac{1}{2},k}}{\delta x_{i-\frac{1}{2}}} \qquad DVT = \frac{v_{i+1,j+\frac{1}{2},k} - v_{i,j+\frac{1}{2},k}}{\delta x_{i+\frac{1}{2}}}$$

$$DVL = \frac{v_{i,j+\frac{1}{2},k} - v_{i,j-\frac{1}{2},k}}{\delta y_j} \qquad DVR = \frac{v_{i,j+\frac{3}{2},k} - v_{i,j+\frac{1}{2},k}}{\delta y_{j+1}}$$

$$DVQ = \frac{v_{i,j+\frac{1}{2},k} - v_{i,j+\frac{1}{2},k-1}}{\delta z_{k-\frac{1}{2}}} \qquad DVH = \frac{v_{i,j+\frac{1}{2},k+1} - v_{i,j+\frac{1}{2},k}}{\delta z_{k+\frac{1}{2}}}$$

$$DWB = \frac{w_{i,j,k+\frac{1}{2}} - w_{i-1,j,k+\frac{1}{2}}}{\delta x_{i-\frac{1}{2}}} \qquad DWT = \frac{w_{i+1,j,k+\frac{1}{2}} - w_{i,j,k+\frac{1}{2}}}{\delta x_{i+\frac{1}{2}}}$$

$$DWL = \frac{w_{i,j,k+\frac{1}{2}} - w_{i,j-1,k+\frac{1}{2}}}{\delta y_{j-\frac{1}{2}}} \qquad DWR = \frac{w_{i,j+1,k+\frac{1}{2}} - w_{i,j,k+\frac{1}{2}}}{\delta y_{j+\frac{1}{2}}}$$

$$DWQ = \frac{w_{i,j,k+\frac{1}{2}} - w_{i,j,k-\frac{1}{2}}}{\delta z_k} \qquad DWH = \frac{w_{i,j,k+\frac{3}{2}} - w_{i,j,k+\frac{1}{2}}}{\delta z_{k+1}}$$

$$\delta x_{au} = \delta x_{i+1} + \delta x_i + \alpha \mathrm{sgn}(u_{i+\frac{1}{2},j,k})(\delta x_{i+1} - \delta x_i)$$

$$\delta y_{au} = \delta y_{j+\frac{1}{2}} + \delta y_{j-\frac{1}{2}} + \alpha \mathrm{sgn}(v_{i,j+\frac{1}{2},k})(\delta y_{j+\frac{1}{2}} - \delta y_{j-\frac{1}{2}})$$

$$\delta z_{au} = \delta z_{k+\frac{1}{2}} + \delta z_{k-\frac{1}{2}} + \alpha \mathrm{sgn}(w_{i,j,k+\frac{1}{2}})(\delta z_{k+\frac{1}{2}} - \delta z_{k-\frac{1}{2}})$$

$$\delta x_{av} = \delta x_{i+\frac{1}{2}} + \delta x_{i-\frac{1}{2}} + \alpha \mathrm{sgn}(u_{i+\frac{1}{2},j,k})(\delta x_{i+\frac{1}{2}} - \delta x_{i-\frac{1}{2}})$$

$$\delta y_{av} = \delta y_{j+1} + \delta y_j + \alpha \mathrm{sgn}(v_{i,j+\frac{1}{2},k})(\delta y_{j+1} - \delta y_j)$$

$$\delta z_{av} = \delta z_{k+\frac{1}{2}} + \delta z_{k-\frac{1}{2}} + \alpha \mathrm{sgn}(w_{i,j,k+\frac{1}{2}})(\delta z_{k+\frac{1}{2}} - \delta z_{k-\frac{1}{2}})$$

$$\delta x_{aw} = \delta x_{i+\frac{1}{2}} + \delta x_{i-\frac{1}{2}} + \alpha \mathrm{sgn}(u_{i+\frac{1}{2},j,k})(\delta x_{i+\frac{1}{2}} - \delta x_{i-\frac{1}{2}})$$

$$\delta y_{aw} = \delta y_{j+\frac{1}{2}} + \delta y_{j-\frac{1}{2}} + \alpha \mathrm{sgn}(v_{i,j+\frac{1}{2},k})(\delta y_{j+\frac{1}{2}} - \delta y_{j-\frac{1}{2}})$$

$$\delta z_{aw} = \delta z_{k+1} + \delta z_k + \alpha \mathrm{sgn}(w_{i,j,k+\frac{1}{2}})(\delta z_{k+1} - \delta z_k)$$

$$\delta x_{i+\frac{1}{2}} = (\delta x_{i+1} + \delta x_i)/2 \qquad \delta y_{j-\frac{1}{2}} = (\delta y_{j-1} + \delta y_j)/2$$

$$\delta x_{i-\frac{1}{2}} = (\delta x_{i-1} + \delta x_i)/2 \qquad \delta z_{k+\frac{1}{2}} = (\delta z_{k+1} + \delta z_k)/2$$

$$\delta y_{j+\frac{1}{2}} = (\delta y_{j+1} + \delta y_j)/2 \qquad \delta z_{k-\frac{1}{2}} = (\delta z_{k-1} + \delta z_k)/2$$

式中，$\mathrm{sgn}(u)$表示 u 的符号，α 为权重因子，其取值范围 $0 \leqslant \alpha \leqslant 1.0$。一般说来，动量方程离散格式的选取根据 Paclet 数的差别可分为五种：中心差分格式、上风格式、指数格式、混合格式及乘方格式。根据实际应用情况和经验，一般采用介于中心差分和上风格式之间的离散方法，并由权重因子予以调整，当 $\alpha = 1$ 时为上风格式；$\alpha = 0$ 时为中心差分格式，因此 α 表示受上游变量的影响程度。

2.3.3 连续性方程的离散

连续性方程 $D = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0$ 的离散形式为

$$D^{n+1}_{i,j,k} = \frac{u^{n+1}_{i+\frac{1}{2},j,k} - u^{n+1}_{i-\frac{1}{2},j,k}}{\delta x_i} + \frac{v^{n+1}_{i,j+\frac{1}{2},k} - v^{n+1}_{i,j-\frac{1}{2},k}}{\delta y_j} + \frac{w^{n+1}_{i,j,k+\frac{1}{2}} - w^{n+1}_{i,j,k-\frac{1}{2}}}{\delta z_k} = 0 \tag{3.5-103}$$

2.4 SOLA - VOF 方法

在 SOLA - VOF 方法中，为了确定自由表面的移动，需要求解体积函数方程：

$$\frac{\partial F}{\partial t} + u \frac{\partial F}{\partial x} + v \frac{\partial F}{\partial y} + w \frac{\partial F}{\partial z} = 0 \tag{3.5-104}$$

2.4.1 体积函数的求值

SOLA - VOF 方法利用体积函数 F 来描述整个流场的流动域，F 的定义为

$$F = 单元内流体的体积/单元总体积 \tag{3.5-105}$$

因此体积函数的取值范围为 $0 \leqslant F \leqslant 1$，当 $F = 0$ 时，该单元为空单元，没有流体；当 $F = 1$ 时为满单元，说明该单

元为内部单元；当$0<F<1$时则表示该单元内有流体流入，但又没有充满，即为表面单元。由此可见，在计算铸件充型、流动时只要计算出每个单元的F值就可以得到该铸件在任一时刻的充型、流动状态。

单元（i，j，k）在δt时间段内体积函数的变化值$\delta F_{i,j,k}$可以由式（3.5-104）离散得：

$$\delta F_{i,j,k}=\frac{\Delta Q_{i,j,k}}{\rho_{i,j,k}\delta x_i\delta y_j\delta z_k}$$
$$=-\delta t\left(\frac{u_{i+\frac{1}{2},j,k}-u_{i-\frac{1}{2},j,k}}{\delta x_i}+\frac{v_{i,j+\frac{1}{2},k}-v_{i,j-\frac{1}{2},k}}{\delta y_j}+\frac{w_{i,j,k+\frac{1}{2}}-w_{i,j,k-\frac{1}{2}}}{\delta z_k}\right) \tag{3.5-106}$$

式中，$\Delta Q_{i,j,k}$为单元（i，j，k）在δt时间段内流体质量的变化，显而易见：

当$\delta F_{i,j,k}>0$时，表示该单元流入量大于流出量，单元内流体量在增加；

当$\delta F_{i,j,k}=0$时，表示该单元流入量等于流出量，单元内流体量不变；

当$\delta F_{i,j,k}<0$时，表示该单元流出量大于流入量，单元内流体量在减少。

计算流动场时，当求解每一个时刻速度场与压力场后，都要利用上式求出当前时刻每个网格单元的δF值，并根据$F+\delta F$值的不同情况做不同处理。

2.4.2　SOLA－VOF计算方法

目前流行的SOLA－VOF有限差分流体力学计算方法，用SOLA求解压力场和速度场，用VOF确定流动域和自由表面。整个计算过程中由速度初值及猜测压力值试算速度场的过程并不参与迭代，因而是一场迭代，其计算步骤如下：

1）由Navier－Stoks方程式（3.5-100）～式（3.5-102），以初始条件或前一时刻值为基础，计算当前时刻的试算速度；

2）根据连续性方程式（3.5-103）对于每一个单元定义散度$D_{i,j,k}$：

$$D_{i,j,k}=\frac{u_{i+\frac{1}{2},j,k}-u_{i-\frac{1}{2},j,k}}{\delta x_i}+\frac{v_{i,j+\frac{1}{2},k}-v_{i,j-\frac{1}{2},k}}{\delta y_i}+\frac{w_{i,j,k+\frac{1}{2}}-w_{i,j,k-\frac{1}{2}}}{\delta z_k} \tag{3.5-107}$$

将第一步的试算速度值代入上式，求出$D_{i,j,k}$；

3）若$D_{i,j,k}=0$（一般当$D_{i,j,k}<10^{-3}$，即认为$D_{i,j,k}=0$），则说明第一步试算速度值能够满足连续性方程式（3.5-103），即此时的速度场与压力场值既满足动量守恒方程又满足质量守恒方程，至此，当前时间步长计算结束。如整个流场中有任一个单元不能满足式（3.5-103），则需要下一步的修正；

4）当$D_{i,j,k}\neq0$时，说明第一步试算的速度值不能满足连续性方程，需要修正，而欲修正速度，必须先修正压力：

$$P^{n+1}=P^n+\delta P^n \tag{3.5-108}$$

其中δP^n为压力修正量，其值可以用下式求得：

$$\delta P^n=-D_{i,j,k}\Big/\frac{\partial D_{i,j,k}}{\partial P} \tag{3.5-109}$$

$$\frac{\partial D_{i,j,k}}{\partial P}=\frac{\delta t}{\delta x_i}\left(\frac{1}{\rho\delta x_{i+\frac{1}{2}}}+\frac{1}{\rho\delta x_{i-\frac{1}{2}}}\right)+\frac{\delta t}{\delta y_j}\left(\frac{1}{\rho\delta y_{j+\frac{1}{2}}}+\frac{1}{\rho\delta y_{i-\frac{1}{2}}}\right)+\frac{\delta t}{\delta z_k}\left(\frac{1}{\rho\delta z_{k+\frac{1}{2}}}+\frac{1}{\rho\delta z_{k-\frac{1}{2}}}\right) \tag{3.5-110}$$

根据校正压力利用下式可以求出校正后的试算速度：

$$\begin{aligned}
u'^{n+1}_{i+\frac{1}{2},j,k}&=u'^{n}_{i+\frac{1}{2},j,k}+\frac{\delta t\delta P^n\omega}{\rho\delta x_{i+\frac{1}{2}}}\\
u'^{n+1}_{i-\frac{1}{2},j,k}&=u'^{n}_{i-\frac{1}{2},j,k}-\frac{\delta t\delta P^n\omega}{\rho\delta x_{i-\frac{1}{2}}}\\
v'^{n+1}_{i,j+\frac{1}{2},k}&=v'^{n}_{i,j+\frac{1}{2},k}+\frac{\delta t\delta p^n\omega}{\rho\delta y_{j+\frac{1}{2}}}\\
v'^{n+1}_{i,j-\frac{1}{2},k}&=v'^{n}_{i,j-\frac{1}{2},k}-\frac{\delta t\delta p^n\omega}{\rho\delta y_{j-\frac{1}{2}}}\\
w'^{n+1}_{i,j,k+\frac{1}{2}}&=w'^{n}_{i,j,k+\frac{1}{2}}+\frac{\delta t\delta p^n\omega}{\rho\delta z_{k+\frac{1}{2}}}\\
w'^{n+1}_{i,j,k-\frac{1}{2}}&=w'^{n}_{i,j,k-\frac{1}{2}}-\frac{\delta t\delta p^n\omega}{\rho\delta z_{k-\frac{1}{2}}}
\end{aligned} \tag{3.5-111}$$

上式中，ω为松弛因子（$0<\omega<2$），n与$n+1$表示校正循环次数；

5）将校正后的试算速度值代入第二步，反复迭代直至所有单元均满足连续性方程；

6）由体积函数方程确定新的流动域，对表面单元做合理设置；

7）返回第一步，进入下一时刻计算，直到流动结束或达到要求。

可以看见，虽然SOLA－VOF是一场迭代，但每次收敛计算过程仍需很多次迭代，特别是计算分析带有自由表面复杂三维流动时，可能需要几千万甚至很多亿次的迭代。由于计算量大，流动场计算一般比较耗时，很多科研工作者正在加快流动场计算收敛速度、缩短流动场计算时间上作相应的研究。

2.5　初始条件与边界条件

2.5.1　初始条件

1）速度初始条件　要根据浇注方式，准确设置浇口处的初始速度值。

2）压力初始条件　当铸型排气条件良好时，铸型内背压可以认为是零，否则应根据背压的计算将初始压力设置为背压。

2.5.2　边界条件

1）自由表面速度边界条件　处理流动场表面区域时，动量守恒方程（Navier－Stokes方程）依旧可应用，但是连续性方程因流动域的变化而不再适用。此时的速度和压力条件必须人为加以设置，而自由表面速度边界条件设置的准确性直接影响到流动场计算的结果，不全面或不正确的速度边界条件将会使计算结果出现人为的不对称性（见图3.5-12所示），甚至导致计算崩溃，因此全面准确设置自由表面速度边界条件是进行流动场模拟的一个关键环节。

一般可采用惯性原理和连续原理相结合的方法，对多达64种情况进行设置。该方法的主要思路是：在同一个网格单元内，如果流体不遇到阻碍，则保持原有的方向和流速继续前进（惯性原理）；如果遇到了阻碍，则将其流量均匀地转移到其他可能的出流方向去，然后再调整出流速度，使新生成的满网格能够满足连续性方程。

2）自由表面压力边界条件　自由表面压力边界条件的设置是指如何处理作用于自由表面压力问题。就铸造而言，自由表面压力主要由两部分组成：一部分是由铸型（模具）排气不畅产生的背压而引起的，另一部分则是液态金属的表面张力。

铸造成形中背压的大小与铸型（模具）的排气条件、铸件的结构及浇注速度等因素紧密相关，实际生产中可用压力传感器来测定背压的大小。

表面张力的存在对金属液充型、流动过程的确有一定的

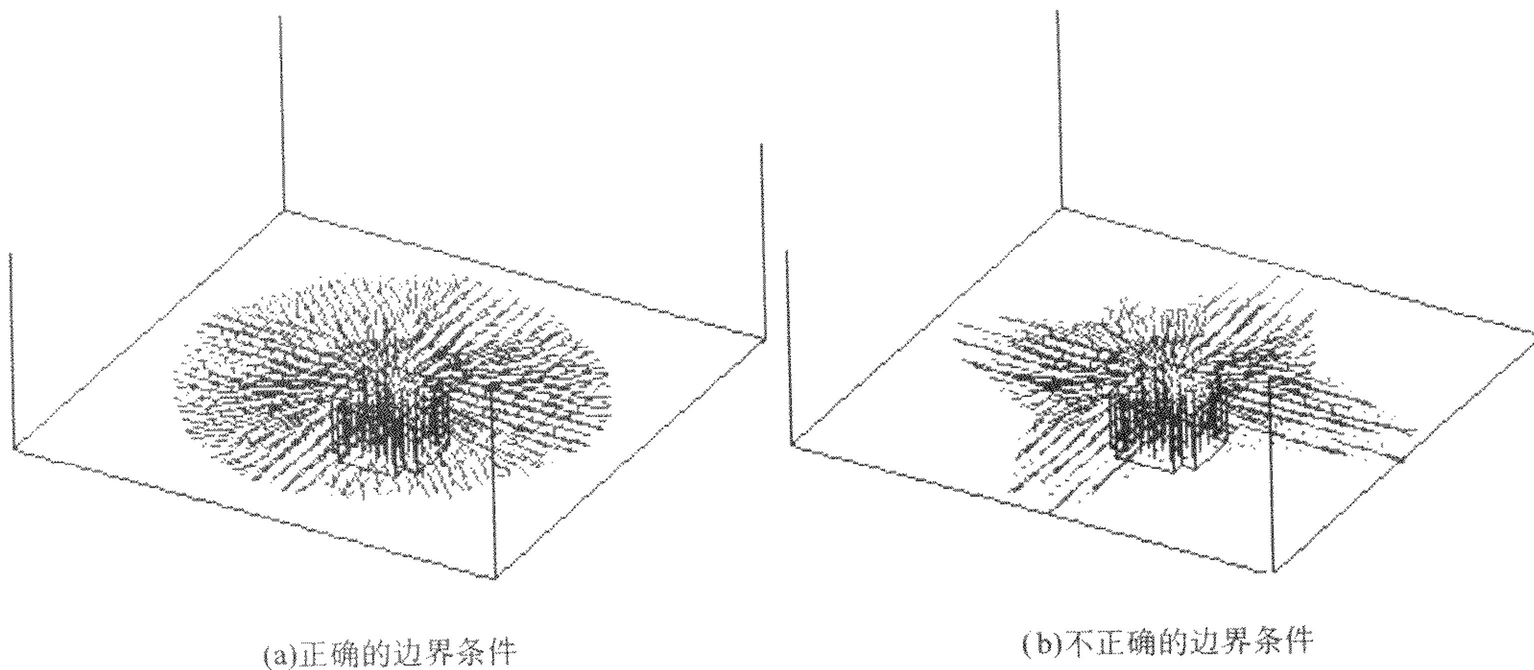


图 3.5-12 速度边界条件设置比较

影响，一些文献也都提及这方面的问题。为了模拟表面张力的作用，一般在动量守恒方程（Navier - Stoks）中加进以下表面力 F_s：

$$F_S = \sum (W\sigma\overline{m})_k \quad (3.5\text{-}112)$$

式中，W 为型腔的厚度，m；σ 为表面张力，N/m；$\overline{m}$ 为自由表面的切向单位矢量。

目前条件下，这个问题在二维条件下已被很好地解决，但有待于进一步发展到三维流场中去。此外，还有必要收集一些熔体表面张力和与型壁接触角的数据，这些数据因受到金属液表面氧化膜层以及金属液与型壁涂料之间交互作用的影响而难以准确获得，从而也限制这方面数值模拟的发展。

既然表面张力现象的模拟存在上述问题，能否将它忽略掉呢?

表面张力引起的压力差 ΔP 与曲率的关系可由下式来描述：

$$\Delta P = \frac{\sigma}{R} \quad (3.5\text{-}113)$$

式中，R 为曲率半径，m；σ 为金属液的表面张力，N/m。

$$R = \frac{L/2}{\cos\theta} \quad (3.5\text{-}114)$$

式中，θ 为金属与铸型的接触角；L 为厚度，m。

为了衡量表面张力的影响程度，将 ΔP 与液体金属内部静压强进行比较。

$$\frac{\Delta P}{P_a} = \frac{\Delta P}{\rho gh} = \frac{\sigma\cos\theta}{0.5L\rho gh}$$

对于铝合金铸件，当 L 为 2.0 cm，$\cos\theta = 0.5$，$\sigma = 0.8$ N/m，$\rho = 2.6\times10^3$ kg/m^2，$g = 9.8$ m/s^2 时：

$$\frac{\Delta P}{P_a} = \frac{1}{h}\times1.56\times10^{-3}$$

可见对于厚度 2.0 cm 这样的薄壁件，充型过程中因表面张力引起的 ΔP 与液体金属自身所产生的静压强 P_a 相比，要小 2 到 3 个数量级，为了节省内存，加快处理速度，使流动场计算能在微机上进行，一般情况下可以省去表面张力的作用，只将自由表面压力设置成型内背压。

3）型壁速度边界条件　在处理液态金属与铸型（模具）相邻的液体单元的速度时，必须通过假想网格层设置速度边界条件。SOLA - VOF 给出了两种极端条件下的边界条件：自由滑动边界和无滑动边界。见图 3.5-13 所示，对于自由滑动边界的情况：

$$u_1 = u \qquad v_b = 0 \qquad w_1 = w \quad (3.5\text{-}115)$$

无滑动边界条件：

$$u_1 = -u \qquad v_b = 0 \qquad w_1 = -w \quad (3.5\text{-}116)$$

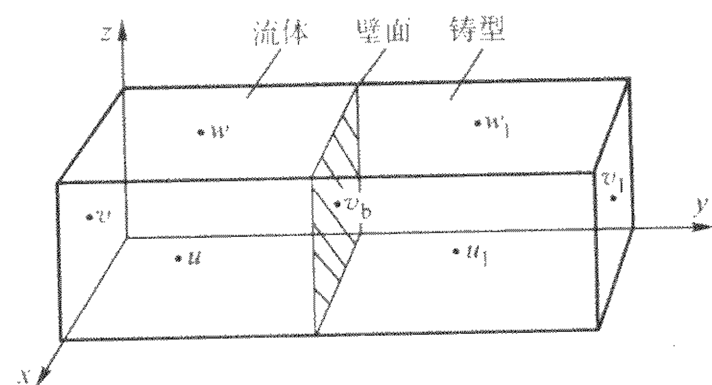


图 3.5-13 型壁速度边界条件

上述两种边界条件，前者认为型壁对液体的流动没有任何阻碍，能够自由滑动；后者认为液体在型壁上根本不能流动。显而易见，实际铸造充型过程并非上述的两种极端情况，而是处于两者之间，为此，定义了一个系数 θ，实际铸造条件下型壁速度边界条件可由下式描述：

$$u_1 = \theta u \qquad v_b = 0 \qquad w_1 = \theta w \quad (3.5\text{-}117)$$

当 $\theta = 1$ 时式（3.5-117）变为式（3.5-115），为自由滑动边界条件；当 $\theta = -1$ 时式（3.5-117）变为式（3.5-116），则为无滑动边界条件。因此实际铸造条件下 θ 的取值范围为 $-1 < \theta < 1$，具体取决于铸型（模具）、涂料及网格相对于速度边界层厚度的大小等因素。

2.6 数值稳定性条件

在利用 SOLA - VOF 方法进行流动场计算时，为了缩短计算时间，时间步长越大越好。但时间步长不能任意大，否则将会导致流动场计算发散，而使整个计算失败，因此要对时间步长进行限制。

时间步长的选择要考虑两个方面的因素，即满足物理方面和数学方面的约束。在物理方面，首先，一个时间步长内液体的流动不能超过一个单元，这是因为差分公式只适用于相邻单元之间，因此必须满足式（3.5-118）；其次，在同一个时间步长内，动量扩散同样不能超过一个单元，也必须满足式（3.5-119）。数学方面，迭代时，权重因子必须控制在一定范围之内，应满足式（3.5-120）。

$$\delta t < \min\left\{\frac{\delta x_i}{\left|u_{i+\frac{1}{2},j,k}\right|}, \frac{\delta y_j}{\left|v_{i,j+\frac{1}{2},k}\right|}, \frac{\delta z_k}{\left|w_{i,j,k+\frac{1}{2}}\right|}\right\} \quad (3.5\text{-}118)$$

$$\frac{\mu}{\rho}\delta t < \min\left[\frac{3}{4}\left(\frac{\delta x_i^2\delta y_j^2\delta z_k^2}{\delta x_i^2\delta y_j^2 + \delta y_j^2\delta z_k^2 + \delta z_k^2\delta x_i^2}\right)\right] \quad (3.5\text{-}119)$$

$$\max\left\{\left|u_{i+\frac{1}{2},j,k}\frac{\delta t}{\delta x_i}\right|,\left|v_{i,j+\frac{1}{2},k}\frac{\delta t}{\delta y_j}\right|,\left|w_{i,j,k+\frac{1}{2}}\frac{\delta t}{\delta z_k}\right|\right\}<\alpha\leqslant 1 \tag{3.5-120}$$

上述三个公式即为流动场数值模拟稳定性条件。

2.7　流动与传热耦合计算

高温液态金属的充型过程总是伴随着热量的散失。巨大的热量散失导致液态金属温度过低，则会形成冷隔、欠浇等严重缺陷，使充型过程无法顺利完成。另一方面，随着热量的散失，温度的下降，金属液的流动特性会发生明显变化，密度、热容、热导率以及黏度因温度不同都有很大差异，这也就决定了液态金属在不同温度下有着不同的流动方式和形态。换句话说，模拟铸造充型过程仅限于理想化的流动场研究而没有考虑热量的散失是不够的，其计算结果与实际生产必然有一定距离，甚至大相径庭。因此，从这种意义上来讲，要准确模拟流动场，就必须考虑热量的散失，必须进行流动与温度耦合计算的研究。

与之相应，欲准确对铸造凝固过程进行数值计算，流动分析必不可少。以往温度场计算时，通常都假设一个理想的温度均匀的初始条件，或凭经验人为地设置初温，这显然会直接影响到计算的准确性。而流动与温度的耦合计算从根本上解决了这一问题，其得出的初始温度场保证了后续凝固过程模拟的准确性与可靠性。

流动与传热耦合计算的数学模型必须能够完整、准确描述液态金属的充型过程以及与之同时发生的换热过程。一般说来，数学模型由下述三个方程组成：

动量守恒方程（N－S方程）

$$\rho\frac{D\boldsymbol{v}}{Dt}=\mu\nabla^2\boldsymbol{v}-\nabla P+\rho\cdot\boldsymbol{G} \tag{3.5-121}$$

质量守恒方程（连续性方程）

$$\nabla\boldsymbol{v}=0 \tag{3.5-122}$$

能量守恒方程

$$c_p\rho\frac{\partial T}{\partial t}=-c_p\rho\boldsymbol{v}\cdot\nabla T-\nabla\cdot\boldsymbol{q} \tag{3.5-123}$$

式中，T 为温度，K；t 为时间，s；ρ 为密度，kg/m^3；c_p 为热容，J/(kg·K)；v 为速度，m/s。

同样可以采用 SOLA－VOF 法求解式（3.5-121）、式（3.5-122），得到求解对象的速度、压力分布，再根据速度分布求解式（3.5-123）的能量守恒方程。

在求解能量守恒方程时，采用美国著名教授帕坦卡提出的幂函数方法，在与充型流动计算一致的交错网格上（见图3.5-11）进行离散，温度变量（T）与压力变量（P）以及体积函数（F）放在网格中心，温度的离散形式如下：

$$\begin{aligned}T_{i,j,k}^{t+\delta t}=(&\alpha_{i+1,j,k}T_{i+1,j,k}^t+\alpha_{i-1,j,k}T_{i-1,j,k}^t+\\&\alpha_{i,j+1,k}T_{i,j+1,k}^t+\alpha_{i,j-1,k}T_{i,j-1,k}^t+\alpha_{i,j,k+1}T_{i,j,k+1}^t+\\&\alpha_{i,j,k-1}T_{i,j,k-1}^t+T_{i,j,k}^tB)/\alpha_{i,j,k}\end{aligned} \tag{3.5-124}$$

其中：

$$\alpha_{i+1,j,k}=D_{i+\frac{1}{2},j,k}A\left(\left|P_{i+\frac{1}{2},j,k}\right|\right)+\max(-Q_{i+1/2,j,k},0)$$

$$\alpha_{i-1,j,k}=D_{i-\frac{1}{2},j,k}A\left(\left|P_{i-\frac{1}{2},j,k}\right|\right)+\max(Q_{i-1/2,j,k},0)$$

$$\alpha_{i,j+1,k}=D_{i,j+\frac{1}{2},k}A\left(\left|P_{i,j+\frac{1}{2},k}\right|\right)+\max(-Q_{i,j+1/2,k},0)$$

$$\alpha_{i,j-1,k}=D_{i,j-\frac{1}{2},k}A\left(\left|P_{i,j-\frac{1}{2},k}\right|\right)+\max(Q_{i,j-1/2,k},0)$$

$$\alpha_{i,j,k+1}=D_{i,j,k+\frac{1}{2}}A\left(\left|P_{i,j,k+\frac{1}{2}}\right|\right)+\max(-Q_{i,j,k+\frac{1}{2}},0)$$

$$\alpha_{i,j,k-1}=D_{i,j,k-\frac{1}{2}}A\left(\left|P_{i,j,k-\frac{1}{2}}\right|\right)+\max(Q_{i,j,k-\frac{1}{2}},0)$$

$$\alpha_{i,j,k}=\rho_{i,j,k}c_{pi,j,k}\delta x\delta y\delta z/\delta t$$

$$B=\alpha_{i,j,k}-\alpha_{i+1,j,k}-\alpha_{i-1,j,k}-\alpha_{i,j+1,k}-\alpha_{i,j-1,k}-\alpha_{i,j,k+1}-\alpha_{i,j,k-1} \tag{3.5-125}$$

式中，t 为计算时间，s；δt 为计算时间步长，s；D 为传导项；$A(|P|)$为幂函数；Q 为对流项。

其中传导项 D 的表达式如下：

$$\begin{aligned}D_{i+\frac{1}{2},j,k}&=\lambda_{i+\frac{1}{2},j,k}\delta y\delta z/\delta x\\D_{i-\frac{1}{2},j,k}&=\lambda_{i-\frac{1}{2},j,k}\delta y\delta z/\delta x\\D_{i,j+\frac{1}{2},k}&=\lambda_{i,j+\frac{1}{2},k}\delta x\delta z/\delta y\\D_{i,j-\frac{1}{2},k}&=\lambda_{i,j-\frac{1}{2},k}\delta x\delta z/\delta y\\D_{i,j,k+\frac{1}{2}}&=\lambda_{i,j,k+\frac{1}{2}}\delta x\delta y/\delta z\\D_{i,j,k-\frac{1}{2}}&=\lambda_{i,j,k-\frac{1}{2}}\delta x\delta y/\delta z\end{aligned} \tag{3.5-126}$$

Q 值由下述式得出：

$$\begin{aligned}Q_{i+\frac{1}{2},j,k}&=(\rho u)_{i+\frac{1}{2},j,k}c_{p_i,j,k}\delta y\delta z\\Q_{i-\frac{1}{2},j,k}&=(\rho u)_{i-\frac{1}{2},j,k}c_{pi,j,k}\delta y\delta z\\Q_{i,j+\frac{1}{2},k}&=(\rho v)_{i,j+\frac{1}{2},k}c_{pi,j,k}\delta x\delta z\\Q_{i,j-\frac{1}{2},k}&=(\rho v)_{i,j-\frac{1}{2},k}c_{pi,j,k}\delta x\delta z\\Q_{i,j,k+\frac{1}{2}}&=(\rho w)_{i,j,k+\frac{1}{2}}c_{pi,j,k}\delta x\delta y\\Q_{i,j,k-\frac{1}{2}}&=(\rho w)_{i,j,k-\frac{1}{2}}c_{pi,j,k}\delta x\delta y\end{aligned} \tag{3.5-127}$$

幂函数$A(|P|)$由下式定义：

$$A(|P|)=\max[0,(1-0.1|P|)^5] \tag{3.5-128}$$

其中 P 为贝克利脱数（Peclet number）：

$$P=UL/a \tag{3.5-129}$$

式中，U 为流体流动速度，m/s；L 为距离，m；a 为热扩散率，m^2/s。

贝克利脱数 P 用来衡量流动传热作用与热传导传热作用的强度比。

上述离散公式具有很明确的物理意义：D 代表热传导作用，Q 代表流动传热的作用，幂函数$A(|P|)$用来调节二者的比例。当流体流动的速度很大时，由式（3.5-129）可以看出，此时的贝克利脱数 P 值很大，幂函数$A(|P|)$趋于零，也就是说，此时热传导传热的作用很小，而流动传热的作用占主导地位。符号 max 体现上风差分格式的思想，可以确保离散系数不会出现负值而导致物理上不真实的解。当流体流动的速度很小时，贝克利脱数 P 值也很小，此时的幂函数$A(|P|)$趋于1，也就是说，此时热传导传热的作用很大，占主导地位，而流动传热则很少。如果出现极端情况，流体流动速度（U）为0，则 P 值为0，而幂函数$A(|P|)=1$。此时流动传热根本不存在，只有热传导传热，即为纯导热问题。

当然流动与传热耦合计算还需处理边界条件、初始条件、潜热问题、稳定性问题等等，因篇幅所限，这里不再详述。

2.8　流动场数值模拟流程图

与温度场模拟分析软件类似，流动场模拟软件亦分为三大部分，即前置处理、计算部分及后置处理。其中前、后处理与温度场模拟软件大致类似，而计算分析部分有很大差异。

图3.5-14为流动场计算分析的流程图。

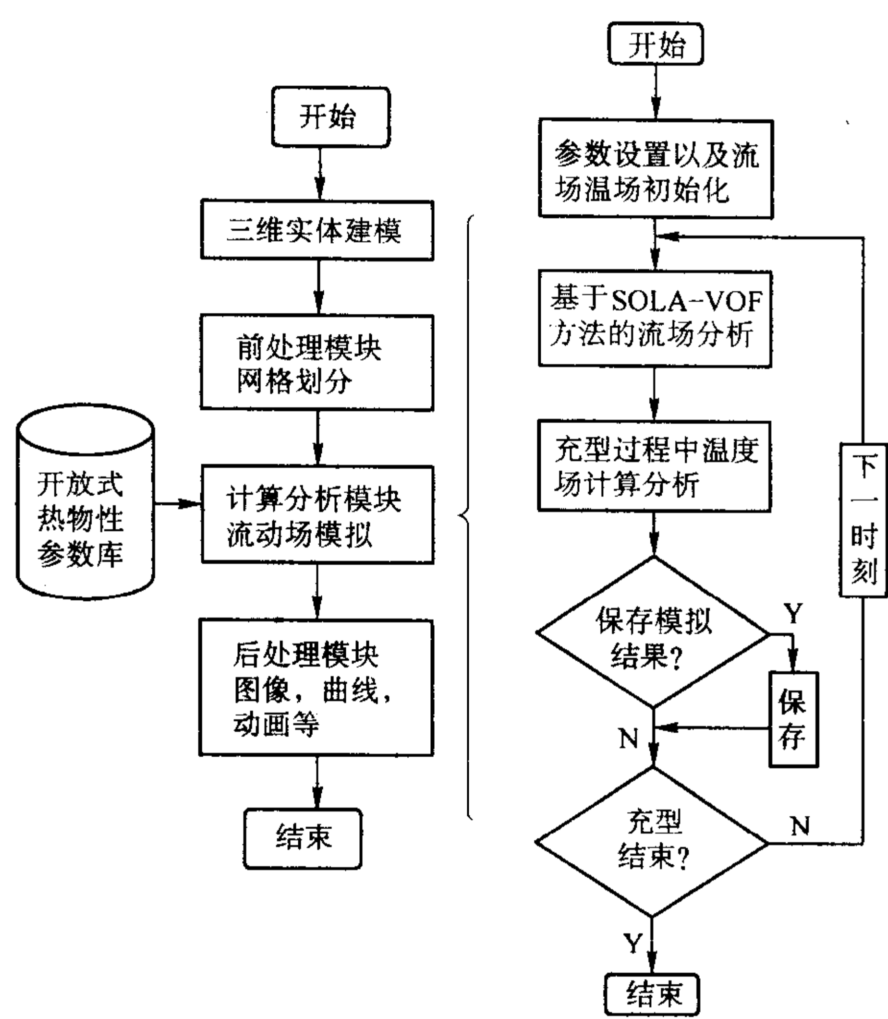


图 3.5-14　三维流动场计算分析的流程图

编写：周建新（华中科技大学）

第6章　塑料注射成形数值模拟方法

1　注射成形CAE概述

1.1　注射成形CAE的概念

CAE（计算机辅助工程）所包含的内容十分广泛。CAE将工程设计、试验、分析乃至制造贯穿于产品研制过程的每一个环节之中，以计算机为辅助工具来指导和预测产品在构思、设计与制造阶段的行为。目前，注射成形CAE限于注射过程的计算机分析，即模拟注射成形中熔体充模、保压与冷却过程以及预测塑料制品在脱模后的翘曲变形。

CAE技术借助于有限元法、有限差分法和边界元法等数值计算方法，分析型腔中塑料的流动、保压和冷却过程，计算制品和模具的应力分布，预测制品的翘曲变形，并由此分析工艺条件、材料参数及模具结构对制品质量的影响，达到优化制品和模具结构、优选成形工艺参数的目的。塑料注射成形CAE软件主要包括流动保压模拟、流道平衡设计、冷却模拟、翘曲预测等功能。其中流动保压模拟软件能提供不同时刻型腔内塑料熔体的温度、压力、切应力等的分布，其预测结果能直接指导工艺参数的选定及流道系统的设计；流道平衡设计软件能帮助用户对一模多腔模具的流道系统进行平衡设计，计算各个流道和浇口的尺寸，以保证塑料熔体能同时充满各个模腔；冷却模拟软件能计算冷却时间、制品及模腔的温度分布，其分析结果可以用来优化冷却系统的设计；应力计算和翘曲预测软件则能计算出制品的收缩情况和内应力的分布，预测制品出模后的变形。

1.2　注射成形CAE的发展概况

近30年来塑料流变学、计算机技术、计算数学、图形学等技术的突飞猛进，为塑料注射成形CAE技术的发展创造了条件。从20世纪70年代起，塑料注射成形CAE技术就成为塑料成形领域的热门研究课题，从分散、零星的研究发展为集中、系统的开发，研究机构层出不穷，如澳大利亚Moldflow公司、美国AC-Tech公司、德国IKV研究所、法国GISIGRAPH公司、台湾清华大学等。但在众多注射成形CAE产品中，最具国际影响、市场占有率最高的应属AC-Tech公司和Moldflow公司。1987年Moldflow公司率先推出商品化的二维流动模拟软件，在生产中发挥了显著作用，其CAE软件包括流动模拟、冷却分析、应力分析和翘曲变形预测等，该公司还开发出一套智能控制器（IPC），它先利用分析软件产生的结果对塑料流动情况进行优化，然后据此来控制注射机。该系统可在较短时间内自动地做好机器的启动准备，通过对充模、保压等工艺参数的优化来减小产品质量的波动。AC-Tech公司的注射成形CAE软件C-Mold包括：优选注射工艺材料，优化成形工艺参数，平衡流道系统设计，三维流动、保压、冷却分析程序以及应力分析和翘曲变形预测程序。该公司于2000年被Moldflow公司兼并。

我国自20世纪80年代以来陆续从国外引进了一批塑料注射成形CAE系统。实践表明这些软件对提高我国模具技术水平有一定的推动作用，但由于这些软件价格昂贵、二次开发能力差、缺乏国产材料数据库、技术支持和售后服务相对不足，因此软件的推广和应用情况并不理想。

我国的一些科研机构，特别是高等院校，自20世纪80年代中期开展注射成形CAE技术的研究，华中科技大学、上海交通大学、浙江大学、天津轻工业学院、成都科技大学等都先后取得了一批喜人的科研成果。如华中科技大学于1988年底研制出了注射模CAD/CAE/CAM系统的原型版本HSC1.0，该系统经过了十余年时间的推广，国内近百家工厂的验证。经过近20年的努力，华中科技大学目前已推出一套适用范围广、运行稳定、结果可靠、使用方便、并具有自主版权的基于实体模型的塑料注射成形模拟软件HSCAE3D。

1.3　注射成形CAE的发展趋势

塑料注射成形CAE技术不论从理论上还是在应用上都取得了长足进展，塑料注射成形模拟软件已有的发展可以划分为独立运行的注射成形过程模拟软件、二维模具设计软件与模拟软件的集成、三维模具设计制造软件与模拟软件的集成三个阶段。尽管注射成形CAE技术已成为计算机在工程应用中的典范，但在以下几个方面仍有待于进一步完善和发展。

1）数学模型和算法的进一步完善　目前的商品化模拟软件一般采用二维半方法，虽然考虑了制品中性层在三维空间的几何位置，但除了用有限差分法求解温度在壁厚方向的差异外，基本上没有考虑物理量在厚度方向上的变化。随着塑料注射成形工艺的发展，塑料制品的结构越来越复杂，壁厚差异越来越大，物理量在壁厚方向上的变化逐渐变得不容忽视，为了进一步提高软件的分析精度和使用范围，必须对目前的分析模型和算法做进一步完善，例如对充填过程进行真正的三维分析。

2）注射成形全过程模拟　目前，注射成形过程的模拟软件包括流动、保压、冷却、应力应变、翘曲分析等模块，各软件的研制与开发是基于各自独立的数学模型，这些模型在很大程度上进行了简化，忽略了相互之间的影响。但是，从塑料注射成形工艺过程来看，塑料注射成形过程中塑料熔体的充模流动、保压和冷却都是交织在一起并相互影响的，因此，充模流动、保压与冷却分析软件必须有机地结合起来，进行耦合分析，才能综合反映实际的成形过程。在此基础上，再通过应力分析、应变分析来预测制品的出模变形，开发翘曲分析软件。

3）塑料注射零件的虚拟制造　塑料注射成形计算机辅助设计单元技术日趋成熟与完善，但是现有的CAD、CAE软件仍然没有从根本上突破传统的模具设计与制造模式，在制造业全球性竞争更加激烈的形势下，必须在塑料注射成形单元技术的基础上，集成基于约束的塑料注射模装配关系模型和模具装配体的评价模型、三维成形过程模拟分析及基本模拟结果的评价与优化、三维模具工作过程仿真、支持注塑件与模具并行设计的多知识表达方式和处理方法，最终通过构造面向产品开发全过程的塑件及模具的特征模型，并以产品数据管理系统为总体集成框架，实现塑料注射零件的虚拟制造。塑料零件的虚拟制造能实现将塑件设计与模具设计紧密结合在一起的协同设计，为全面提高信息技术在塑料产品与模具设计中的应用水平提供了一条可行途径，促进产品开发周期和开发成本的最小化，推动塑料注射零件及模具的设计向可持续、知识化、数字化、可视化方向发展。

4）进一步拓宽CAE技术在塑料成形其他领域中的应用　气体辅助注射成形是国外20世纪80年代开始使用的一种新工艺，它比普通注射成形具有更多的工艺参数，同时由于气体的参与，塑料熔体的流动变得更加复杂，许多模具设计师由于缺乏经验和设计准则的指导，无法熟练地使用这种技

术。因此，气体辅助注射成形流动模拟成为目前世界各家注射成形 CAE 软件公司和大学、研究机构的研究热点。除此以外，热固性塑料的注射成形 CAE、热流道系统 CAE 分析技术的研究也方兴未艾。

5）知识工程在塑料注射成形中的应用　高层次智能化的 CAE 软件应具有思维和推理能力，能完成一系列需要设计人员经验的过程，如自动提供工艺条件的初始参数，根据模拟结果诊断制品缺陷等。目前，国外已有不少公司致力于这方面的研究，并相继展示了他们的成果，如美国 AC - Tech 公司的流动模拟软件使用了启发式知识来自动判断熔接缝和气穴的数目和位置，并给出一些解释和建议。澳大利亚 Moldflow 公司推出的一套用于联机分析处理成形过程的控制系统 IPC（Intelligent Process Control），能将模拟结果和实时工艺参数的测量结果进行比较，并运用专家系统监测产品的质量，调整成形参数。国内华中科技大学应用人工神经网络和专家系统技术来评价模拟结果，提出改进意见，生成分析报告。上海交通大学在这方面也进行了较深入的研究。

2　充模过程模拟

2.1　充模过程的数学描述

注射成形充模过程属非牛顿流体、非等温、非稳态的流动与传热过程，满足黏性流体力学的基本方程，但该方程过于复杂，只有引入合理而必要的假设与简化，并结合特定的边界条件，才能采用数值方法求解。

2.1.1　充模过程的基本方程

1）广义牛顿定律，反映了一般工程问题范围内黏性流体的应力张量与应变速率张量之间的关系，数学表达式为本构方程：

$$[\tau]=2\eta[\varepsilon]-(p-\lambda\nabla\cdot V)[I] \qquad (3.6\text{-}1)$$

式中，$[\tau]$ 为应力张量；η 为流体的动力黏度系数；p 为流体的静压强；λ 为体胀黏度系数，也称为第二黏度系数；V 为速度矢量；$[I]$ 为单位张量；$[\varepsilon]$ 为应变速率张量。定义为：

$$[\varepsilon]=\begin{bmatrix}\varepsilon_x & \varepsilon_{xy} & \varepsilon_{xz}\\ \varepsilon_{yx} & \varepsilon_y & \varepsilon_{yz}\\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_z\end{bmatrix}=\begin{bmatrix}\frac{\partial u}{\partial x} & \frac{1}{2}(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}) & \frac{1}{2}(\frac{\partial w}{\partial x}+\frac{\partial u}{\partial z})\\ \frac{1}{2}(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}) & \frac{\partial v}{\partial y} & \frac{1}{2}(\frac{\partial w}{\partial y}+\frac{\partial v}{\partial z})\\ \frac{1}{2}(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}) & \frac{1}{2}(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}) & \frac{\partial w}{\partial z}\end{bmatrix} \qquad (3.6\text{-}2)$$

式中，u、v、w 为速度矢量 V 沿 x，y，z 三个坐标方向的分量。

本构方程的直角坐标形式为：

$$\sigma_x=-p+\lambda\nabla\cdot V+2\eta\frac{\partial u}{\partial x} \qquad (3.6\text{-}3a)$$

$$\sigma_y=-p+\lambda\nabla\cdot V+2\eta\frac{\partial v}{\partial y} \qquad (3.6\text{-}3b)$$

$$\sigma_z=-p+\lambda\nabla\cdot V+2\eta\frac{\partial w}{\partial z} \qquad (3.6\text{-}3c)$$

$$\tau_{xy}=\tau_{yx}=\eta\ (\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}) \qquad (3.6\text{-}3d)$$

$$\tau_{yz}=\tau_{zy}=\eta\ (\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}) \qquad (3.6\text{-}3e)$$

$$\tau_{zx}=\tau_{xz}=\eta\ (\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}) \qquad (3.6\text{-}3f)$$

式中，σ_x、σ_y、σ_z 为直角坐标系中 x，y，z 方向的正应力；τ_{xy}、τ_{yz}、τ_{zx} 为直角坐标系中的切应力。

2）质量守恒定律，其含义是流体的质量在运动过程中保持不变，数学表达式为连续性方程：

$$\frac{\partial p}{\partial t}+\nabla\cdot\ (\rho V)\ =0 \qquad (3.6\text{-}4)$$

式中，ρ 为密度；t 为时间。

在直角坐标系中上式可表示为：

$$\frac{\partial p}{\partial t}+\frac{\partial\ (\rho u)}{\partial x}+\frac{\partial\ (\rho v)}{\partial y}+\frac{\partial\ (\rho w)}{\partial z}=0 \qquad (3.6\text{-}5)$$

式（3.6-5）中的第一项代表单位时间内单位体积的质量增加，第二、三、四项合起来代表单位时间内单位体积内质量的净流出。

3）动量守恒定律，其含义是流体动量的时间变化率等于作用于其上的外力总和，数学表达式为运动方程：

$$\rho\frac{DV}{Dt}=\rho F+\nabla\cdot\ [\tau] \qquad (3.6\text{-}6)$$

式中，F 为单位体积流体的质量力。

将本构方程代入上式并写成直角分量形式为：

$$\rho\frac{Du}{Dt}=\rho F_x-\frac{\partial(p-\lambda\nabla\cdot V)}{\partial x}+\frac{\partial}{\partial x}(2\eta\frac{\partial u}{\partial x})+\frac{\partial}{\partial y}[\eta(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y})]+\frac{\partial}{\partial z}[\eta(\frac{\partial w}{\partial x}+\frac{\partial u}{\partial z})] \qquad (3.6\text{-}7a)$$

$$\rho\frac{Dv}{Dt}=\rho F_y-\frac{\partial\ (p-\lambda\nabla\cdot V)}{\partial y}+\frac{\partial}{\partial x}\ [\eta\ (\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y})]\ +\frac{\partial}{\partial y}\ (2\eta\frac{\partial v}{\partial y})\ +\frac{\partial}{\partial z}\ [\eta\ (\frac{\partial w}{\partial y}+\frac{\partial v}{\partial z})] \qquad (3.6\text{-}7b)$$

$$\rho\frac{Dw}{Dt}=\rho F_z-\frac{\partial\ (p-\lambda\nabla\cdot V)}{\partial z}+\frac{\partial}{\partial x}\ [\eta\ (\frac{\partial w}{\partial x}+\frac{\partial u}{\partial z})]\ +\frac{\partial}{\partial y}\ [\eta\ (\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y})]\ +\frac{\partial}{\partial z}\ [2\eta\ (\frac{\partial w}{\partial z})] \qquad (3.6\text{-}7c)$$

上面三式中等式左端代表单位体积流体的惯性力，等式右端第一项代表单位体积流体的质量力，第二项代表作用于单位体积流体的压力差，第三至第五项代表黏性体胀力和黏性变形力，它们只与流体的黏度系数和应变速率张量有关。

4）热力学第一定律，其含义是系统内能的增加等于对该系统所作的功与加给该系统的能量之和，数学表达式为能量方程：

$$\rho c_V\frac{DT}{Dt}=-P\nabla\cdot V+\nabla\cdot(K\nabla T)+\rho q+\Phi \qquad (3.6\text{-}8)$$

式中，c_V 为定容比热容，$c_V\approx c_p$（定压比热容）；T 为流体温度；q 为单位质量流体的热源强度；K 为流体导热系数。

上式等式左边为系统内能的增加，等式右边第一项为流体体积变化时压力所作的功，第二项为外界传给系统的能量，第三项为系统的内热源给系统提供的能量，第四项 Φ 为黏性耗散功，可表示为：

$$\Phi=\eta\Big[2\ (\frac{\partial u}{\partial x})^2+2\ (\frac{\partial v}{\partial y})^2+2\ (\frac{\partial w}{\partial z})^2+\ (\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x})^2+(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x})^2+\ (\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y})^2\Big]-\frac{2\eta}{3}\ (\nabla\cdot\vec{V})^2 \qquad (3.6\text{-}9)$$

将上式代入能量方程，并写成直角坐标形式为：

$$\rho c_p(\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}+w\frac{\partial T}{\partial z})=-P(\nabla\cdot\vec{V})+\eta\Big[2(\frac{\partial u}{\partial x})^2+2(\frac{\partial v}{\partial y})^2+2(\frac{\partial w}{\partial z})^2+(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x})^2+(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x})^2+(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y})^2\Big]-\frac{2\eta}{3}(\nabla\cdot\vec{V})^2+\frac{\partial}{\partial x}(K\frac{\partial T}{\partial x})+\frac{\partial}{\partial y}(K\frac{\partial T}{\partial y})+\frac{\partial}{\partial z}(K\frac{\partial T}{\partial z})+\rho q \qquad (3.6\text{-}10)$$

2.1.2 塑料熔体充模流动的简化和假设

1）由于型腔壁厚（z向）尺寸远小于其他两个方向（x和y方向）的尺寸，且塑料熔体黏性较大，所以熔体的充模流动可视为扩展层流，z向的速度分量可忽略不计，即 $w=0$，且认为压力不沿 z 向变化，即$\frac{\partial p}{\partial z}=0$。

2）充模过程中熔体压力不是很高，因此可视熔体为未压缩流体，即$\frac{\partial \rho}{\partial t}=\frac{\partial \rho}{\partial x}=\frac{\partial \rho}{\partial y}=\frac{\partial \rho}{\partial z}=0$，因而$\nabla \cdot \boldsymbol{V}=0$。

3）由于熔体黏性较大，相对于黏性切应力而言，惯性力和质量力都很小，可忽略不计，即 $\rho \frac{D\boldsymbol{V}}{Dt}=0$，$\rho \vec{F}=0$。

4）在熔体流动方向（x 和 y 方向）上，相对于热对流项而言，热传导项很小，可忽略不计，即$\frac{\partial}{\partial x}(K\frac{\partial T}{\partial x})=0$，$\frac{\partial}{\partial y}(K\frac{\partial T}{\partial y})=0$。

5）熔体不含内热源，即 $q=0$。

6）在充模过程中，熔体温度变化不大，可认为比热容 c_V、c_P 和导热系数 K 是常数。

7）熔体前沿采用平面流前模型，忽略喷泉效应的影响。

2.1.3 塑料熔体充模流动的控制方程及边界条件

利用上述假设和简化，可由黏性流体力学的基本方程导出塑料熔体充模流动的控制方程：

$$\frac{\partial}{\partial x}(b\bar{u})+\frac{\partial}{\partial y}(b\bar{v})=0 \tag{3.6-11a}$$

$$\frac{\partial}{\partial z}(\eta \frac{\partial u}{\partial z})-\frac{\partial P}{\partial x}=0 \tag{3.6-11b}$$

$$\frac{\partial}{\partial z}(\eta \frac{\partial v}{\partial z})-\frac{\partial P}{\partial y}=0 \tag{3.6-11c}$$

$$\rho c_p(\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y})=k\frac{\partial^2 T}{\partial z^2}+\eta\dot{\gamma}^2 \tag{3.6-11d}$$

式中，b 为型腔半厚；$\bar{u}$，$\bar{v}$ 为x 和y 方向的平均流速；其他符号与前述相同。

图 3.6-1 示出流动平面内的边界条件。在型腔界面 C_o 和型芯界面 C_i 上应满足无渗透边界条件，即

$$\frac{\partial p}{\partial n}=0 \quad 在\ C_o\ 和\ C_i\ 上 \tag{3.6-12}$$

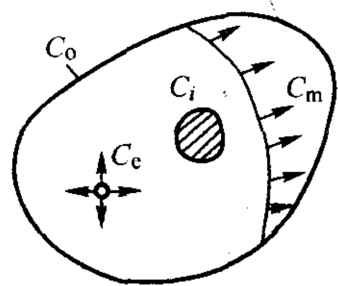


图 3.6-1 流动平面内的边界条件

通过熔体入口边界的塑料流量应等于注射机的流量，即当熔体入口边界 C_e 上压力均布时应满足

$$\int_{C_e} q\mathrm{d}l = Q(t) \tag{3.6-13}$$

式中，q 为边界 C_e 上的流量分布；$Q(t)$ 为 t 时刻的熔体注射流量。

熔体入口边界 C_e 上的温度为

$$T=T_e \tag{3.6-14}$$

式中，T_e 为熔体注射温度。

关于壁厚方向的边界条件，可假定为：

$$u=v=0,\ T=T_W \quad z=b\ 处 \tag{3.6-15}$$

$$\frac{\partial u}{\partial z}=\frac{\partial v}{\partial z}=0,\ \frac{\partial T}{\partial z}=0 \quad z=0\ 处 \tag{3.6-16}$$

式中，T_W 为模壁温度（如图 3.6-2 所示）。

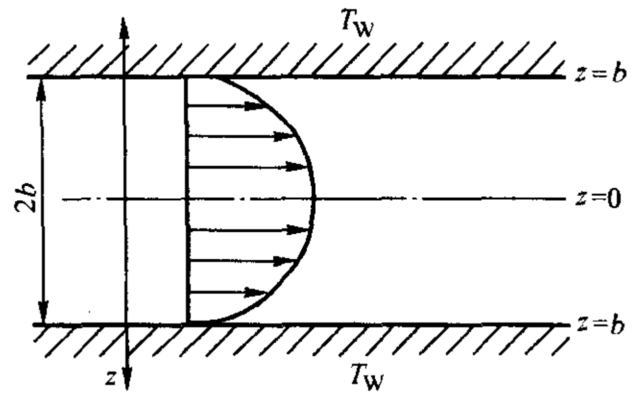


图 3.6-2 壁厚方向上的边界条件

2.1.4 塑料熔体的黏度模型

塑料材料的流变参数主要是确定聚合物的黏度与熔体压力、温度、剪切速率之间的定量关系，塑料熔体基本的流动性能用黏度来表征。目前常用的黏度数学模型主要是 Cross 模型：

$$\eta=\frac{\eta_0(T,P)}{1+(\eta_0\dot{\gamma}/\tau^*)^{1-n}} \tag{3.6-17}$$

式中，τ^* 为材料剪切常数；n 为非牛顿指数；$\dot{\gamma}$ 为切应变速率；η_0 为零切变黏度。η_0 一般采用 Arrhenius 型表达式（3.6-18）或 WLF 型表达式（3.6-19）：

$$\eta_0(T,P)=Be^{\beta P}e^{T_b/T} \tag{3.6-18}$$

$$\eta_0=D_1 e\left[\frac{-A_1(T-(D_2+D_3P))}{A_2+T-D_2}\right] \tag{3.6-19}$$

式（3.6-17）和式（3.6-18）构成了五参数（η，τ^*，B，T_b，β）黏度模型，式（3.6-17）和式（3.6-19）构成了七参数（n，τ^*，D_1，D_2，D_3，A_1，A_2）黏度模型。

特别需要指出的是，在 CAE 软件中，塑料材料的流变性能参数十分重要，不同的性能参数将导致完全不同的模拟结果。同时，塑料材料的性能又因品种不同、牌号不同、生产厂家不同、甚至批次不同而差异较大。因此，获得所用材料准确的性能参数是使用 CAE 软件的前提条件。

2.2 压力场的计算

分别对式（3.6-11b）、式（3.6-11c）积分，并利用边界条件式（3.6-15）和式（3.6-16），可得：

$$u=-\frac{\partial P}{\partial x}\int_0^z \frac{\bar{z}}{\eta}\mathrm{d}\bar{z} \tag{3.6-20a}$$

$$v=-\frac{\partial P}{\partial y}\int_0^z \frac{\bar{z}}{\eta}\mathrm{d}\bar{z} \tag{3.6-20b}$$

对式（3.6-20）沿 z 向积分，得到熔体的平均流速为：

$$\bar{u}=\frac{\partial P}{\partial x}\times\frac{S}{b} \tag{3.6-21a}$$

$$\bar{v}=\frac{\partial P}{\partial y}\times\frac{S}{b} \tag{3.6-21b}$$

式中，S 为流动率，其计算公式为：

$$S=\int_0^b \frac{(b-z)^2}{\eta}\mathrm{d}z \tag{3.6-22}$$

将式（3.6-21）代入连续性方程式（3.6-11a），得到如下的压力场的控制方程：

$$\nabla\cdot(S\nabla P)=0 \tag{3.6-23}$$

在数值计算实施之前，必须对型腔实体模型的表面进行有限元网格剖分。三角形单元具有对复杂边界逼近程度好、剖分算法成熟、无需等参转换等优点，因此本文采用三角形线性单元。三角形单元内的压力分布采用线性插值表示，用有限元法处理压力场控制方程式（3.6-23），并利用伽辽金法（Galerkin），可导出任一已充满节点的压力场有限元方程：

$$\sum_{l} S^{(l)} \sum_{j=1}^{3} D_{ij}^{(l)} P_{N'} = 0 \qquad (3.6\text{-}24)$$

式中，l 为遍历包含节点 N 的所有三角形单元；i 为单元 l 中对应于总体节点 N 的局部节点号；j 为单元 l 中对应于总体节点 N' 的局部节点号；$D_{ij}^{(l)}$ 为与单元 l 几何形状有关的系数。

对每个已被熔体充满的节点写出式（3.6-24），得到以熔体充填区域节点压力为未知量的代数方程组，由于整体刚度矩阵为一大型稀疏矩阵，并采用迭代法求解此方程组即可求得到各节点的压力分布。

2.3 熔体流动前沿位置的确定

在注射成形中熔体填充区域是在不断扩展的，熔体流动前沿是熔体填充区域的移动边界，需要在计算中确定其在流动方向上的位置，为此采用控制体积法处理。对每个节点的控制体积，引入熔体充填系数 f_{melt}，$f_{melt} = V_{melt}/V$，V 和 V_{melt} 分别表示节点的控制体积和控制体积中已被塑料熔体充填的部分的体积大小。在注射过程阶段，整个型腔可分为三个区域（如图 3.6-3 所示）：未被塑料熔体充填的控制体积，$f_{melt} = 0$；已被塑料熔体充填但尚未被完全充满的控制体积，即流前控制体积，$1 > f_{melt} > 0$；被塑料熔体完全充满的控制体积，$f_{melt} = 1$。熔体流动前沿的更新步骤为：在每一时刻的当前熔体流动前沿状况下，根据压力场的分布计算出流入每一流前节点控制体积的熔体流量率，算出在此流量率下充满每一未充满流前控制体积所需的时间，取最小时间为时间增量，计算在该时间增量过程中，流入各流前控制体积的熔体流量和各控制体积的充填状况，这时必有一个流前控制体积被充满，其他流前控制体积的熔体充填系数 f_{melt} 也在增加，这样就获得了新的熔体流动前沿。重复上述过程直至熔体注射结束。

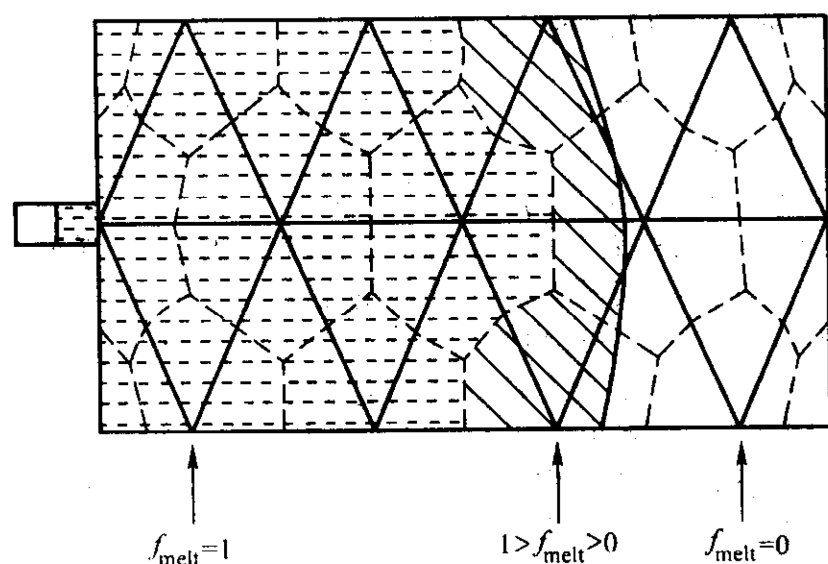


图 3.6-3　注射阶段的充填状况和移动边界

2.4 温度场数值求解

熔体温度在流动平面内以及沿型腔壁厚方向均发生变化。因此在求解温度场时，除了在流动平面划分有限元网格外，还需沿厚度方向划分差分网格，并通过有限元/有限差分混合法求解。如图 3.6-4 所示，将三角形单元沿厚度方向进行有限差分网格剖分，对每一个三角形单元在厚度方向共分成 NZ 层。

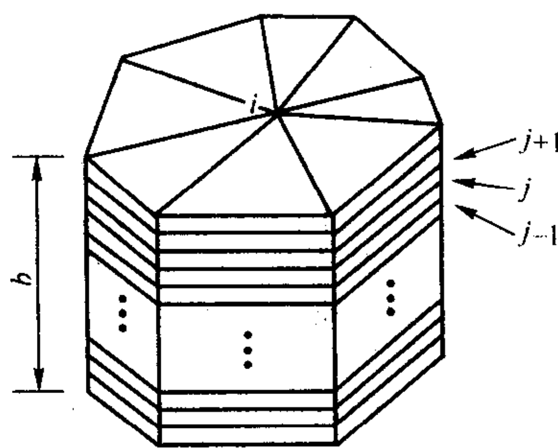


图 3.6-4　三角形单元有限差分层

对于控制方程中的能量方程，各项的有限元/有限差分形式分述如下。

1）瞬态项　瞬态项采用向后差分近似，即：

$$\rho c_p \left.\frac{\partial T}{\partial t}\right|_{N,j,t+1} = \rho c_p \frac{T_{N,j,t+1} - T_{N,j,t}}{\Delta t} \quad j = 1, 2, 3, \cdots, NZ \qquad (3.6\text{-}25)$$

2）热传导项　对三角形单元，有：

$$K \left.\frac{\partial^2 T}{\partial z^2}\right|_{N,j,t+1} = \frac{1}{\sum_{l} V_{i,j}^{(l)}} \sum_{l} \left\{ \frac{V_{i,j}^{(l)} K}{3\Delta z_j^2} \left[\sum_{m=1}^{3} (T_{N',j+1,t+1} - 2T_{N',j,t+1} + T_{N',j-1,t+1}) \right] \right\}$$

$$j = 1, 2, 3, \cdots, NZ \qquad (3.6\text{-}26)$$

式中，l 为遍历所有包含节点 N 的单元；N' 为单元，l 中局部节点号为 m 的节点的总体节点号，即 $N' = NOD\ (l,\ m)$，NOD 表示单元局部节点号与整体节点号的关系。

3）热对流项　为保证数值计算过程的稳定性，采用一种类似于向后差分的方法——"上风法（up wind）"来处理热对流项，即在热对流项进行加权平均计算时仅包含来自节点上游单元的贡献，可将热对流项的加权函数定义为：

$$\hat{V}_{i,j}^{(l)} = \begin{cases} V_{i,j}^{(l)} & \text{如果 } DOT > 0 \\ 0 & \text{如果 } DOT \leqslant 0 \end{cases} \qquad (3.6\text{-}27)$$

式中 DOT 为上下游判定系数，可采用下式来计算：

$$DOT = q = -S^{(l)} \sum_{j=1}^{3} D_{ij}^{(l)} P_{N'} \qquad (3.6\text{-}28)$$

热对流项的加权平均计算公式为：

$$\rho c_p \left.\left(u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y}\right)\right|_{N,j,t+1} = \frac{\rho c_p}{\sum_{l} \hat{V}_{i,j}^{(l)}} \sum_{l} \left\{ \hat{V}_{i,j}^{(l)} \frac{\int_{-b}^{z_j} \left(\frac{\tilde{z}}{\eta}\right)^{(l)} d\tilde{z}}{2A^{(l)}} \cdot \sum_{m=1}^{3} \left[\left(-\frac{\partial P}{\partial x}\right)^{(l)} b_{2m}^{(l)} + \left(-\frac{\partial P}{\partial y}\right)^{(l)} b_{3m}^{(l)} \right] \cdot T_{N',j,t+1} \right\}$$

$$j = 1, 2, 3, \cdots, NZ \qquad (3.6\text{-}29)$$

4）黏性热项　黏性热项的处理同热对流项，也采用上风法。

$$\eta \dot{\gamma}^2 \Big|_{N,j,t+1} = \frac{1}{\sum_{l} \hat{V}_{i,j}^{(l)}} \sum_{l} \hat{V}_{i,j}^{(l)} \left[z_j \sqrt{\left(\frac{\partial P}{\partial x}\right)^2 + \left(\frac{\partial P}{\partial y}\right)^2}^{(l)} \right]^2 \frac{1}{\eta^{(l)}}$$

$$j = 1, 2, 3, \cdots, NZ \qquad (3.6\text{-}30)$$

对熔体充填区域的每个差分节点写出式（3.6-25）~式（3.6-30），将各式代入温度控制方程式（3.6-11d），就得到以各差分节点的温度为未知量的线性方程组：

$$[K]\{T\} = \{F\} \qquad (3.6\text{-}31)$$

采用低松弛法迭代求解此方程组即可得到塑料熔体充填区域的温度场。

2.5　数值计算过程

注射成形充填的数值计算过程如图3.6-5所示。程序在开始时，需要读入的初始数据包括型腔形状、网格单元、塑料材料特性数据及流动工艺条件等参数。开始分析时，假定塑料熔体入口的第一个控制体积已被熔体充满。此时可以近似认为，熔体仍处于等温状态，其温度即为熔体的入口温度，这样便可获得初始时刻熔体的前沿位置和温度场，然后再求解压力场。利用压力场的计算结果，更新流动前沿位置（增加一个控制体积），确定时间增量，求解新时刻的温度场、压力场，如此循环下去，直至整个型腔被熔体充满为止。

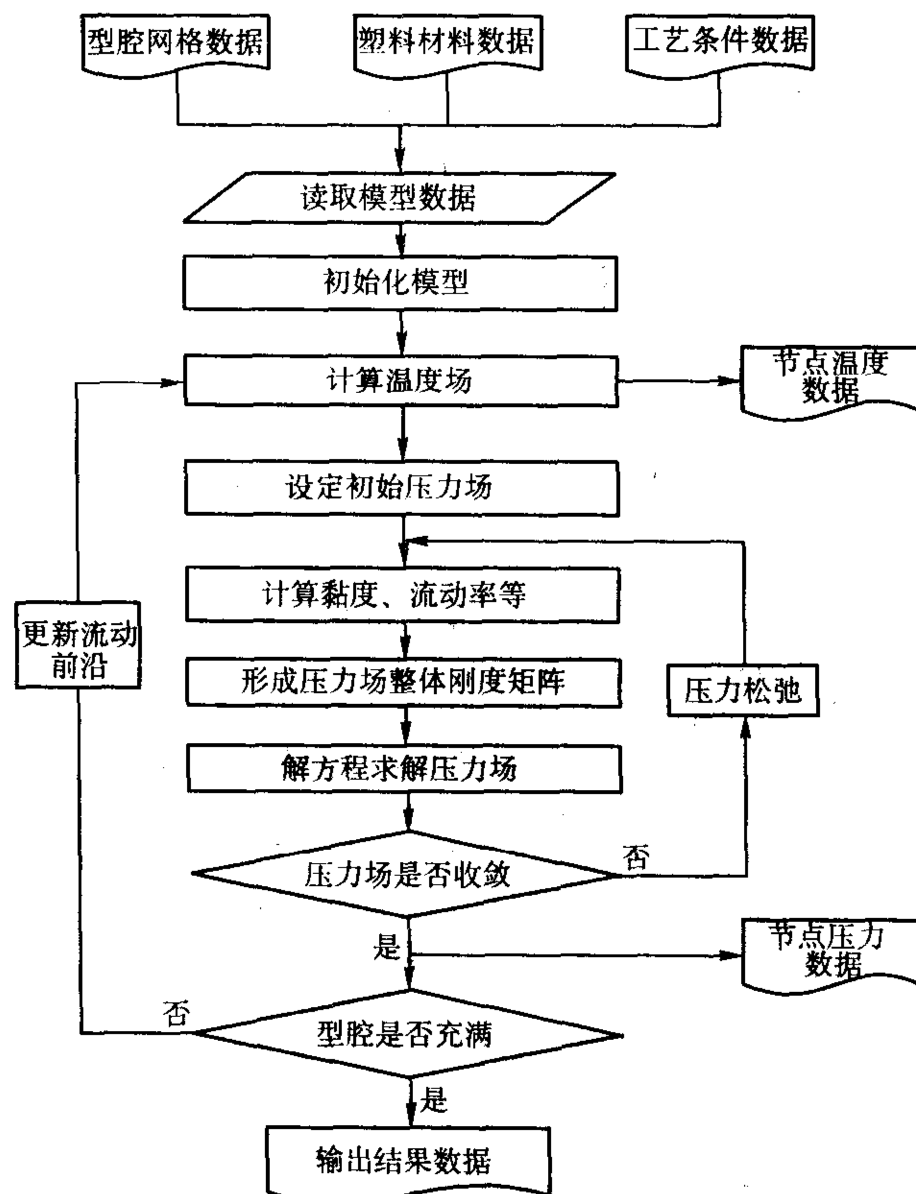


图3.6-5　注射成形充填分析的数据逻辑流程

3　保压过程模拟

3.1　保压模拟的重要性

由于塑料从融熔态冷却到固态时体积变化很大，可达25%，因此熔体充满型腔后仍需要有一部分塑料熔体在高压下被继续注入型腔，以补充熔体因冷却而引起的收缩。若保压不足（保压压力不够或保压时间太短），制品会由于收缩而形成凹陷、缩孔或收缩过大等缺陷；反之，若保压压力太高，一方面制品可能会产生飞边，另一方面制品的内应力增大，可能导致制品发生过量的翘曲变形。因此，在对充模流动过程进行模拟之后，必须对保压阶段进行深入的研究，预测保压阶段型腔内熔体压力、温度、密度、切应力等的变化，为合理确定保压压力、保压时间、浇口尺寸、注射温度等提供科学依据。

3.2　保压过程的数学模型

保压模拟的实质是求解可压缩、非牛顿流体的非等温流动问题，其分析原理与流动模拟类似，但有以下几点不同。

1）在充模流动阶段，熔体密度变化很小，模拟中假定熔体是未被压缩的。而在保压模拟中，熔体密度变化相对较大，这个假设不再成立。实质上，保压过程正是利用熔体的可压缩性来补入新的塑料，故保压模拟中必须考虑熔体密度的变化。

2）由于增加了密度参量，必须在保压模拟中引入状态方程（$P-V-T$ 关系）才能求解。

3）在保压过程中，熔体温度变化范围较大，因此，必须采用适应范围更广的黏度模型，并考虑熔体的比热容、导热系数随温度的变化。

4）充模流动模拟时熔体流量是已知的，压力为待求量，而在保压中入口压力为设定的保压压力，需求熔体的流量，两者的边界条件不同。

类似于充模流动模拟中的推导，考虑熔体的密度以及比热容和导热系数的变化，可导出保压过程的控制方程：

$$\frac{\partial \rho}{\partial t}+\frac{\partial}{\partial x}(\rho u)+\frac{\partial}{\partial y}(\rho v)+\frac{\partial}{\partial z}(\rho w)=0 \qquad (3.6\text{-}32a)$$

$$\frac{\partial P}{\partial x}-\frac{\partial}{\partial z}\left(\eta\frac{\partial u}{\partial z}\right)=0 \qquad (3.6\text{-}32b)$$

$$\frac{\partial P}{\partial y}-\frac{\partial}{\partial z}\left(\eta\frac{\partial v}{\partial z}\right)=0 \qquad (3.6\text{-}32c)$$

$$\rho c_p(T)\left(\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}\right)=\frac{\partial}{\partial z}\left[K(T)\frac{\partial T}{\partial z}\right]+\eta\left[\left(\frac{\partial u}{\partial z}\right)^2+\left(\frac{\partial v}{\partial z}\right)^2\right] \qquad (3.6\text{-}32d)$$

式中，ρ 为密度；w 为 z 方向的速度分量；

其他参数与充模流动模拟中相同。

经过与充模流动的类似推导，可得如下的压力场控制方程：

$$G\frac{\partial P}{\partial t}-\frac{\partial}{\partial x}\left(\overline{S}\frac{\partial P}{\partial x}\right)-\frac{\partial}{\partial y}\left(\overline{S}\frac{\partial P}{\partial y}\right)=-F \qquad (3.6\text{-}33)$$

式中

$$\overline{S}=\int_0^b\rho\left[-\int_0^z\frac{\overline{z}}{\eta}\mathrm{d}\overline{z}\right]\mathrm{d}z \qquad (3.6\text{-}34)$$

$$G=\int_0^{\hat{z}}\left(\frac{\partial\rho_s}{\partial P}\right)_T\mathrm{d}z+\int_{\hat{z}}^b\left(\frac{\partial\rho_l}{\partial P}\right)_T\mathrm{d}z \qquad (3.6\text{-}35)$$

$$F=\int_0^{\hat{z}}\left(\frac{\partial\rho_s}{\partial T}\right)_P\frac{\partial T}{\partial t}\mathrm{d}z+\int_{\hat{z}}^b\left(\frac{\partial\rho_l}{\partial T}\right)_P\frac{\partial T}{\partial t}\mathrm{d}z+(\rho_s-\rho_l)_{z=\hat{z}}\frac{\partial\hat{z}}{\partial t} \qquad (3.6\text{-}36)$$

式中，ρ_l、ρ_s 为液、固相的密度；$\hat{z}$ 为固－液相交界面位置。

3.3　塑料熔体的特性模型

在充模流动模拟中，一般采用了五参数黏度模型，即式（3.6-17）和式（3.6-18）。从式（3.6-18）可以看出，参数 T_b 反映了零切变黏度 η_0 对温度的敏感程度，该参数与熔体温度有关，尤其是当熔体温度接近玻璃化温度时，关联性更强。一般而言，充模过程中熔体的温度远高于玻璃化温度，T_b 可认为不变，因此可采用五参数模型。但在保压模拟中，熔体可以从注射温度冷却到玻璃化温度左右，温度变化范围大，因此 η_0 不能再采用 Arrhenius 型表达式，WLF 型表达式对温度的适应范围更广，应采用由式（3.6-17）和（3.6-19）构成的七参数黏度模型。

保压过程中的 $P-V-T$ 关系一般采用 Tait 提出的经验公式：

$$V(P,T)=V_0(T)\left[1-0.0894\times\ln\left(1+\frac{P}{B(T)}\right)\right]-V_t(T,P) \qquad (3.6\text{-}37)$$

$$V_0(T)=\begin{cases}b_{1,l}+b_{2,l}(T-b_5) & T\geqslant T_g\\ b_{1,s}+b_{2,s}(T-b_5) & T<T_g\end{cases} \qquad (3.6\text{-}38)$$

$$B(T)=\begin{cases}b_{3,l}e^{-b_{4,l}(T-b_5)} & T\geqslant T_g\\ b_{3,s}e^{-b_{4,s}(T-b_5)} & T<T_g\end{cases} \qquad (3.6\text{-}39)$$

$$V_t(T,P)=\begin{cases}0 & T\geqslant T_g\\ b_7 e^{b_8(T-b_5)-b_9 P} & T<T_g\end{cases} \tag{3.6-40}$$

式中，T_g 为玻璃化温度（对非结晶型材料）或结晶温度（对结晶型材料），可认为 T_g 是压力的线性函数：

$$T_g(P)=b_5+b_6 P \tag{3.6-41}$$

上述各式中 $b_{1,1}\sim b_9$ 为材料常数。

采用如下公式来描述 c_p 和 K 与温度的关系：

$$c_p(T)=C_1+C_2(T-C_5)+C_3\tanh[C_4(T-C_5)]$$
对非结晶型材料　(3.6-42a)

$$c_p(T)=C_1+C_2(T-C_5)+C_3\exp[C_4(T-C_5)^2]$$
对结晶型材料　(3.6-42b)

$$K(T)=\lambda_1+\lambda_2(T-\lambda_5)+\lambda_3\tanh[\lambda_4(T-\lambda_5)] \tag{3.6-43}$$

式中，$C_1\sim C_5$、$\lambda_1\sim\lambda_5$ 为取决于材料的常数。

3.4 保压模拟数值计算过程

从以上公式可知，保压过程的分析与充模流动一样，归结为求解一组满足给定边界条件的偏微分方程。实践表明，采用有限元与有限差分的混合法是求解保压过程最有效的数值方法。因保压分析的计算方法与充模流动过程极为相似，故略去其推导过程及求解步骤。图 3.6-6 示出保压过程数值计算流程图。

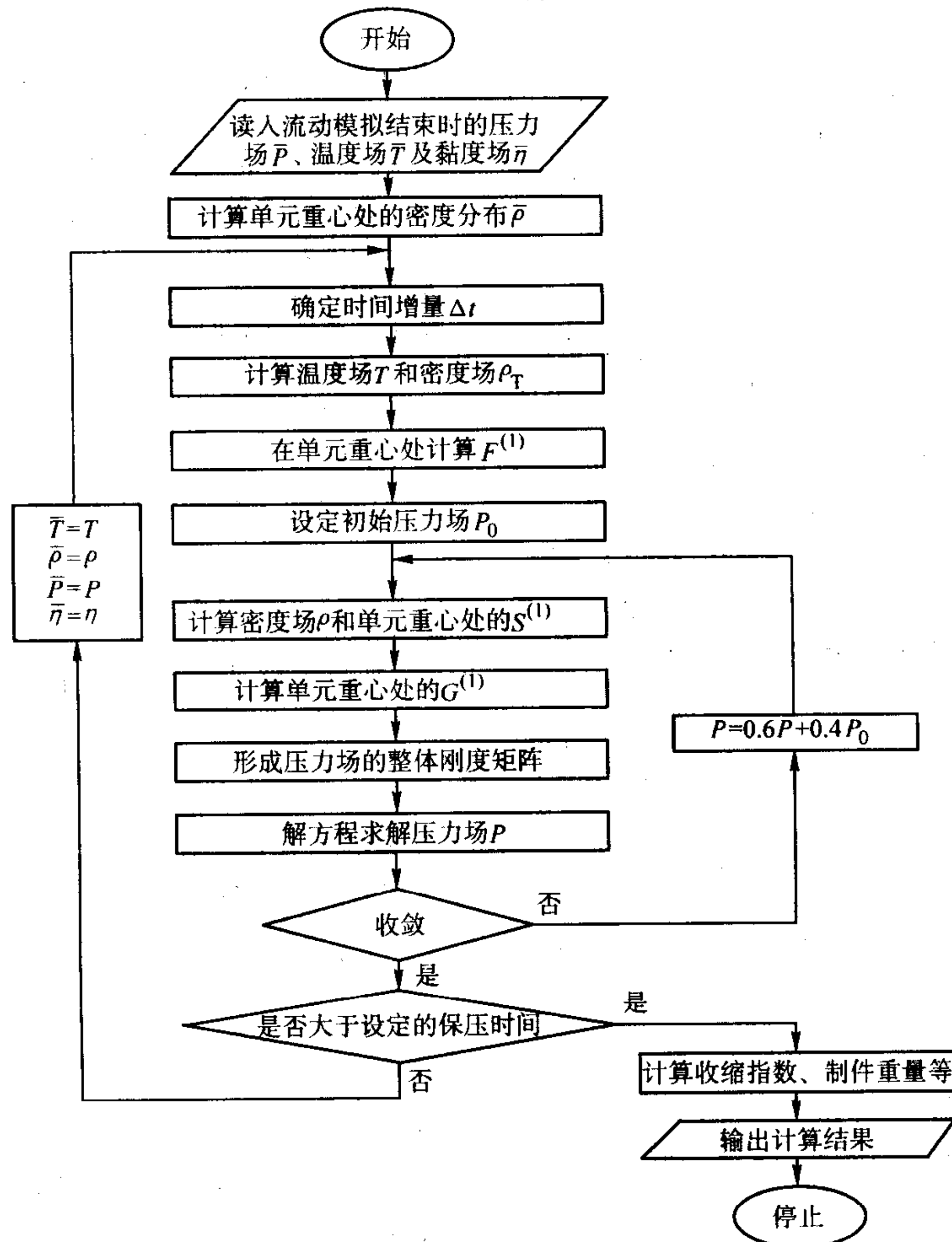


图 3.6-6　保压过程数值计算流程图

4 冷却过程模拟

实践表明，塑料注射模冷却时间约占整个注射循环周期的三分之二。塑料制品的翘曲变形与局部凹痕等弊病常常是冷却不良所致。因此，注射模冷却系统的设计直接影响到制品的生产效率和质量，模具设计师必须充分重视模具冷却系统的设计。

影响冷却系统的因素很多，诸如制品形状、冷却管道类型、尺寸与位置、冷却介质的流速与温度、制品与模具之间的非稳定热交换等，都对制品的冷却过程影响很大。

长期以来，注射模冷却系统的设计主要依赖于经验。模具设计师因缺乏经验而造成注射模具冷却效果不好的例子屡见不鲜。计算技术的飞速发展使得我们能借助于计算机，分析影响冷却系统的各个因素，模拟制品在模具内的冷却过程，从而提高模具冷却系统的设计质量。

4.1 一维冷却分析

对于具有简单扁平形状的大型注射模型腔，常采用如图 3.6-7 所示的等间距冷却管道布置方案。

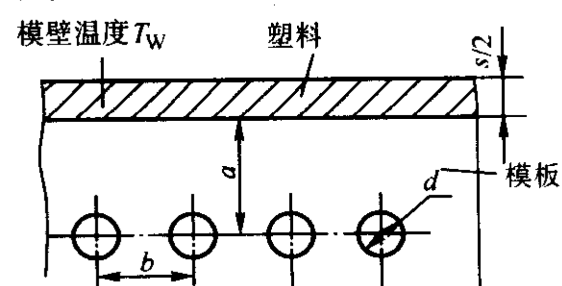


图 3.6-7　一维冷却系统的管道布置

因为模具的型腔壁较长，这样的冷却系统仅用图示的 a 和 b 两个参数就可确定，通常将这样的管道布置称之为一维冷却系统。

若设定冷却管道中允许的温度波动率为 J，可用如下公式迭代求解一维冷却系统中的管边距 a 和管间距 b：

$$J = 2.4\left(\frac{h_C d}{\lambda_W}\right)^{0.22}\left(\frac{b}{a}\right)^{2.8}\ln\left|\frac{b}{a}\right| \tag{3.6-44}$$

$$Q_C = \frac{0.87\ (T_W - T_C)\ \lambda_W h_C d}{0.876\lambda_W + h_C d\ [a - 0.13\ (\pi d - b)]} \tag{3.6-45}$$

式中，J 为冷却管道中允许的温度波动率，对于无定形塑料，$J<5\%$，对于结晶形塑料，$J<2.5\%$；d 为设定的冷却管道直径；h_C 为冷却水与管壁之间的传热系数；λ_W 为模具材料的传热系数；T_W 为型腔壁的平均温度；T_C 为冷却水的平均温度；Q_C 为冷却水应带走的热量。

由于直接利用上述两式求解 a 和 b 不方便，在编制程序时可按如下步骤计算 a 和 b。

1）设定允许的水管温度波动率 j、水管直径 d 以及 b/a 的初始值（如选定一个大值 $b/a=5.0$）。

2）将 b/a 的初始值代入式（3.6-44）中，计算 J 值。比较 J 值与 j 值的接近程度，若 J 值不接近于 j 值，则依次减少 b/a（如每次减小 0.1）。

3）通过制品与冷却水的热平衡分析获得冷却水应带走的热量 q_C。初次利用式（3.6-45）计算 Q_C 时，可假定 $a=d$（管道直径）。比较 Q_C 与 q_C 的大小，若两值有较大差距，逐次增加 a 值（如每次增加 1 mm），重新计算 Q_C 值，一直计算到 Q_C 与 q_C 接近为止，以此方法得到管边距 a 与管间距 b。

此算法简单，易于编程。对于大型、规则的塑料制品，如电视机外壳、箱类零件等，一维冷却分析的结果是令人满意的。

在冷却系统设计时，模具设计师除了水管的距离外，还希望能够获得冷却管道的长度、预测制品在脱模时的温度以及计算所需冷却时间等，这些均可采用近似或经验公式计算，具体算法可参阅有关文献。

4.2 二维冷却分析

由于大多数塑料制品均具有复杂的形状，以等间距水管布置为基础的一维冷却系统设计方案显然是无法胜任的。一种行之有效的方法是，对注射模典型截面上的冷却系统进行二维冷却分析。

注射模的冷却过程具有非稳态性。模具内某点的测量温度在循环注射过程中呈周期性变化。但这种非稳态温度波动较小，且主要表现在型腔表面区域。在二维冷却分析中，经常忽略温度的这种周期性变化，认为温度场是稳定的。

二维稳态传热在数学上归结为求解拉普拉斯方程：

$$\nabla^2 T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 \tag{3.6-46}$$

其边界条件为：

$$T = T_0 \qquad x \in \Gamma_1 \tag{3.6-47}$$

$$\frac{\partial T}{\partial n} = 0 \qquad x \in \Gamma_2 \tag{3.6-48}$$

$$-\lambda\frac{\partial T}{\partial n} = h(T - T_0) \qquad x \in \Gamma_3 \tag{3.6-49}$$

式中，T 为温度；λ 为模具材料的导热系数；Γ_1、Γ_2 和 Γ_3 为不同的边界条件。例如，模具与空气接触的部分属于 Γ_1，有时可简单地认为模具外壁温度是恒定的，且等于室温。边界条件 Γ_2 常用于对称的模具，此时因对称性可以只分析半个模具，模具的对称面便成为模具边界的一部分，因为在对称面上热流量为零，故温度的法向梯度也为0。边界条件 Γ_3 分别对应如下三种情况：

1）模具与空气的边界，此时不认为模具外壁温度是恒定的，式（3.6-49）中的 h 为模具与室温 T_0 之间空气的传热系数；

2）模具与冷却管道的边界，此时式（3.6-49）中的 h 为水管壁与冷却水交界面的传热系数，T_0 此时为冷却水的温度；

3）塑料制品与型腔壁的边界，此时 h 为塑料与模具之间的传热系数，可取制品与型腔壁之间在一个冷却周期内传热系数的平均值，T_0 为塑料制品冷却时的温度。

式（3.6-46）既可用有限元法求解，又可用边界元法求解。在求解注射成形冷却问题时，常用边界元法。因为在注射成形中关注的主要是型腔与制品表面的温度分布，同时边界元法仅需离散二维截面的边界而不是整个截面，能够简化操作，节省计算时间。

边界元积分公式借助于格林第二公式得到。格林第二公式可写为：

$$\int_D [T\nabla^2 K - K(\nabla^2 T)]\mathrm{d}t = \int_\Omega (T\frac{\partial K}{\partial n} - K\frac{\partial T}{\partial n})\mathrm{d}A \tag{3.6-50}$$

式中，T、K 为任意两个在区域 D 内二次可微函数；Ω 为区域 D 的边界。

如果能够选择函数 T 和 K 均满足拉普拉斯方程，即 $\nabla^2 K = \nabla^2 T = 0$，则有：

$$\int_\Omega (T\frac{\partial K}{\partial n} - K\frac{\partial T}{\partial n})\mathrm{d}A = 0 \tag{3.6-51}$$

T 选择满足拉普拉斯方程式（3.6-46）的温度场函数，K 选择满足拉普拉斯方程的“自由空间格林函数”，K 称为基本解。

在选定基本解后，通过积分计算，由上式可得到如下形式的公式：

$$C(P)T(P) = \int_\Gamma [K_n(P,Q)T(Q) - K(P,Q)T_n(Q)]\mathrm{d}S \tag{3.6-52}$$

式中，P、Q 为区域中和边界上的任意点，如图 3.6-8a 所示，$K(P,Q)$ 为基本解，二维分析时，$K(P,Q)=\ln r$（r 为 PQ 两点间的空间距离），$T_n = \frac{\partial T}{\partial n}$，$K_n = \frac{\partial K}{\partial n}$，$n$ 为点 Q 处的单位外法向向量，$C(P)$ 为 P 点内角，假定边界 S 平滑，当 P 在边界上时 $C(P)=\pi$，P 在区域内时 $C(P)=2\pi$，S 为边界 Γ 的弧长。

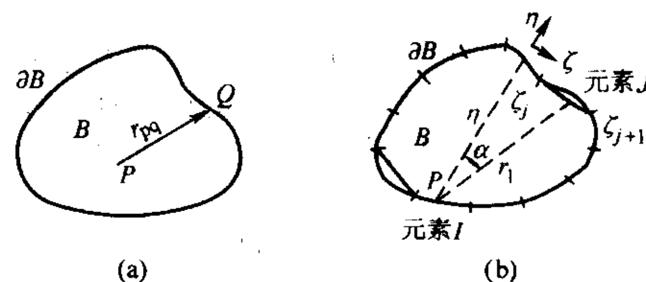


图 3.6-8 二维区域及其离散

式（3.6-52）又可改为：

$$C(P)T(P) = \int_\Gamma (T\frac{\partial \ln r}{\partial n} - \ln r\frac{\partial T}{\partial n})\mathrm{d}S \tag{3.6-53}$$

式（3.6-53）表明，区域内任一点温度 $T(P)$，皆能用

边界积分项定义，即已知边界上的 T、$\frac{\partial T}{\partial n}$，便可利用上式求得区域内任一点的温度。

如图3.6-8b所示，若将 P 点移至由 N 个线性单元组成的边界上，且假定 T 和 $\frac{\partial T}{\partial n}$ 在每个单元内为常数，则式（3.6-53）可改写为

$$\pi T_i = \sum_{j=1}^{N}\left[T_j\left(\frac{\partial \ln r_{ij}}{\partial n}\right)\mathrm{d}s - \left(\frac{\partial T}{\partial n}\right)_j \ln r_{ij}\mathrm{d}s\right] \tag{3.6-54}$$

上式的积分项可利用图3.6-8b中所示的局部坐标系（η，ζ）求出。这样，随着点 P 在边界上的移动，便可得到 N 个代数方程，加上 N 个边界条件，就能惟一地确定 $2N$ 个未知数（T 和 $\frac{\partial T}{\partial n}$）。

二维冷却分析软件可将分析结果（温度场或热流量场）用图形的形式显示在模具边界上，同时可以输出冷却时间、冷却水管进出口温度等数据。模具设计师利用二维冷却分析软件，可交互地改变冷却管道尺寸和位置，或者改变冷却介质的流速和温度，或者改变制品推出温度等，来选择合理的冷却系统设计方案，以便获得均匀的温度场分布和较短的模具冷却时间。

4.3 三维冷却分析

用二维典型截面代替三维模具进行冷却分析，具有编程简单、计算量小等许多优点，但很多模具的分析不能用二维典型截面来代替，特别是大型、精密的复杂注射模具，需要更精确的分析结果，此时应采用三维冷却分析。尤其是近年来计算机容量与计算速度的飞速发展，为三维冷却分析创造了条件，三维分析已成为注射成形冷却模拟的主流技术。

在三维冷却分析时，采用边界元法的优点除了能将三维问题化作二维问题处理外，还能与三维注射流动分析程序共用同一几何模型，大大简化几何造型的工作量。下面分别简述三维稳态和非稳态的冷却分析原理及实现方法。

4.3.1 三维稳态冷却分析

当考虑稳态传热时，三维边界积分方程与二维相同［式（3.6-50）］，差别仅是三维基本解不同，但与二维冷却分析相比，三维冷却分析的数值求解要困难得多，具体表现如下。

1）模具尺寸比例相差悬殊。例如，模具外表面比冷却管道内表面尺寸大若干数量级，冷却管道本身长度又比管道直径大许多倍，型腔厚度和长度的尺寸也相差甚远，除非将单元划分得很细，否则边界元法无法适应单元尺寸的这种大幅度变化。

2）离散管道圆周一般需要16个三角单元。由于模具内冷却管道很长，当划分管道表面网格时，就会产生大量单元。

3）边界元法产生非对称满秩矩阵。当模具具有复杂型腔时往往导致数据量与计算量过大，必须寻求合理的解决方案。

由于注射模具内的封闭型腔与金属零件内的封闭裂纹较为相似，一直以来比较流行的一种解决办法是采用断裂力学的研究手法，用型腔的中心层代替其封闭表面来克服数值求解的困难，这种简化方法可以使计算量成倍地减少。但是目前三维流动分析软件已普遍采用制品的三维实体模型，同时数据量的大小已不再是模拟软件需要考虑的主要约束条件，为了使冷却分析的几何模型及网格划分能与三维流动分析软件保持一致，新的数学模型不再用型腔的中心层代替封闭表面，而是直接采用制品实体模型的表面。

如果忽略温度和热流量沿冷却管道圆周方向上的变化，可以将冷却管道划分为特殊的二维单元（圆柱单元），这可以避免产生大量冷却管道表面三角单元。模具外表面通常用一个常温绝热大单元来处理。

在三维冷却分析中，同时考虑模具外表面、型腔表面和冷却管道，用类似于二维分析的推导，可以得到一组点 P 在模具外表面、点 P 在型腔面、点 P 在冷却管道上的联合方程式，然后用数值方法求解。

4.3.2 三维非稳态冷却分析

当对精密注射模作冷却系统分析时，需要考察型腔壁的各点温度随时间变化的情况，此时就需采用三维非稳态冷却分析，非稳态问题的求解比稳态问题要复杂得多。非稳态分析的传热方程为：

$$\frac{\lambda}{\rho c}\nabla^2 T = \frac{\partial T}{\partial t} \tag{3.6-55}$$

式中，t 为时间；ρ 为密度；c 为比热容。

对式（3.6-55）的计算机求解，国内外许多学者进行了大量研究，最早是采用有限差分法，用差分和差商代替微分和微商。为了提高求解的精度，20世纪60年代又流行有限元法。从理论上讲，将非稳态热传导微分方程求解转化为泛函变分问题，将温度对时间的微分用差分法展开，用数值方法求出各时间步长的温度场是可行的。然而，有限元法需对整个模具进行三维网格划分，会导致数据量过大、计算时间过长、操作复杂等弊病。

对于三维非稳态传热问题，国外学者在70年代曾提出用直接边界元法求解，采用与时间有关的基本解。但采用与时间有关的基本解进行边界元计算时仍要求三维网格划分，失去了边界元法中只计算边界积分和只划分边界网格的优越性。

美国SDRC公司的三维非稳态冷却分析软件，在求解式（3.6-55）时采用了傅里叶转换消去其时间变量，然后采用边界元法求解。对应于不同的傅里叶数的温度场，该方法或者由于精度问题，或者由于计算量大，在使用中仍受到一定限制。

华中科技大学模具技术国家重点实验室塑料模研究室，采用特解边界元法将非稳态传热问题转化为完全边界积分的形式，其处理过程如下。

引入满足式（3.6-55）的基本解：

$$K(P,Q) = \frac{\rho c}{4\pi\lambda r} \tag{3.6-56}$$

根据格林公式得到边界积分方程为：

$$\int_V \frac{\partial T}{\partial t}K(P,Q)\mathrm{d}V = \int_\Gamma [K_n(P,Q)T(Q) - K(P,Q)T_n(Q)]\mathrm{d}S - c(P)T(P) \tag{3.6-57}$$

式中 V 为边界所包围的体积。

等式左边包含体积分。为了将体积分化为面积分，将温度对时间的微分表示为一系列与空间有关的函数 f 和与时间有关的函数 a 的乘积：

$$\frac{\partial T}{\partial t} = \sum_{j=1}^{N} f_j(x)\frac{\partial a_j(t)}{\partial t} \tag{3.6-58}$$

假定 f 为某函数 ψ 的二阶导数 $f = \nabla^2\psi$，式（3.6-57）左端的体积分可表为：

$$\int_V \frac{\partial T}{\partial t}K\mathrm{d}V = \sum_{j=1}^{N}\frac{\partial a_j(t)}{\partial t}\int_V f_j K\mathrm{d}V = \sum_{j=1}^{N}\frac{\partial a_j(t)}{\partial t}\int_V \nabla^2\psi_j K\mathrm{d}V \tag{3.6-59}$$

将上式用格林公式变换，整理后得到

$$\int_V \nabla^2\psi_j K\mathrm{d}V = -\int_\Gamma [K(P,Q)\varphi_j - K_n(P,Q)\psi_j]\mathrm{d}S - C_j(P)\psi_j(P) \tag{3.6-60}$$

式中，$\varphi=\frac{\partial\psi}{\partial n}$。将上式代入式（3.6-57）可得：

$$c(P)T(P)+\int_\Gamma[K(P,Q)T_n(Q)-K_n(P,Q)T(Q)]\mathrm{d}S=$$
$$\frac{\rho c}{\lambda}\sum_{j=1}^{N}\left\{\int_\Gamma[K(P,Q)\varphi_j(Q)-K_n(P,Q)\psi_j]\mathrm{d}S+C_j(P)\psi_j(P)\right\}\tag{3.6-61}$$

至此，已将温度对时间微分有关的体积分转化为面积分，因此可用纯边界积分求解。

在冷却过程中，制品与模具不断地进行热交换。理想的非稳态分析应同时考虑四方面的热交换：制品内部、制品与模具之间、模具内部、模具与冷却介质之间。由于在制品和模具的冷却分析中分别采用了不同的数值分析方法，需采用迭代法对制品和模具的冷却进行耦合分析，其计算步骤如下。

1）选定时间步长。由于在制品及模具计算中采用的有限差分格式皆为无条件稳定，该时间步长的确定主要以考虑计算精度为主。

2）读入流动分析结果。以流动结束时刻的温度作为非稳态冷却分析的初始温度。

3）对制品进行一维非稳态传热的有限差分计算，确定本时间步长内制品与模具界面的热通量。

4）根据特解边界元法，对模具进行三维非稳态冷却分析，求出模具内的温度分布。

5）根据模具内的温度分布，重复步骤3和步骤4，直至制品与模具上的每一节点温度值均收敛。

6）将上一时间步长的模具温度分布作为本时间步长的初始条件，重复步骤3至步骤5，直至所有时间步长全部计算完毕。

5　应力分析

注射成形过程包括充模流动、保压及冷却三个阶段，注射成形是一个相当复杂的物理过程；一方面，非牛顿的高温塑料熔体在压力作用下，注入温度较低的型腔，产生流动应力；另一方面，在保压压力作用下，大约25%的塑料继续在高压下进入型腔以弥补因体积变化引起的收缩，同时由于熔体与模具型腔的温差导致注入的熔体不均匀冷却，引起热应力和保压应力。由于热应力和保压应力比因流动产生的应力大两个数量级以上，因此，注射成形应力分析主要考虑由温度和保压压力引起的应力。这些在成形过程中产生的应力，不仅导致了制品发生翘曲变形，而且影响制品的外观以及性能，因而，分析塑料制品在流动、保压及冷却过程中的内应力积累，具有重要的理论和实际意义。

5.1　基本假设

根据塑料制品的面方向尺寸远大于厚度方向尺寸的几何特点及注射成形过程中的变形约束情况，可以做出如下假设。

1）塑料制品为简单热流变材料。

2）在保压过程中，塑料制品紧贴在模壁上，由于模具的限制，塑料制品在面内很难发生变形，特别是形状复杂的塑料制品更是如此，这样可以将面内的应变假设为0，即：$\varepsilon_x=\varepsilon_y=0$。计算中将塑料制品进行三角形网格划分，在每个离散的三角形网格中，可以认为其纤维定向不变，为了研究的方便，可以在每个网格内建立如下的特殊坐标系：x坐标轴与塑料的纤维方向一致，在这种坐标系下x和y轴是应变主轴，ε_x和ε_y即为主应变。

3）塑料制品在厚度方向尺寸上没有约束，可以发生变形，应变场为$\varepsilon_z=\frac{\partial U_z}{\partial z}$。

4）由于塑料制品通常为薄壁件，同时保压过程中凝固层内应力产生的$x-z$和$y-z$面内的畸变很小，因而可假设$x-z$和$y-z$面内的畸变为0，即：$\varepsilon_{xz}=\varepsilon_{yz}=0$。

5）假设塑料制品两侧面在型腔内承受的压力为保压压力，压力值采用保压模拟中的结果：$\sigma_z(\pm b,t)=P(t)=P(t)$，式中$b$为塑料制品厚度的一半，$P$为保压压力，它与注射工艺条件和塑料黏度的材料参数有关。

6）在塑料成形过程中，假设塑料制品的各向异性主要表现在收缩率的差异上，黏弹性松弛函数与各向同性材料的松弛函数相同，基于这种假设，松弛函数可以写成如下形式：

$$G_{ijkl}=\frac{3[G_2(t)-G_1(t)]\delta_{ij}\delta_{kl}}{3}+\frac{[G_1(t)(\delta_{ik}\delta_{il}+\delta_{il}\delta_{ik})]}{2}\tag{3.6-62}$$

式中，δ为Kronecker符号，G_1为剪切松弛模量，G_2为体积松弛模量。

7）在保压阶段，由于压力的作用，做出了制品外侧完全限制的假定。在冷却过程中，由于浇口已完全凝固，则来自于注射机的压力自行拆除。此时，应该允许模内制品的自由收缩。横向应力为0的条件则替代了横向应变分量为0的条件。即

$$\int_h\sigma(z,t)\mathrm{d}z=0\tag{3.6-63}$$

5.2　保压过程中的内应力模型

根据上述对保压过程给出的假设，可以计算出其应变球张量和偏张量为

$$\begin{cases}\varepsilon_m=\varepsilon_z/3;\ e_x=e_y=-\varepsilon_m=-\varepsilon_z/3\\ e_z=\varepsilon_z-\varepsilon_m=2\varepsilon_z/3;\ e_{xz}=e_{yz}=0\end{cases}\tag{3.6-64}$$

利用热黏弹性力学原理，可以推出塑料在保压过程中的内应力计算式：

$$\sigma_{ij}(x,t)=\int_{-\infty}^{t}G_{ijkl}(\xi-\xi^n)\frac{\partial}{\partial t_n}[\varepsilon_{kl}(t_n)-\theta_{ij}(x,t_n)]\mathrm{d}t_n\tag{3.6-65}$$

式中，ξ为与温度有关的材料时间，θ为热应变和保压引起的体积应变，即：

$$\theta_{ij}(x,t')=\alpha_{ij}\Theta(x,t')+\varepsilon_p(x,t')\tag{3.6-66}$$

式中，α为热胀系数；Θ为热胀系数变换函数，其定义式为：

$$\Theta(x,t)=\frac{1}{\alpha_0}\int_{T_0}^{T}\alpha(t)\mathrm{d}T,\alpha_0=\alpha(T_0)\tag{3.6-67}$$

ε_p表示保压过程中体积变化引起的应变为：

$$\varepsilon_p(x,t')=-\frac{1}{3}\ln\frac{v[P_s(z),T_s(z)]}{v[0,T_s(z)]}\tag{3.6-68}$$

将式（3.6-64）代入式（3.6-65），整理后得出其三个主方向的内应力为：

$$\begin{cases}\sigma_x(x,t)=\int_{-\infty}^{t}G_1(\xi-\xi^n)\dfrac{\partial e_x}{\partial t_n}\mathrm{d}t_n+\\ \int_{-\infty}^{t}G_2(\xi-\xi^n)\dfrac{\partial}{\partial t_n}[\varepsilon_m(t_n)-\theta_x(x,t_n)]\mathrm{d}t_n\\ \sigma_y(x,t)=\int_{-\infty}^{t}G_1(\xi-\xi^n)\dfrac{\partial e_y}{\partial t_n}\mathrm{d}t_n+\\ \int_{-\infty}^{t}G_2(\xi-\xi^n)\dfrac{\partial}{\partial t_n}[\varepsilon_m(t_n)-\theta_y(x,t_n)]\mathrm{d}t_n\\ \sigma_z(x,t)=\int_{-\infty}^{t}G_1(\xi-\xi^n)\dfrac{\partial e_z}{\partial t_n}\mathrm{d}t_n+\\ \int_{-\infty}^{t}G_2(\xi-\xi^n)\dfrac{\partial}{\partial t_n}[\varepsilon_m(t_n)-\theta_y(x,t_n)]\mathrm{d}t_n\end{cases} \tag{3.6-69}$$

由以上假设可以看出，应变分量只与 z 轴坐标及时间 t 有关，因而内应力也只与 z 和 t 有关。

内应力计算式（3.6-69）中包含松弛函数、热胀系数、应变等变量，在松弛函数中包括材料时间 ξ，其他项中包括物理时间 t，为了计算此积分，需要把它表示成卷积形式。将 t 写成 ξ 的函数：

$$t=g(x,\xi),t'=g(x,\xi') \tag{3.6-70}$$

假设 F 为位置和时间的函数，引入以下变换形式：

$$\tilde{F}(x,\xi)=F(x,t) \tag{3.6-71}$$

将式（3.6-70）和式（3.6-71）应用于式（3.6-69）中的 σ、ε 和 θ，然后将式（3.6-69）进行拉氏变换，可以得到变换后的热应力计算式

$$\begin{cases}\tilde{\sigma}_x^*(z,\eta)=-\eta G_1^*(\eta)\tilde{\varepsilon}^*(z,\eta)/3+\\ \quad\eta G_2^*(\eta)[\tilde{\varepsilon}^*(z,\eta)/3-\tilde{\theta}_x^*(z,\eta)]\\ \tilde{\sigma}_y^*(z,\eta)=-\eta G_1^*(\eta)\tilde{\varepsilon}^*(z,\eta)/3+\\ \quad\eta G_2^*(\eta)[\tilde{\varepsilon}^*(z,\eta)/3-\tilde{\theta}_y^*(z,\eta)]\\ \tilde{\sigma}_z^*(z,\eta)=2\eta G_1^*(\eta)\tilde{\varepsilon}^*(Z,\eta)/3+\\ \quad\eta G_2^*(\eta)[\tilde{\varepsilon}^*(z,\eta)-\tilde{\theta}_y^*(z,\eta)]/3\end{cases} \tag{3.6-72}$$

式中，η 为 ξ 的变换参量，* 号表示经过 Laplace 变换后的函数形式。

由应变球张量及应变偏量的计算式（3.6-64）及应力应变关系可得：

$$S_{yz}=S_{xz}=0 \tag{3.6-73}$$

式中 S 为应力偏张量。将它代入关于 z 向的应力平衡方程式中，可以得出：

$$\frac{\partial\sigma_z(z,t)}{\partial z}=0 \tag{3.6-74}$$

利用应力边界条件假设，对式（3.6-74）求积分，可以得到 σ_z 为

$$\sigma_z(z,t)=\tilde{\sigma}_z(z,\xi)=-P(t)=-\tilde{P}(\xi) \tag{3.6-75}$$

将式（3.6-75）代入式（3.6-72），消去应变项 $\tilde{\varepsilon}^*$，可以得到计算公式：

$$\begin{cases}\tilde{\sigma}_x^*=\dfrac{[G_1^*(\eta)-G_2^*(\eta)]}{2G_1^*(\eta)+G_2^*(\eta)}\tilde{P}^*(\eta)+\\ \left\{\dfrac{G_2^*(\eta)-G_1^*(\eta)}{3[2G_1^*(\eta)+G_2^*(\eta)]}\tilde{\theta}_y^*(Z,\eta)-\tilde{\theta}_x^*(Z,\eta)\right\}\eta G_2^*(\eta)\\ \tilde{\sigma}_y^*=\dfrac{[G_2^*(\eta)-G_1^*(\eta)]}{2G_1^*(\eta)+G_2^*(\eta)}\tilde{P}^*(\eta)-\\ \dfrac{G_2^*(\eta)\tilde{\theta}_y^*(Z,\eta)\eta[7G_1^*(\eta)+2G_1^*(\eta)]}{3[2G_1^*(\eta)+G_2^*(\eta)]}\end{cases} \tag{3.6-76}$$

对于高聚物，可以将与球应力有关的松弛模量假设为弹性的，根据弹性力学中的定义，体积松弛模量可以取为：

$$G_2^*(\eta)=3K/\eta \tag{3.6-77}$$

式中，K 为体积模量。由弹性力学可知，与弹性模量的关系式为：$K=E/3(1-2\nu)$。由简单拉伸时的应力应变关系，可以推导出简单拉伸松弛模量与拉伸模量的关系式为：

$$G_1^*(\eta)=\frac{6KE^*(\eta)}{9K-\eta E^*(\eta)} \tag{3.6-78}$$

将式（3.6-77）、式（3.6-78）代入式（3.6-76），可以得到塑料在保压成形过程中的应力模型

$$\begin{cases}\tilde{\sigma}_x^*(z,\eta)=A_1\tilde{P}^*(\eta)-A_1K\tilde{\theta}_x^*(Z,\eta)-3K\tilde{\theta}_y^*(Z,\eta)\\ \tilde{\sigma}_y^*(z,\eta)=A_2\tilde{P}^*(\eta)-A_2\tilde{\theta}_y^*(Z,\eta)\end{cases} \tag{3.6-79}$$

式中 A_1、A_2 为计算参数：

$$\begin{cases}A_1=\dfrac{\eta E^*(\eta)-3K}{3K+\eta E^*(\eta)}\\ A_2=\dfrac{6K[3K+2\eta E^*(\eta)]}{3K+\eta E^*(\eta)}\end{cases} \tag{3.6-80}$$

由于式（3.6-79）是拉普拉斯变换结果，因此求解时需要进行拉普拉斯逆变换，然后采用数值方法求解。通过对保压时间划分为多个时间步，采用有限差分求解。

5.3 冷却过程中的内应力模型

根据上述对制品在冷却过程中提出的假设，有

$$\begin{cases}\sigma_{xy}=\sigma_{yz}=\sigma_{xz}=\sigma_{zx}=\sigma_z=0\\ \sigma_x(z,t)=\sigma_y(z,t)=\sigma_x(X,t)=\sigma_y(X,t)=\sigma(z,t)\\ \varepsilon_z(X,t)=\varepsilon_z(z,t)\\ \varepsilon_y(X,t)=\varepsilon_x(X,t)=\varepsilon(t)\end{cases} \tag{3.6-81}$$

同理可以计算出其应变球张量和偏张量为

$$\begin{cases}s_x=s_y=-\dfrac{s_z}{2}=\dfrac{\sigma(z,t)}{3}\\ e_x=e_y=-\dfrac{e_z}{2}=\dfrac{\varepsilon(t)-\varepsilon_z(z,t)}{3}\end{cases} \tag{3.6-82}$$

由一般的黏弹性材料本构方程可求得：

$$\sigma(z,\xi)=\int_{-\infty}^{\xi}G_1(\xi-\xi')\frac{\partial}{\partial\xi'}[\varepsilon(\xi')-\varepsilon_z(z,\xi')]\mathrm{d}\xi' \tag{3.6-83}$$

式（3.6-83）中 G_1 为切变模量，对于弹性部分，采用如下形式的材料本构关系：

$$\sigma_{kk}=3G_2[\varepsilon_{kk}-\alpha\Theta] \tag{3.6-84}$$

式（3.6-84）中 K 为体积模量，α 及 Θ 的定义同式（3.6-66）。

若 $\alpha(t)=\alpha_0$，由式（3.6-82）和式（3.6-84），则可得

$$2\sigma(z,\xi)=3G_2\{2\varepsilon(\xi)+\varepsilon_z(z,\xi)-3\alpha_0[T(z,\xi)-T(z,0)]\} \tag{3.6-85}$$

利用式（3.6-85），从式（3.6-83）中消去 $\varepsilon_z(z,\xi)$，可得：

$$\frac{4}{3G_2}\int_{-\infty}^{\xi} G_1(\xi-\xi')\frac{\partial}{\partial\xi'}[\sigma(z,\xi')]\mathrm{d}\xi' + \sigma(z,\xi) =$$
$$3\int_{-\infty}^{\xi} 2G_1(\xi-\xi')\frac{\partial}{\partial\xi'}[\varepsilon(\xi') - \alpha_0(T(z,\xi') - T(z,0))]\mathrm{d}\xi' \tag{3.6-86}$$

将式（3.6-86）进行拉氏变换，可以得到辅助模量$R(\xi)$，满足Volterra型方程

$$\frac{4}{3G_2}\int_{-\infty}^{\xi} G_1(\xi-\xi')\frac{\partial R(\xi')}{\partial\xi'}\mathrm{d}\xi' + R(\xi) = 2G_1(\xi) \tag{3.6-87}$$

所以，可以得到塑料在冷却成形过程中的应力模型

$$\sigma(z,\xi) = 3\int_{-\infty}^{\xi} R(\xi-\xi')\frac{\partial}{\partial\xi'}[\varepsilon(z,\xi') - \alpha_0 T(z,\xi')]\mathrm{d}\xi' \tag{3.6-88}$$

可采用梯形法对式（3.6-88）联合应力平衡条件进行数值求解，由此可得到制品在冷却过程中的面内应力。式（3.6-88）中的辅助模量同样可采用梯形法求解式（3.6-87）得到。由于材料时间初始变化很快，易造成数值上的不稳定。因此，在求解式（3.6-88）时可采用一种准数值化求解过程，先由式（3.6-88）联合应力平衡条件式（3.6-63）推导出如下方程求取沿制品壁厚上不同瞬间的应变：

$$\varepsilon_n = \frac{1}{\sum\limits_n S_{n,n-1}\Delta z}\{\sum_n \Delta z S_{n,n-1}[\varepsilon_{n-1} + \alpha_0(T_n - T_{n-1})] - \sum_n \Delta z \sum_{i=1}^{n-1} S_{i,i-1}[\varepsilon_i - \varepsilon_{i-1} - \alpha_0(T_i - T_{i-1})]\} \tag{3.6-89}$$

其中
$$S_{i,i-1} = \frac{\int_{\xi_{i-1}}^{\xi_i} R(\xi_n - \xi')\mathrm{d}\xi'}{\xi_i - \xi_{i-1}} \tag{3.6-90}$$

一旦求出制品不同瞬间的面内应变，由式（3.6-88）就能求出沿制品壁厚上相应的应力分布；初始应变取保压结束时的体积应变。

6　翘曲变形计算

6.1　翘曲变形计算的基本原理

注塑成形过程中，聚合物的力学性能和热物性发生了很大改变，从而导致了多种残余应力的形成。又由于出模后制品的温度变化不大，可忽略塑料材料的非线性问题，而认为注塑件是线性弹性体。因此，制品的应力-应变本构方程为

$$\{\sigma\} = [D](\{\varepsilon\} + \{\varepsilon^{T}\}) + \{\sigma_1\} + \{\sigma_t\} + \{\sigma_o\} \tag{3.6-91}$$

式中，$\{\sigma\}$ 为应力向量；$[D]$ 为弹性矩阵；$\{\varepsilon\}$ 为应变向量；$\{\sigma_1\}$ 为流动残余应力；$\{\sigma_t\}$ 为模具内不均匀冷却所产生的残余应力；$\{\sigma_o\}$ 为由于模具排气或制品被顶出时可能产生的应力；$\{\varepsilon^T\}$ 为制品出模后的继续冷却至室温时所产生的热应变，这主要是由于出模时制品内部的温度分布不均匀而引起的。

注塑制品出模后的应变与位移几何关系满足

$$\{\varepsilon\} = [B][d] \tag{3.6-92}$$

式中，$[B]$ 为描述弹性体应变与位移几何关系的几何矩阵；$[d]$ 为位移。

根据上述应力-应变本构关系和应变-位移几何关系，再利用虚功原理可推导以下平衡方程

$$[F] = \int_V [B]^{T}[D][B]d_v[d] + [F_T] + [F_1] + [F_t] + [F_0] \tag{3.6-93}$$

式中，$[F]$ 为外载荷；$[B]^T$ 为 $[B]$ 的转置矩阵；$[F_T]$ 为制品出模后继续冷却时所产生热应力的等效载荷；$[F_1]$ 为流动残余应力所形成的等效载荷；$[F_t]$ 为模具内不均匀冷却所产生温度应力的等效载荷；$[F_0]$ 为出模时可能产生应力的等效载荷。$[F_T]$，$[F_1]$，$[F_t]$，$[F_0]$ 均可由其应力计算出，即

$$[F_T] = \int_v [B]^T[D][\varepsilon^T]\mathrm{d}v$$
$$[F_1] = \int_v [B]^T[\sigma_1]\mathrm{d}v$$
$$[F_t] = \int_v [B]^T[\sigma_t]\mathrm{d}v$$
$$[F_0] = \int_v [B]^T[\sigma_0]\mathrm{d}v$$

由于制品出模后没有受到外力作用，根据假设可忽略流动残余应力和模具排气或被顶出时所产生应力对制品翘曲变形的影响，即流动残余应力等效载荷 $[F_1]$ 和出模时可能产生的等效载荷 $[F_0]$ 不予以考虑，且令

$$[k]^e = \int_v [B]^T[D][B]\mathrm{d}v \tag{3.6-94}$$

式中，$[k]^e$ 为单元刚度矩阵，则方程式（3.6-93）可简化为

$$[F_e] = -([F_T] + [F_t]) = [k]^e[d] \tag{3.6-95}$$

式中，$[F_e]$ 为单元节点等效载荷。

方程式（3.6-94）揭示了单元刚度矩阵的计算方法，方程式（3.6-95）则是离散单元的平衡方程，也即它们反映了单个单元的单元特性。对整个制品而言，应分析其整体结构的特性。假设制品离散网格的总数为 N，将其所有单元的平衡方程进行集合，即可形成系统方程组

$$[F] = \sum_i^N [F_e]_i = \sum_i^N [k]_i^e[d]_i = [K]\{d\} \tag{3.6-96}$$

式中，$[F]$ 为整体等效节点载荷；$[K]$ 为整体刚度矩阵；$\{d\}$ 为节点位移向量；求解系统方程即可计算翘曲变形。

当制品变形较小时，采用小变形线性有限元理论模拟翘曲变形，一般不会引起很大误差。但是，当制品变形较大时，若仍按小变形线性问题来处理，则会带来较大的误差，此时应该考虑翘曲变形中的几何非线性问题。

根据几何非线性有限元理论和虚功原理，可推导出非线性方程组

$$[K']\{X\} = [F'] \tag{3.6-97}$$

式中，$[K']$ 为几何非线性翘曲变形的整体刚度矩阵，它与节点位移有关，$[F']$ 为几何非线性翘曲变形的整体节点载荷，$\{X\}$ 为大变形节点位移，求解非线性方程即可获得大变形位移。

6.2　线性翘曲变形有限元分析的数学模型

为进行注塑制品翘曲变形有限元分析，可作如下假设。

1）注塑制品可看作理想薄壳结构。这是因为，注塑制品是薄壁件，其平面内两个方向尺寸往往比厚度方向尺寸大一个数量级以上。

2）假设采用三角板单元组成的折板系统代替实际的制品结构，则可采用基于三角板单元的小挠度薄壳理论模拟翘曲变形。采用三角单元的另一好处是能保证注塑成形CAD/CAE分析软件几何模型的一致性，和流动/保压/冷却/应力分析结果的可用性。

3）注塑制品的翘曲变形可以看作是面内变形和弯曲变形的组合，这两种变形状态互相独立，互不耦合。

4）离散制品所采用的三角板单元由常应变三角单元和克希霍夫三角单元组合而成，如图3.6-9所示。这样，面内载荷仅由常应变三角单元的面内刚度承受，只引起面内变形，侧向载荷由克希霍夫三角单元的弯曲刚度承受，只引起弯曲变形。因此，根据平面应力问题和薄板弯曲问题求解面

内变形和弯曲变形后，进行线性叠加可计算注塑制品的翘曲变形。

如图3.6-9所示，以三角单元的1，2顶点连线作为局部坐标系的y轴，垂直于三角单元面的直线为z轴，在整体坐标系（x'，y'，z'）中建立三角单元的局部坐标系（x，y，z）。由于翘曲变形可看作是由面内变形和弯曲变形叠加而成，故制品上任一点在局部坐标系内沿坐标轴方向的位移为用平面内位移和弯曲位移表示为：

$$[u \quad v \quad w]^{\mathrm T}=[u^{\mathrm p}\ v^{\mathrm p}\ 0^{\mathrm p}]^{\mathrm T}+[u^{\mathrm b}\ v^{\mathrm b}\ w]^{\mathrm T} \qquad (3.6\text{-}98)$$

式中，$u^{\mathrm p}$，$v^{\mathrm p}$分别为平板中面内x，y方向位移；w为平板弯曲中面挠度；$u^{\mathrm b}$，$v^{\mathrm b}$是由弯曲变形而产生的面内位移。

对上式进行微分，并根据应变定义可以推导出

$$\{\varepsilon\}=\{\varepsilon\}^{\mathrm p}+\{\varepsilon\}^{\mathrm b} \qquad (3.6\text{-}99)$$

式中，$\{\varepsilon\}$为制品内任一点的应变；$\{\varepsilon\}^{\mathrm p}$为平面应变；$\{\varepsilon\}^{\mathrm b}$为弯曲应变。该方程说明了翘曲变形中任意点的应变是由平面应变和弯曲应变组合而成的。

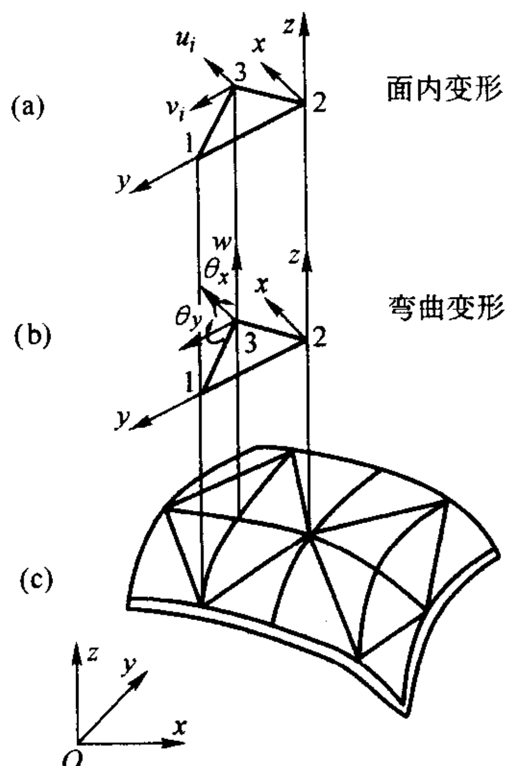


图3.6-9　三角板单元在局部坐标系内变形的分解

设u_i，v_i，w_i分别表示节点i在x，y，z方向的位移，θ_{ix}，θ_{iy}，θ_{iz}分别表示节点i绕x，y，z方向的转角，则局部坐标系内节点i的位移$\{\delta_i\}$可表示如下

$$\{\delta_i\}=[u_i\ v_i\ w_i\ \theta_{ix}\ \theta_{iy}\ \theta_{iz}]^{\mathrm T}=[\{\delta_i^{\mathrm p}\}\{\delta_i^{\mathrm b}\}]^{\mathrm T},\ i=1,2,3 \qquad (3.6\text{-}100\mathrm a)$$

$$\{\delta_i^{\mathrm p}\}^{\mathrm T}=[u_i\ v_i]^{\mathrm T},\ \{\delta_i^{\mathrm b}\}^{\mathrm T}=[w_i\ \theta_{ix}\ \theta_{iy}]^{\mathrm T},\ i=1,2,3 \qquad (3.6\text{-}100\mathrm b)$$

式中，$\{\delta_i^{\mathrm p}\}$为平面内位移；$\{\delta_i^{\mathrm b}\}$为板弯曲位移。对于平板弯曲，三角板单元实际上只有五个自由度，θ_{iz}总是等于零，引入的目的是为了进行整体坐标系与局部坐标系之间的转换。因此，单元节点位移$\{\delta\}_{\mathrm e}$可表示为

$$\{\delta\}_{\mathrm e}=[\{\delta_i\}\ \{\delta_j\}\ \{\delta_m\}]^{\mathrm T} \qquad (3.6\text{-}101)$$

对于离散网格单元，选定了位移函数后，可推出应变与位移之间的几何关系

$$\{\varepsilon\}=[B]\{\delta\}_{\mathrm e}=([B^{\mathrm p}]+[B^{\mathrm b}])\{\delta\}_{\mathrm e} \qquad (3.6\text{-}102)$$

式中，$[B]$为几何矩阵，反映了翘曲变形应变与位移之间的几何关系；$[B^{\mathrm p}]$为平面应力问题的几何矩阵；$[B^{\mathrm b}]$为薄板弯曲变形的几何矩阵。利用方程式（3.6-99）~式（3.6-102），得到翘曲变形的单元刚度矩阵

$$[k]^{\mathrm e}=\int([B^{\mathrm p}]^{\mathrm T}+[B^{\mathrm b}]^{\mathrm T})[D]([B^{\mathrm p}+[B^{\mathrm b}]])\mathrm dv \qquad (3.6\text{-}103)$$

式中，$[k]^{\mathrm e}$为单元刚度矩阵。由于面内变形和弯曲变形互不耦合，所以单元刚度矩阵$[k]^{\mathrm e}$的6×6阶子阵$[k_{rs}]$也可由平面变形和弯曲变形的刚度矩阵表示，即

$$[k_{rs}]_{6\times6}=\begin{bmatrix}k_{rs}^{\mathrm p} & 0 & 0\\ 0 & k_{rs}^{\mathrm b} & 0\\ 0 & 0 & 0\end{bmatrix} \qquad (3.6\text{-}104)$$

式中，$[k_{rs}^{\mathrm p}]$为平面应力问题单元刚度矩阵的子阵；$[k_{rs}^{\mathrm b}]$为薄板弯曲变形单元刚度矩阵的子阵。从方程式（3.6-104）可看出，只要计算出$[k_{rs}^{\mathrm p}]$和$[k_{rs}^{\mathrm b}]$，即可计算$[k]^{\mathrm e}$。

对于平面应力问题的单元刚度矩阵，其求解过程比较简单。选择三角单元节点坐标的线性函数为单元位移函数，即

$$u^{\mathrm p}=\alpha_1+\alpha_2x+\alpha_3y \qquad (3.6\text{-}105\mathrm a)$$

$$v^{\mathrm p}=\alpha_4+\alpha_5x+\alpha_6y \qquad (3.6\text{-}105\mathrm b)$$

式中，α_1，…，α_6是待定系数；$u^{\mathrm p}$，$v^{\mathrm p}$是单元内任意一点在x和y方向的位移分量。将三角单元三节点的位移代入上面两式，可求出所有待定系数。再根据单元刚度矩阵的求解思路，可推导出平面应力问题单元刚度矩阵及其子阵的计算公式

$$[k^{\mathrm p}]^{\mathrm e}=\iint[B^{\mathrm p}]^{\mathrm T}[D^{\mathrm p}][B^{\mathrm p}]t\mathrm dx\mathrm dy \qquad (3.6\text{-}106\mathrm a)$$

$$[k_{ij}^{\mathrm p}]=t\Delta[B_i^{\mathrm p}]^{\mathrm T}[D^{\mathrm p}][B_j^{\mathrm p}](i,j=1,2,3) \qquad (3.6\text{-}106\mathrm b)$$

式中，$[k^{\mathrm p}]^{\mathrm e}$为平面应力问题的单元刚度矩阵；$[k_{ij}^{\mathrm p}]$为$[k^{\mathrm p}]^{\mathrm e}$的$2\times2$阶子阵；$t$为制品厚度；$\Delta$为三角单元面积；$[B_i^{\mathrm p}]$为平面应力问题几何矩阵的$3\times2$阶子阵。

至于薄板弯曲问题的单元刚度矩阵，其计算过程则比较复杂，比平面应力问题单元刚度矩阵的计算困难得多，下面将阐述其求解思路和方法。

6.2.1　平板弯曲变形单元刚度矩阵的计算

薄板弯曲时，其应力应变本构关系仍是线弹性的，但几何关系却不同，由基于克希霍夫假设的小挠度薄壳理论可知平面内位移$u^{\mathrm b}$，$v^{\mathrm b}$与挠度w的关系满足

$$u^{\mathrm b}=-z\frac{\partial w}{\partial x},\quad v^{\mathrm b}=-z\frac{\partial w}{\partial y} \qquad (3.6\text{-}107)$$

式中，z为所考察点到中面的距离。

对上式进行微分，可推出弯曲变形的应变与位移的几何关系

$$\{\varepsilon\}^{\mathrm b}=\begin{Bmatrix}\varepsilon_x^{\mathrm b}\\ \varepsilon_y^{\mathrm b}\\ \gamma_{xy}^{\mathrm b}\end{Bmatrix}=\begin{Bmatrix}\partial u^{\mathrm b}/\partial x\\ \partial v^{\mathrm b}/\partial y\\ \partial u^{\mathrm b}/\partial y+\partial v^{\mathrm b}/\partial x\end{Bmatrix}=\begin{Bmatrix}-z\partial^2w/\partial x^2\\ -z\partial^2w/\partial y^2\\ -2z\partial^2w/\partial x\partial y\end{Bmatrix} \qquad (3.6\text{-}108)$$

薄板弯曲变形有限元分析，采用的是非完全协调三角薄板单元，为保证计算的收敛性，必须对位移函数有一定的要求，合理地选择位移函数是有限元法求解薄板弯曲问题的关键。位移函数必须能反映单元刚性位移，能反映单元常量形变，且应尽可能反映位移的连续性。这里采用面积坐标系下的位移函数，即

$$w=\beta_1L_1+\beta_2L_2+\beta_3L_3+\beta_4L_2L_3+\beta_5L_3L_1+\beta_6L_1L_2+\beta_7(L_2L_3^2-L_3L_2^2)+\beta_8(L_3L_1^2-L_1L_3^2)+\beta_9(L_1L_2^2-L_2L_1^2) \qquad (3.6\text{-}109)$$

式中，L_1，L_2，L_3分别为单元内任意点的三个面积坐标值；β_1，…，β_9为待定系数。将三角单元三节点位移代入上式，即可求出所有待定系数。

根据面积坐标和直角坐标的性质，可推出面积坐标和直角坐标的偏导数关系

$$\begin{Bmatrix}\partial^2/\partial x^2\\ \partial^2/\partial y^2\\ 2\partial^2/\partial x\partial y\end{Bmatrix}=[E]\begin{Bmatrix}\partial^2/\partial L_1^2\\ \partial^2/\partial L_2^2\\ 2\partial^2/\partial L_1\partial L_2\end{Bmatrix} \qquad (3.6\text{-}110)$$

式中，$[E]$ 为系数矩阵。

利用方程式（3.6-110），可确定位移函数和形函数如下式

$$w=[N^b]\{\delta^b\}^e=[N_1^b\ N_2^b\ N_3^b]\{\delta^b\}^e \quad (3.6\text{-}111)$$

式中，$[N^b]$ 为薄板弯曲的形函数；$[N_i^b]$ 为形函数的子阵；$\{\delta^b\}^e$ 为单元位移。

利用方程式（3.6-108）和式（3.6-111），可推出薄板弯曲的几何方程

$$\{\varepsilon\}^b=[B^b]\{\delta^b\}^e \quad (3.6\text{-}112)$$

式中

$$[B^b]=\begin{bmatrix}\partial^2N_1^b/\partial x^2 & \partial^2N_2^b/\partial x^2 & \partial^2N_3^b/\partial x^2\\ \partial^2N_1^b/\partial y^2 & \partial^2N_2^b/\partial y^2 & \partial^2N_3^b/\partial y^2\\ 2\partial^2N_1^b/\partial x\partial y & 2\partial^2N_2^b/\partial x\partial y & 2\partial^2N_3^b/\partial x\partial y\end{bmatrix}=[B_1^b\ B_2^b\ B_3^b] \quad (3.6\text{-}113)$$

式中，$[B^b]$ 为薄板弯曲变形的几何矩阵；$[B_i^b]$ 是 $[B^b]$ 的 3×3 阶子阵。

由于方程式（3.6-111）中的形函数 $[N^b]$ 是用面积坐标表示的，而几何矩阵中的元素都是对直角坐标的偏导数，因此，应将几何矩阵 $[B^b]$ 写成对面积坐标的偏导数形式。利用方程式（3.6-110），可得几何矩阵的子阵

$$[B_i^b]=\frac{1}{4\Delta}[E][\Phi][G_i](i=1,2,3) \quad (3.6\text{-}114)$$

式中，$[\Phi]$ 为 3×9 阶面积坐标矩阵；$[G_i]$ 为 9×3 阶系数矩阵，且

$$[G_i]=[G_{i1}\ G_{i2}\ G_{i3}]^T \quad (3.6\text{-}115)$$

式中，$[G_{ij}]$ 为 $[G_i]$ 的子阵，且 $i=1,2,3$，$j=1,2,3$。

利用方程式（3.6-104）和式（3.6-114），可推出薄板弯曲变形单元刚度矩阵的子阵

$$[k_{ij}^b]=\frac{D}{192\Delta^3}\{[G_{i1}]^T[C]\sum_{m=1}^{3}H_{1m}[G_{jm}]+[G_{i2}]^T[C]\sum_{m=1}^{3}H_{2m}[G_{jm}]+[G_{i3}]^T[C]\sum_{m=1}^{3}H_{3m}[G_{jm}]\} \quad (3.6\text{-}116)$$

式中，$D=\frac{Et^3}{12(1-\mu^2)}$；$E$ 为弹性模量；μ 为泊松比；$[C]$ 为 3×3 阶常数矩阵；$[H_{ij}]$ 为 3×3 阶系数矩阵。至此，即导出了薄板弯曲变形单元刚度矩阵的计算模型。

6.2.2　等效节点载荷与整体载荷列阵的计算

对于注塑制品，可将其看作多层薄片在厚度方向上累加而成，即将制品沿厚度方向分为多层。为简单起见，假设每层材料的力学性能参数相同，且每层各单元内温度应力相同，但是，制品同一单元不同层之间的温度应力分布并不相同，正是这种不均匀的分布状况，才导致注塑制品翘曲变形。

如前所述，注塑制品在流动过程中产生流动应力，在冷却过程中形成温度应力。但考虑到流动应力的计算比较困难，且对翘曲变形影响相对较小，因而，注塑应力分析模块只计算模具内形成的温度应力，而忽略流动应力。因此，如何将这些温度应力转化为等效节点载荷，是翘曲变形有限元数值模拟中必须解决的问题之一。

设 $[\sigma_t]$ 为注塑制品在模具内不均匀冷却而形成的温度应力，对于温度应力等效节点载荷 $[F_t]^e$，由于制品在厚度方向上应力分布不同，故不可直接套用经典公式 $[F_t]^e=\int_t\int_s[B]^T[\sigma_t]dsdt$，而应对此式进行改造。由于制品内温度应力沿厚度方向分层分布，故可沿厚度方向对温度应力进行积分，从而求解等效节点载荷，即

$$[F_t]^e=\sum_{i=1}^{M}\int_{Z_{i-1}}^{Z_i}(\int_s[B]^T([\sigma_t])_i ds)dt \quad (3.6\text{-}117)$$

式中，M 为注塑制品沿厚度方向所分的层数；Z_i 为第 i 层距中性面的距离；$([\sigma_t])_i$ 为单元在第 i 层的应力分布，该部分数据来自于应力分析模块。由于三角单元的等效节点载荷可由其节点的等效节点载荷表示，即

$$[F_t]^e=[F_{t1}\ F_{t2}\ F_{t3}]^T \quad (3.6\text{-}118)$$

式中，$[F_{ti}]$ 为节点 i 的等效节点载荷，且可分解为平面应力问题的等效节点载荷和薄板弯曲变形的等效节点载荷，即

$$[F_{ti}]=[F_{ti}^pF_{ti}^b0]^T\quad(i=1,2,3) \quad (3.6\text{-}119)$$

式中，$[F_{ti}^p]$ 为节点 i 的平面应力问题的等效节点载荷；$[F_{ti}^b]$ 为节点 i 的薄板弯曲变形的等效节点载荷。

分别将平面应力的几何矩阵和板弯曲的几何矩阵代人方程式（3.6-119），即可求得上式中的 $[F_{ti}^p]$ 和 $[F_{ti}^b]$ 为

$$[F_{ti}^p]=\sum_{j=1}^{M}\int_{Z_{j-1}}^{Z_j}(\int_s[B_i^p]^T([\sigma_t])_j ds)dt(i=1,2,3) \quad (3.6\text{-}120a)$$

$$[F_{ti}^b]=\sum_{j=1}^{M}\int_{Z_{j-1}}^{Z_j}(\int_s[B_i^b]^T([\sigma_t])_j ds)dt(i=1,2,3) \quad (3.6\text{-}120b)$$

注塑成形过程中，塑料制品冷却到一定时间即被顶出模具。制品出模时刻的温度要比环境温度高，而且，制品内部的温度分布不均匀。在这种情况下，制品出模后的继续冷却过程相当于一个非均匀的自然冷却过程，所以会产生一定的热应力，从而对制品的翘曲变形有一定影响。

由于制品出模时与环境温度的温差分布不均匀，从而产生了热应变，即

$$([\varepsilon^T])_i=\begin{pmatrix}\alpha_x\\ \alpha_y\\ \alpha_z\end{pmatrix}(\Delta T)_i \quad (3.6\text{-}121)$$

式中，$([\varepsilon^T])_i$ 为单元在第 i 层的热应变；$(\Delta T)_i$ 为单元在第 i 层的温度差，该部分数据可从注塑冷却分析结果中获得；α_x，α_y，α_z 为塑料在 x，y，z 方向的热胀系数。

根据应力－应变本构关系，将制品分为多层，利用方程式（3.6-121）、式（3.6-117），可得

$$[F_T]^e=\sum_{i=1}^{M}\int_{Z_{i-1}}^{Z_i}(\int_s[B]^T[D]([\varepsilon^T])_i ds)dt \quad (3.6\text{-}122)$$

式中，$[F_T]^e$ 为出模后不均匀冷却所产生热应力的等效节点载荷，根据前面所述模具内温度应力等效节点载荷的求解思路，将出模后不均匀冷却所产生热应力的等效节点载荷分解为平面应力问题的等效节点载荷和薄板弯曲变形的等效节点载荷，则很容易计算出模后形成热应力的等效节点载荷。

至于整体节点载荷，在单元等效节点载荷已知的情况下是很容易计算的。将所有节点的等效节点载荷在公共节点上叠加后，即可计算出整体节点载荷

$$[F_t]_N=\sum_{i=1}^{N}[F_t]_i^e=[F_{t1},F_{t2},\cdots,F_{tn}] \quad (3.6\text{-}123a)$$

$$[F_T]_N=\sum_{j=1}^{N}[F_T]_j^e=[F_{T1},F_{T2},\cdots,F_{Tn}] \quad (3.6\text{-}123b)$$

式中，N 为制品所离散的单元个数；n 为网格单元的节点个数；$[F_t]_N$、$[F_T]_N$ 分别为模具内不均匀冷却所产生温度应力的整体等效节点载荷和出模后继续不均匀冷却所引起的热

应力的整体等效节点载荷。

6.3 几何非线性翘曲变形有限元分析的数学模型

当采用线性翘曲变形分析软件所计算的变形比较大时，说明此时的变形是大变形，线性分析结果的误差比较大，应该利用线性分析结果，进一步进行几何非线性分析。几何非线性有限元理论和线性有限元理论的分析思路是一致的，都遵循一定的研究步骤，即先将产品结构离散化，选择单元位移插值函数，根据非线性几何方程，推导应变与位移之间关系，分析单元特性（单元刚度矩阵和单元载荷信息），组集整体结构特性，形成、求解非线性方程组。两者的最大差别主要在于前者的几何方程（应变与位移的关系）是非线性的。

几何非线性翘曲变形有限元分析的数学模型推导是从虚功方程开始，对于整个制品来说，其虚功方程是每个单元的虚功方程之和，即

$$F(\{\delta\}) = \sum\int_e [\overline{B}]^T\{\sigma\}_e dv - \{R\} = 0 \qquad (3.6\text{-}124)$$

式中，$[\overline{B}]$ 为大变形的几何矩阵；$[\overline{B}]^T$ 为 $[\overline{B}]$ 的转置矩阵；$\{\sigma\}_e$ 为应力向量；$\{R\}$ 是由每个单元的节点载荷列阵集成的总载荷列阵；$\{\delta\}$ 为大变形位移；$F(\{\delta\})$为制品变形时的总功。

对于几何非线性问题，由于大变形几何方程是线性部分（与小变形理论相同）与二次项部分之和，故 $[\overline{B}]$ 可写为

$$[\overline{B}] = [B_0] + [B_L] \qquad (3.6\text{-}125)$$

式中，$[B_0]$ 为几何方程的线性部分，是一个常数矩阵；$[B_L]$ 为几何方程的非线性部分，且为位移对坐标导数的二次项函数。对上式求导，可得

$$d[\overline{B}] = d[B_L] \qquad (3.6\text{-}126)$$

设载荷 $\{R\}$ 与位移 $\{\delta\}$ 无关，对方程式（3.6-124）求导，并利用方程式（3.6-125）和式（3.6-126），可推出

$$dF = \sum[K_T]_e d\{\delta\}_e = [K_T]d\{\delta\} \qquad (3.6\text{-}127)$$

$$[K_T]_e = [K_E]_e + [K_L]_e + [K_G]_e \qquad (3.6\text{-}128)$$

式中，$[K_T]_e$ 是单元的切向刚度矩阵；$[K_T]$ 是由 $[K_T]_e$ 集成的总切向刚度矩阵；$[K_E]_e$ 是作为小位移问题处理时的单元刚度矩阵；$[K_L]_e$ 与几何方程中的非线性项相关，称为“大位移矩阵”，$[K_G]_e$ 为几何刚度矩阵。这几种刚度矩阵可按下列公式求得

$$[K_E]_e = \int_e [B_0]^T[D][B_0]dV \qquad (3.6\text{-}129)$$

$$[K_L]_e = \int_e([B_0]^T[D][B_L] + [B_L]^T[D][B_0] + [B_L]^T[D][B_L])dV \qquad (3.6\text{-}130)$$

$$\int_e d[\overline{B}]^T\{\sigma\}dv = \int_e d[B_L]^T\{\sigma\}dv = [K_G]_e d\{\delta\}_e \qquad (3.6\text{-}131)$$

由方程式（3.6-127）~式（3.6-131）可求出几何非线性翘曲变形的位移增量。如果已知初始位移，那么几何非线性变形的最终位移便可确定。从方程式（3.6-129）~式（3.6-131）可知，要求解几何非线性翘曲变形，关键问题是如何确定应变与大位移之间的几何关系 $[B_L]$。

对于几何非线性翘曲变形问题，不能沿用小挠度薄板理论描述其变形特点，而应采用大挠度薄板理论描述应变与位移关系的几何方程。其离散单元仍可采用前述的常应变三角单元和克希霍夫三角单元组合而成的三角板单元，几何非线性变形问题仍可看作是平面变形和薄板弯曲变形的组合。对于大挠度薄板，其中面应变由两部分引起，一部分是中面力引起的中面位移所产生的应变，另一部分是由挠度所引起的应变。中面位移所产生的应变（ε_x^p，ε_y^p，γ_{xy}^p）可由弹性力学平面问题的几何方程来表示：

$$\varepsilon_x^p = \frac{\partial u^p}{\partial x},\ \varepsilon_y^p = \frac{\partial v^p}{\partial y},\ \gamma_{xy}^p = \frac{\partial v^p}{\partial x} + \frac{\partial u^p}{\partial y} \qquad (3.6\text{-}132)$$

而挠度所引起的应变（ε_x^L，ε_y^L，γ_{xy}^L）与位移间的几何关系为

$$\varepsilon_x^L = \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2,\ \varepsilon_y^L = \frac{1}{2}\left(\frac{\partial w}{\partial y}\right)^2,\ \gamma_{xy}^L = \frac{\partial w}{\partial x}\frac{\partial w}{\partial y} \qquad (3.6\text{-}133)$$

由于薄板大挠度变形仍然满足克希霍夫理论中的直法线假设，所以大挠度薄板内任一点的全应变（ε_x，ε_y，γ_{xy}）为

$$\varepsilon_x = \frac{\partial u^p}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2 - z\frac{\partial^2 w}{\partial x^2} \qquad (3.6\text{-}134a)$$

$$\varepsilon_y = \frac{\partial v^p}{\partial y} + \frac{1}{2}\left(\frac{\partial w}{\partial y}\right)^2 - z\frac{\partial^2 w}{\partial y^2} \qquad (3.6\text{-}134b)$$

$$\gamma_{xy} = \frac{\partial v^b}{\partial x} + \frac{\partial u^b}{\partial y} + \frac{\partial w}{\partial x}\frac{\partial w}{\partial y} - 2z\frac{\partial^2 w}{\partial x\partial y} \qquad (3.6\text{-}134c)$$

上式也可写成矩阵形式，如

$$\{\varepsilon\} = \{\varepsilon\}^p + \{\varepsilon\}^b + \{\varepsilon\}^L \qquad (3.6\text{-}135)$$

式中，$\{\varepsilon\}$ 是几何非线性翘曲变形的应变列阵；$\{\varepsilon\}^p$ 是平面应力问题的应变列阵；$\{\varepsilon\}^b$ 是薄板弯曲问题的应变列阵；可参见几何线性模型的相关部分；而 $\{\varepsilon\}^L$ 是薄板大挠度位移所产生的应变列阵。

根据式（3.6-133）以及平面变形、板弯曲变形有限元公式，有

$$d\{\varepsilon\}^p = [B^p]d\{\delta\}_e = \sum_{1,2,3}[B_i^p]d\{\delta_i\} \qquad (3.6\text{-}136a)$$

$$d\{\varepsilon\}^b = [B^b]d\{\delta\}_e = \sum_{1,2,3}[B_i^b]d\{\delta_i\} \qquad (3.6\text{-}136b)$$

$$d\{\varepsilon\}^L = [B_L]d\{\delta\}_e = \sum_{1,2,3}[B_i^L]d\{\delta_i\} \qquad (3.6\text{-}136c)$$

$$[B_0] = [B^p] + [B^b] \qquad (3.6\text{-}137)$$

$$[B_i^L] = [0_{3\times2}\ B_i^{Lb}\ 0_{3\times1}] \qquad (3.6\text{-}138)$$

式中，$[B_i^p]$ 是平面单元按线性分析所得的应变矩阵，是 $[B^p]$ 的 3×6 阶子阵；$[B_i^b]$ 为薄板单元按线性分析所得到的应变几何矩阵，是 $[B^b]$ 的 3×6 阶子阵；$[B_i^{Lb}]$ 是板单元大变形几何矩阵 $[B_i^L]$ 的子阵，来自几何方程的非线性项。如果令

$$[B^{Lb}] = [B_1^{Lb}\quad B_2^{Lb}\quad B_3^{Lb}] \qquad (3.6\text{-}139)$$

那么要计算 $[B_L]$ 矩阵，也就是要计算 $[B^{Lb}]$，它满足

$$d\{\varepsilon\}^L = [B^{Lb}]d\{\delta\}_e^b \qquad (3.6\text{-}140)$$

其具体计算步骤如下所述。

首先，由式（3.6-133）可写出

$$\{\varepsilon\}^L = \frac{1}{2}\begin{bmatrix}\partial w/\partial x & 0\\ 0 & \partial w/\partial y\\ \partial w/\partial y & \partial w/\partial x\end{bmatrix}\begin{bmatrix}\frac{\partial w}{\partial x}\\ \frac{\partial w}{\partial y}\end{bmatrix} = \frac{1}{2}[C]\{\theta\} \qquad (3.6\text{-}141)$$

式中，$\{\theta\}$ 是 w 的一阶导数列阵。由于有

$$w = [N^b]\{\delta\}_e^b = [N_1^b\quad N_2^b\quad N_3^b]\{\delta\}_e^b \qquad (3.6\text{-}142)$$

式中，$[N^b]$ 为板弯曲形函数矩阵；$\{\delta\}_e^b$ 为单元所有节点的弯曲位移列阵。所以，$\{\theta\}$ 可用上式中的节点位移与形函数的导数表示为

$$\{\theta\} = \begin{bmatrix}\partial w/\partial x\\ \partial w/\partial y\end{bmatrix} = [G]\{\delta^b\}_e \qquad (3.6\text{-}143)$$

$$[G] = \begin{bmatrix}\partial N_1^b/\partial x & \partial N_2^b/\partial x & \partial N_3^b/\partial x\\ \partial N_1^b/\partial y & \partial N_2^b/\partial y & \partial N_3^b/\partial y\end{bmatrix} \qquad (3.6\text{-}144)$$

式中，$[G]$ 为 2×9 阶系数矩阵，且 $[G]$ 中的元素是对直角坐标的偏导数，而形函数是用面积坐标表示的，因此，应将矩阵 $[G]$ 表示为对面积坐标的偏导数，即

$$[G]=\frac{1}{2\Delta}\begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}\begin{bmatrix} \partial N_1^b/\partial L_1 & \partial N_2^b/\partial L_1 & \partial N_3^b/\partial L_1 \\ \partial N_1^b/\partial L_2 & \partial N_2^b/\partial L_2 & \partial N_3^b/\partial L_2 \end{bmatrix} \quad (3.6\text{-}145)$$

方程式（3.6-145）是几何非线性问题中求解矩阵［G］的经典公式，但对于几何非线性翘曲变形有限元分析，矩阵［G］的求解非常复杂。这是因为，其中任何一个元素$\partial N_i^b/\partial L_j$（$i=1，2，3，j=1，2$）都是关于面积坐标$L_1$，$L_2$，$L_3$的非齐次多项式。如果不能采用一个行之有效、简洁的方式表示$\partial N_i^b/\partial L_j$，那么矩阵［$G$］将变得极其复杂，以至于不能继续推导后续的许多公式和进行程序编制。因此，对方程式（3.6-145）的表示方式进行如下变换

$$[G]=\frac{1}{2\Delta}\begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}\begin{bmatrix} LL & 0 \\ 0 & LL \end{bmatrix}\begin{bmatrix} [CC][A_1] & [CC][A_2] & [CC][A_3] \\ [DD][A_1] & [DD][A_2] & [DD][A_3] \end{bmatrix} \quad (3.6\text{-}146)$$

式中，b_1，b_2，c_1，c_2为节点坐标函数；［LL］为1×10阶面积坐标矩阵；［A_i］为9×3阶系数矩阵；［CC］，［DD］均为10×9阶常系数矩阵，且

$$[LL]=[1\ L_1\ L_2\ L_3\ L_1L_2\ L_2L_3\ L_1L_3\ L_1^2\ L_2^2\ L_3^2] \quad (3.6\text{-}147a)$$

$$[CC]=\begin{bmatrix} 1&0&-1&0&0&0&0&0&0\\ 0&0&0&0&-1&0&0&0&0\\ 0&0&0&-1&0&1&0&0&0\\ 0&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&0&-2\\ 0&0&0&0&0&0&-2&0&0\\ 0&0&0&0&0&0&0&4&0\\ 0&0&0&0&0&0&0&-1&0\\ 0&0&0&0&0&0&1&0&1\\ 0&0&0&0&0&0&0&-1&0 \end{bmatrix} \quad (3.6\text{-}147b)$$

$$[DD]=\begin{bmatrix} 0&1&-1&0&0&0&0&0&0\\ 0&0&0&0&-1&1&0&0&0\\ 0&0&0&-1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&2\\ 0&0&0&0&0&0&-4&0&0\\ 0&0&0&0&0&0&0&2&0\\ 0&0&0&0&0&0&0&-1&-1\\ 0&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0&0 \end{bmatrix} \quad (3.6\text{-}147c)$$

比较方程式（3.6-145）和式（3.6-146），可以发现变换后的矩阵［G］的计算很方便。计算矩阵［G］后，利用方程式（3.6-110）中直角坐标和面积坐标的偏导数之间的关系，并由方程式（3.6-141）可推导出矩阵［C］的表达式

$$[C]=\frac{1}{2\Delta}\begin{bmatrix} b_1 & b_2 & 0 & 0 \\ 0 & 0 & c_1 & c_2 \\ c_1 & c_2 & b_1 & b_2 \end{bmatrix}\begin{bmatrix} \frac{\partial w}{\partial L_1} & \frac{\partial w}{\partial L_2} & 0 & 0 \\ 0 & 0 & \frac{\partial w}{\partial L_1} & \frac{\partial w}{\partial L_2} \end{bmatrix}^{\mathrm{T}} \quad (3.6\text{-}148)$$

方程式（3.6-148）是几何非线性有限元分析中求解矩阵［C］的经典方式，矩阵中的$\partial w/\partial L_1$和$\partial w/\partial L_2$也是关于面积坐标L_1，L_2，L_3的非齐次多项式，同样也存在着计算复杂的现象。因此，借用矩阵［G］的变换思路，将矩阵［C］中的$\partial w/\partial L_1$和$\partial w/\partial L_2$变成基于矩阵乘积的简单表达方式，即

$$\frac{\partial w}{\partial L_1}=\sum_{1,2,3}[LL][CC][A_i][\delta_i^b] \quad (3.6\text{-}149a)$$

$$\frac{\partial w}{\partial L_2}=\sum_{1,2,3}[LL][DD][A_i][\delta_i^b] \quad (3.6\text{-}149b)$$

至此，利用方程式（3.6-149），可以方便地推导矩阵［C］。对于矩阵［C］，还可证明有下列性质

$$(\mathrm{d}[C])\{\theta\}=[C]\mathrm{d}\{\theta\} \quad (3.6\text{-}150)$$

再将方程式（3.6-141）代入上式可得

$$\mathrm{d}\{\varepsilon\}^{\mathrm{L}}=\frac{1}{2}(\mathrm{d}[C])\{\theta\}+\frac{1}{2}[C]\mathrm{d}\{\theta\}=[C]\mathrm{d}\{\theta\}=[C][G]\mathrm{d}\{\delta\}_e^b \quad (3.6\text{-}151)$$

对照式（3.6-140）和上式，不难看出

$$[B^{\mathrm{Lb}}]=[C][G] \quad (3.6\text{-}152)$$

因此，根据矩阵［C］和［G］，很容易计算出矩阵［B^{Lb}］。至此，几何矩阵［B_0］，［B_{L}］，［$\overline{B}$］就不难计算了。

7　注射成形模拟技术新进展

7.1　传统CAE软件的局限性

塑料注射成形CAE软件的发展十分迅速，其在全面提升模具设计水准中的显著效果正逐渐为模具界所认识。由于算法的局限性，一直以来，充填和保压模拟软件都是采用基于“中心层”模型（Middle Plane）的有限元/有限差分方法来分析。所谓中心层是假想的位于模具型腔和型芯中间的层面，其模拟过程如图3.6-10所示。基于这种型腔模型的CAE软件在应用中具有很大的局限性，主要表现在：

1）CAE软件的使用人员必须理解中心层的概念，用户直接由产品模型构造中心层感到困难；

2）独立开发的CAE系统造型功能往往很差，依据模腔的CAD模型自动生成中心层模型又十分困难，这极大地妨碍了CAE软件的推广和普及；

3）由于CAD、CAE软件的模型不统一，二次建模不可避免，设计效率因此大打折扣，CAD、CAE的集成也不可能实现。

新近发展起来的基于实体/表面模型的模拟分析引入型腔的表面模型来取代其中心层模型，这一创新使得由现行的模具设计CAD系统产生的型腔模型可以直接应用于CAE软件，突破了传统软件的建模瓶颈。华中科技大学模具技术国家重点实验室和澳大利亚MOLDFLOW公司这方面的研究已走在世界前列，分别推出了成熟的商品化软件。基于型腔实体/表面模型的模拟过程如图3.6-11所示。

7.2　理论与实现

如前所述，注射成形的充填和保压模拟在型腔壁厚方向上通常采用对称边界条件，为讨论方便，复述如下：

$$u=v=0,\ T=T_{\mathrm{W}}\quad z=b\text{ 处} \quad (3.6\text{-}153a)$$

$$\frac{\partial u}{\partial z}=\frac{\partial v}{\partial z}=0,\ \frac{\partial T}{\partial z}=0\quad z=0\text{ 处} \quad (3.6\text{-}153b)$$

式中，T_{W}为模壁温度（如图3.6-12a所示）。

结合充填和保压过程的控制方程与边界条件可以推出：速度、温度、压力等主要参量在壁厚方向上关于中心层（$z=0$）对称，并由此可得：型腔半壁厚上的平均速度$\overline{u}$，$\overline{v}$与型腔全壁厚上的平均速度$\overline{u}$，$\overline{v}$相等。基于这一重要特征，

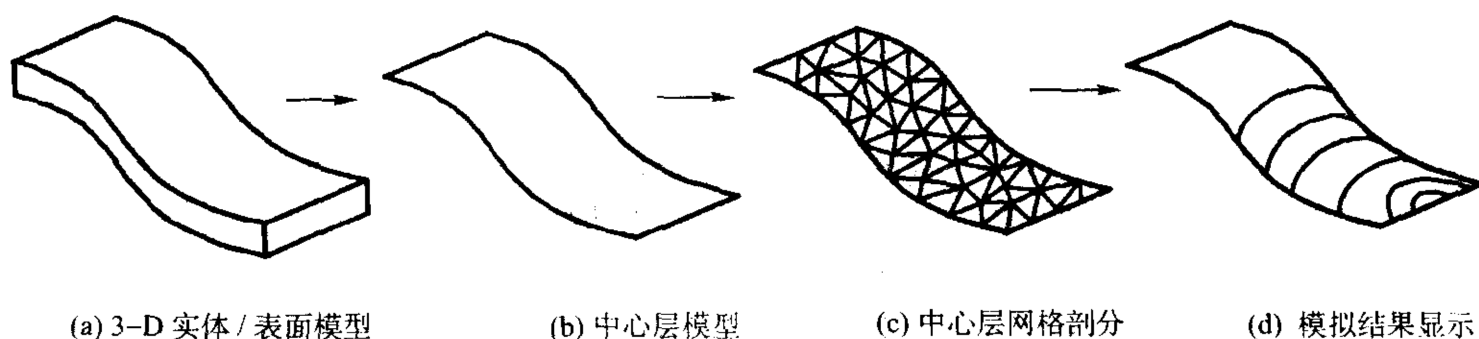


图 3.6-10　基于中心层模型的模拟过程

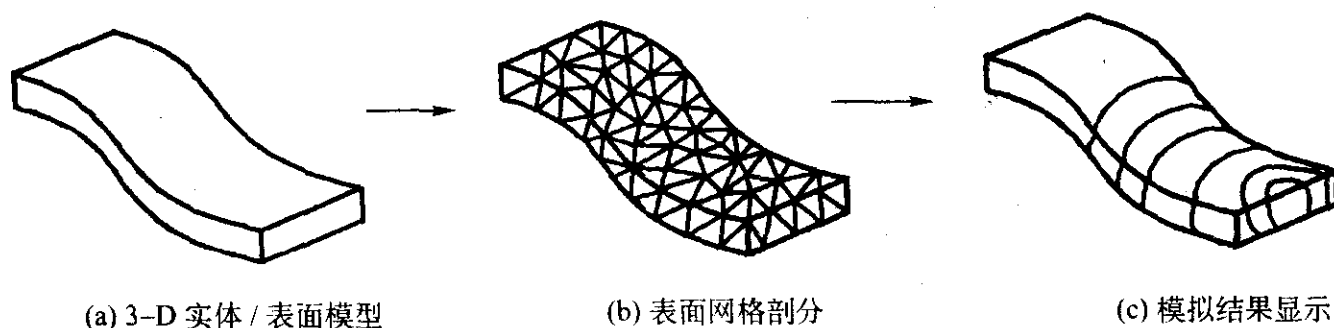
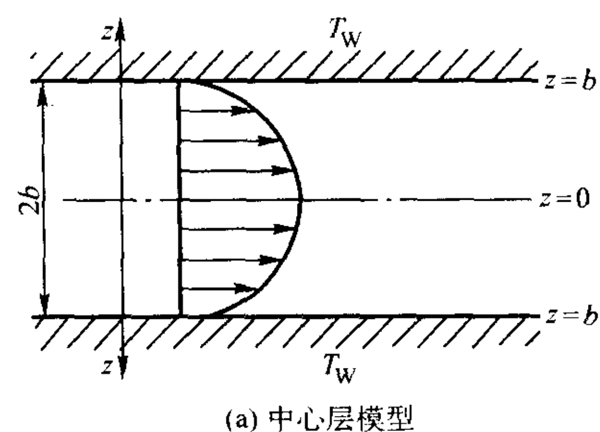
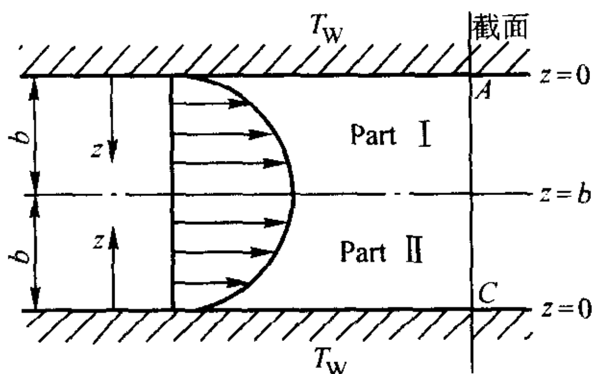


图 3.6-11　基于实体/表面模型的模拟过程



(a) 中心层模型



(b) 表面模型

图 3.6-12　壁厚方向上的边界条件

可以将整个型腔在壁厚方向上分成两部分，如图 3.6-12b 中的 Part Ⅰ和 Part Ⅱ。与此同时，三角形有限元网格在型腔的表面产生（图 3.6-12b 中的 $z=0$ 处），而不是中心层（图 3.6-12a 中的 $z=0$ 处）。相应地，与基于中心层的有限差分是在中心层的两侧进行（从中心层至两模壁）不同，此处壁厚方向上的有限差分仅在表面的内侧（从模壁至中心层）进行，即从图 3.6-12b 中的 $z=0$ 处至 $z=b$ 处。考虑到上述修正，整个分析的坐标系也同时由图 3.6-12a 所示调整为图 3.6-12b 所示。进行上述的处理后，熔体流动过程的控制方程并没有改变，而壁厚方向上的边界条件则为：

$$u=v=0,\ T=T_W\quad z=0\text{处} \tag{3.6-154a}$$

$$\frac{\partial u}{\partial z}=\frac{\partial v}{\partial z}=0,\ \frac{\partial T}{\partial z}=0\quad z=b\text{处} \tag{3.6-154b}$$

同时，还需要在壁厚方向上引进新的边界条件，以保证同一截面处对应的两部分能够协调流动，即：

$$u_{\mathrm{I}}=u_{\mathrm{II}},\ v_{\mathrm{I}}=v_{\mathrm{II}},\ T_{\mathrm{I}}=T_{\mathrm{II}},\ P_{\mathrm{I}}=P_{\mathrm{II}}\ z=b\text{处} \tag{3.6-155}$$

$$C_{\mathrm{m-I}}=C_{\mathrm{m-II}} \tag{3.6-156}$$

式中，下标Ⅰ，Ⅱ分别表示同一截面处对应的两部分的物理量；$C_{\mathrm{m-I}}$、$C_{\mathrm{m-II}}$ 表示同一截面处对应两部分的自由移动流动前沿（如图 3.6-13 所示）；C_o，C_i，C_e 分别为型腔、型芯和浇口的边界。

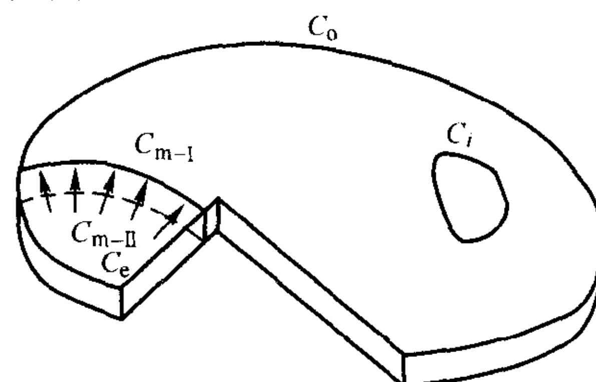


图 3.6-13　对应表面的流动前沿示意图

新引入的边界条件式（3.6-155）和式（3.6-156）是基于实体/表面模型模拟的关键，必须在数值实施过程中予以保证，这给数值计算过程带来了新的复杂性和难度，主要表现在：

1）同一截面处对应表面的有限元三角网格不可能完全对称（网格节点是独立生成的），如图 3.6-12b 中处于同一截面的 A、C 点，其对应的网格可能是错开的，因此在 u，v，T，P 等物理量的计算中必须进行插值处理；

2）由于对应表面具有各自的流动前沿，考虑同一截面处的两对应节点（如图 3.6-12b 中的 A、C 点），两点所对应的控制体积都充满或者一点已满另一点尚未充满的可能性都存在，这两种情况需要分别进行相应的处理；

3）由于与 1）相同的原因，对应表面的流动前沿总存在微小差别，其允差的调节可通过流前节点的时间差别控制或位置差别控制来实现；

4）每次流动前沿的更新都会扩展熔体的流动范围，因此，每次流动前沿更新后，都必须检查式（3.6-155）是否满足，若不满足需做相应的处理，直至满足为止；

5）上述1）~4)中的各种处理操作都需要对同一截面处三角网格的物理量进行比较与调整，因此，要在模拟前将处于同一截面的三角网格挑选出来，即必须在所有三角网格中进行配对操作。

6）由于在计算过程有一些节点的压力、温度等参数直接由其对应节点给定，这实际上改变了有限元方法中形成的整体刚度矩阵，导致数值计算稳定性下降，因此有必要寻求稳定性更优的数值算法。

基于上述的处理，通过与中心层模型模拟分析相类似的推导与求解，可以获得基于表面/实体模型的分析结果，这里不再重复。

7.3 软件介绍

在基于实体模型的模拟算法的基础上，华中科技大学模具技术国家重点实验室在国内首次研究开发了新一代的三维真实感模拟软件 HSCAE3D。该软件无需根据零件的 CAD 模型再构造产品的中心层模型，而是直接利用三维实体模型进行分析和显示。系统首先利用实体的表面信息生成模型表面的三角网格。对于薄壁塑料制品，可对所有三角网格的节点进行厚度方向的配对，产生节点配对信息及网格厚度信息。然后再利用网格数据、节点配对信息进行流动保压模拟，产生的分析结果直接在三维实体表面上显示。

虽然本系统采用的数值模拟本质上仍然是二维有限元/一维有限差分方法，但是由于采用了表面/实体模型，这种方法回避了中心层的建模瓶颈，成功解决了现行 CAE 软件的不足，CAD、CAE 软件的集成也切实可行。在结果显示方面，由于三维真实感模拟直接将分析结果显示在实体上，这比在中心层面上的显示更加直观、逼真、生动，具有三维真实感的效果。

软件的执行的流程如图 3.6-14 所示。首先读取塑料制品的 STL 文件，显示制品图形，构造邻接拓扑关系，搜索制品边界（用于线框模型的显示）。然后进行表面重整、网格划分、网格优化以及节点配对。在此基础上用户定义浇口、设置工艺条件后即可进行流动保压分析。流动和保压分析结束后用户可以直观地浏览各项分析结果，包括流动前沿、气穴、熔接缝、温度场、压力场、剪切力场、剪切速率场、表面纤维定向、收缩率场、密度场等数据。

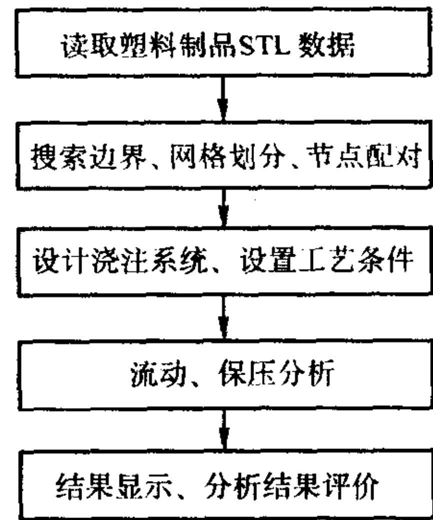


图 3.6-14 软件执行流程图

编写：周华民（华中科技大学）

第7章　常用材料成形软件简介

1　冲压成形模拟软件

1.1　FASTAMP软件

1.1.1　简介

FASTAMP软件是由华中科技大学塑性成形模拟及模具技术国家重点实验室开发研制的，具有完全自主版权的板料冲压成形与模具设计快速分析软件。FASTAMP软件采用了基于有限元逆算法（Inverse Approach）的求解器，集成了国际水平的有限元前后处理模块，实现了真正意义上的板料成形模拟CAE系统。

FASTAMP软件采用全中文界面，易懂易学。前处理模块具备兼容性极强的标准CAD接口，高质量曲面自动网格生成器，以及功能强大的点/线/曲面/单元编辑功能，可以帮助用户建立起完整的有限元计算模型。

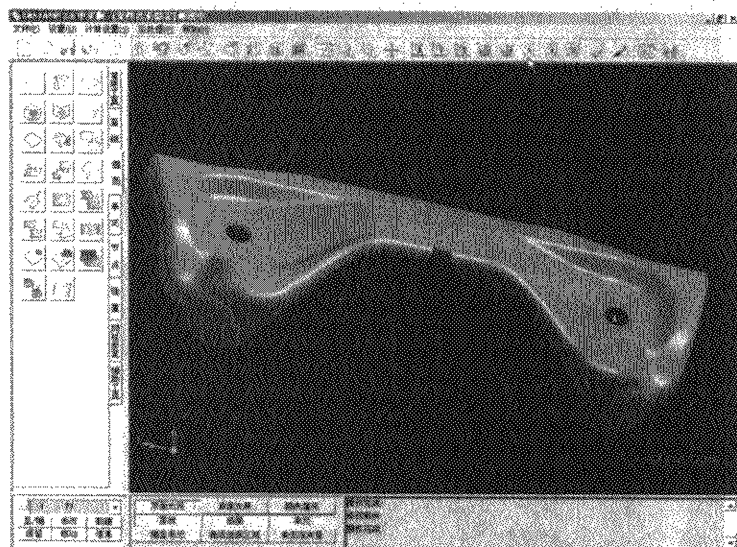
FASTAMP软件计算速度快，计算精度高，模拟功能强大，主要面向产品设计、选材、工艺设计、模具设计人员，突破传统覆盖件CAE软件只能提供给少数的专业分析人员使用的制约瓶颈。由于采用了独特的处理方式，能够在数值仿真过程中实时动态地处理摩擦、压边力和拉深肋等工艺参数，可以真实地反映它们的作用效果，从而突破了以往逆算法只能用于粗略预测简单零件坯料形状而不能应用于汽车覆盖件成形精确模拟的限制。

FASTMP软件还可以模拟两步成形和多步成形过程，可以近似地模拟单动压机、双动压机、三动压机等类型压机成形过程的问题，是对传统逆算法模拟功能的一个大的突破。FASTAMP另一个突出特点是可以快速预测覆盖件三维翻边过程的可成形性，精确确定三维修边线，彻底改变传统修边模和翻边模设计过程中依靠简单的解析理论和经验公式的现象，大幅度提高修边模和翻边模的设计效率。

由于集成了完整的有限元前处理模块，在算法理论上取得了新的突破，FASTAMP软件无论是从计算精度、速度，还是软件的功能和易用性方面来看，都和国外同类知名软件达到了同一水平。可广泛应用于板料冲压成形模具设计与成形性校核，为模具工艺方案选择、模具结构设计、工艺参数优化提供快速有效的数值分析，也可以应用于冲压件的可成形性快速模拟。精确反算冲压件或零件的坯料形状，快速预测冲压件的厚度分布、应变分布、破裂位置、起皱位置等。

1.1.2　应用实例

（1）精确的坯料形状反算



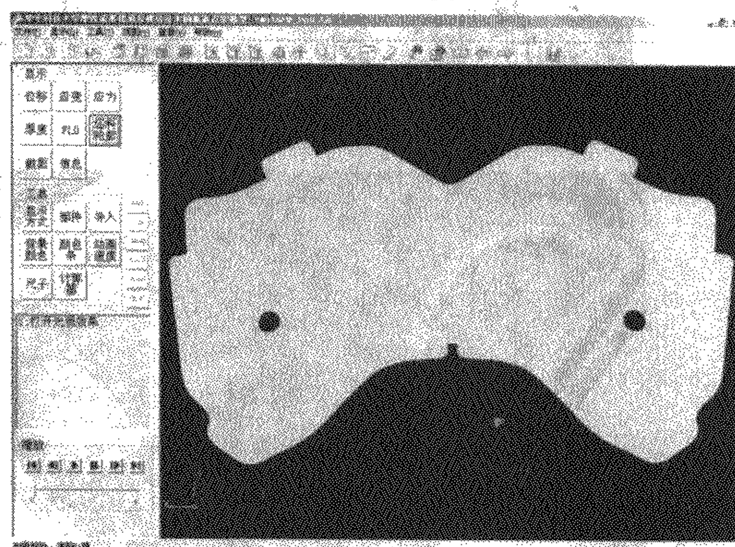
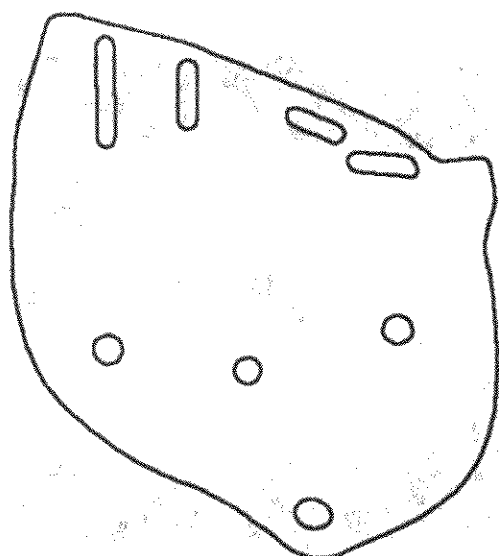


图3.7-1　坯料形状精确反算

图3.7-1零件是某汽车模具厂实际生产的翻边与拉深混合成形件。由于成形条件复杂，难以准确把握变形规律，利用传统的方法无法确定出该零件的合理坯料尺寸。利用FASTAMP软件进行求解，得出了零件的初始坯料形状，经实际实验对比，仅在个别区域存在轻微误差0.2 mm，完全可以满足零件的质量要求。

（2）翻边成形模拟（图3.7-2）

翻边成形是FASTAMP软件最成功的工业应用，FASTAMP可以精确地反算毛坯形状，预测翻边过程中的起皱和破裂，提供最佳的工艺方案，包括预估压料力大小、毛坯定位、冲孔工序等。



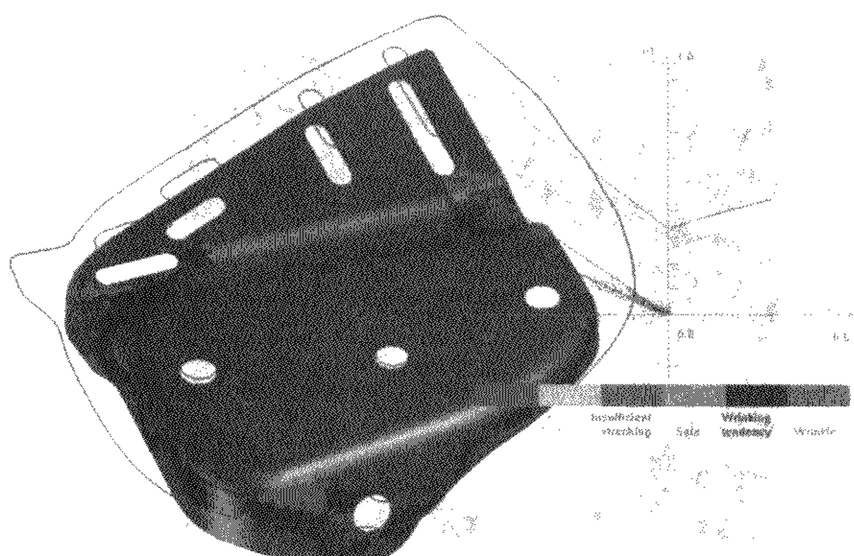


图3.7-2　坯料形状精确反算

（3）两步及多步成形模拟（图3.7-3）

FASTAMP软件对传统的逆算法进行了改进，可以精确

模拟两步成形和多步成形，并可以将零件中间展开到曲压料面上，真实模拟压边过程。

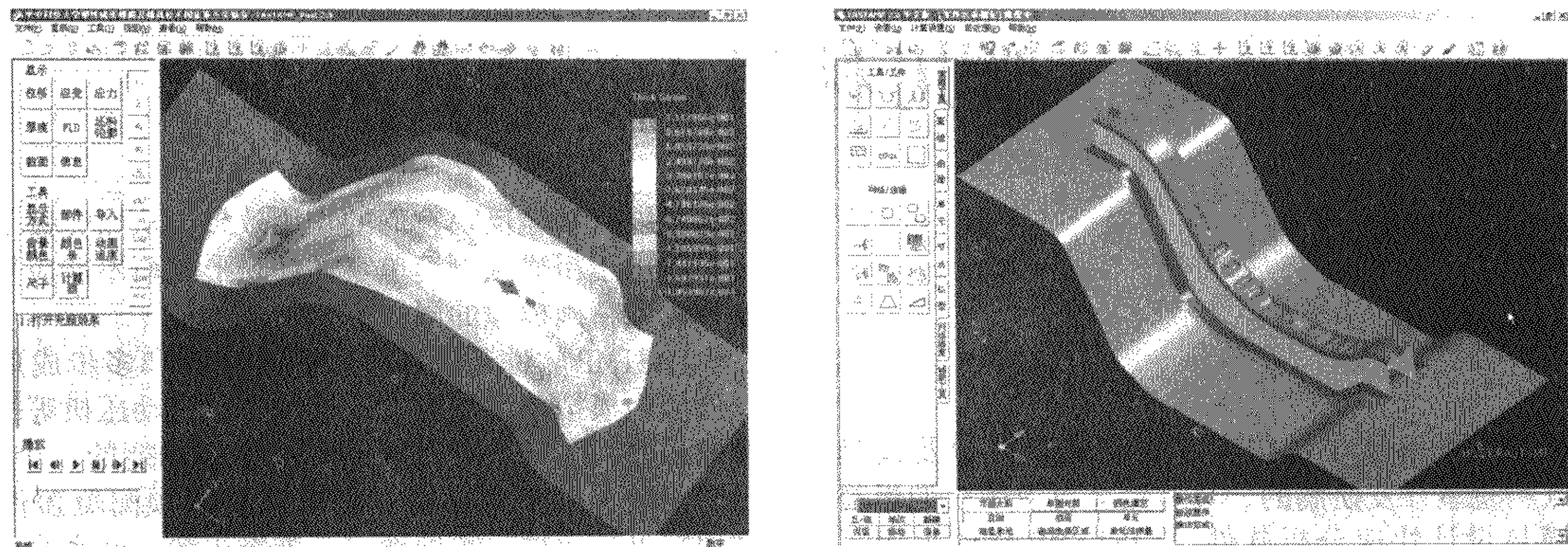


图 3.7-3　两步成形模拟

(4) 三维翻边成形与三维修边线模拟（图 3.7-4）

FASTAMP 软件可以快速预测覆盖间的三维翻边过程的可成形性，精确确定三维修边线。并可将计算得出的修边线直接导入到 UG 等造型软件中，帮助设计人员制定模具设计方案，从而大幅度提高了修边模具和翻边模具的设计效率。

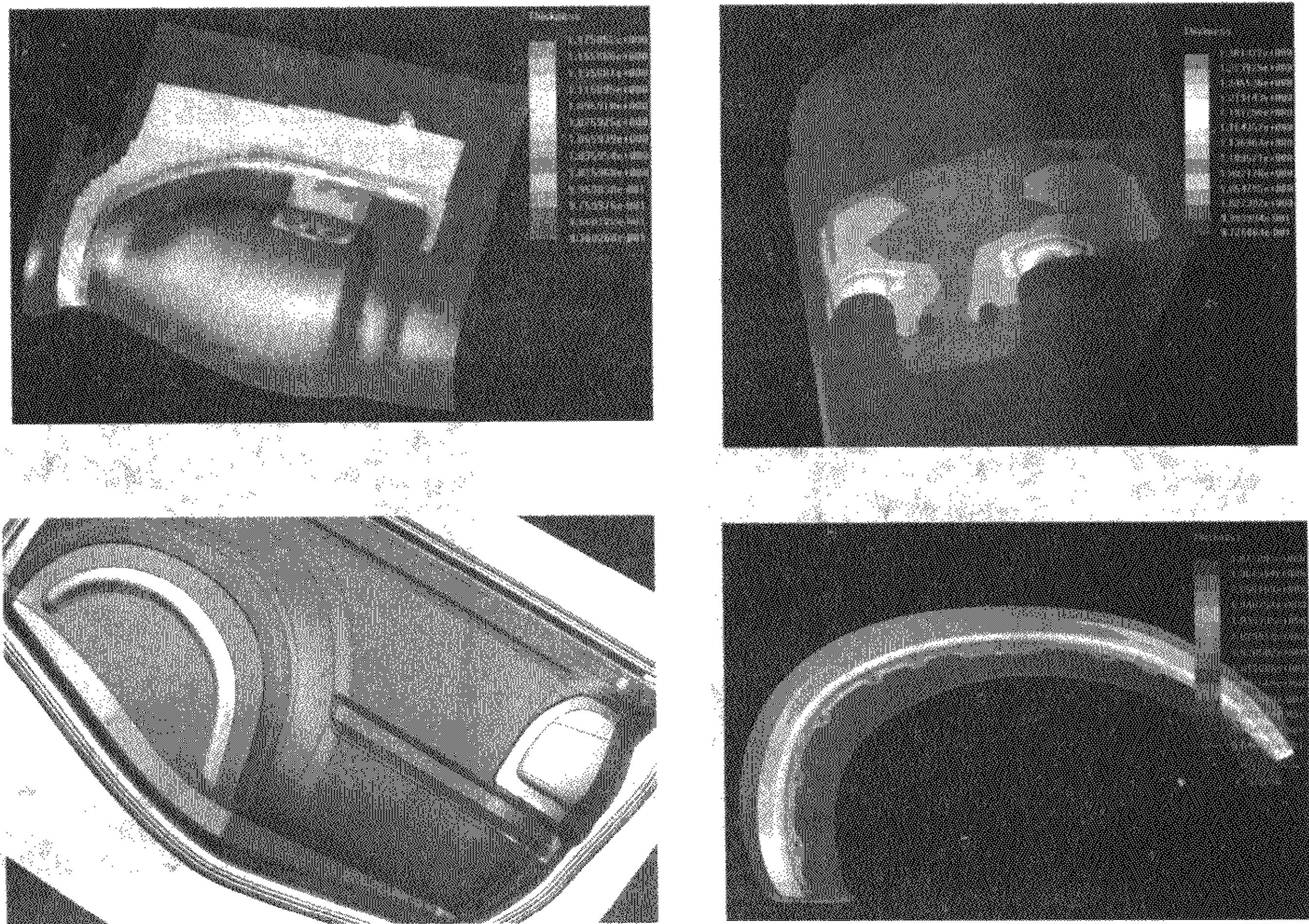
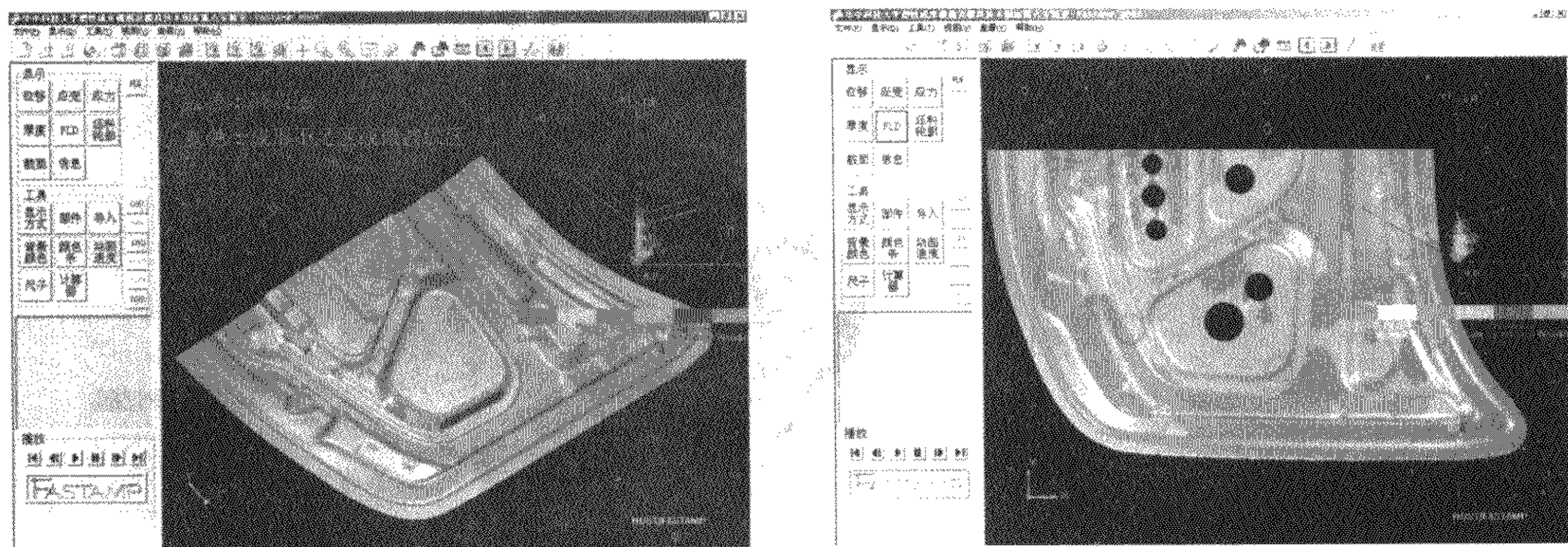


图 3.7-4　三维翻边成形与三维修边线模拟

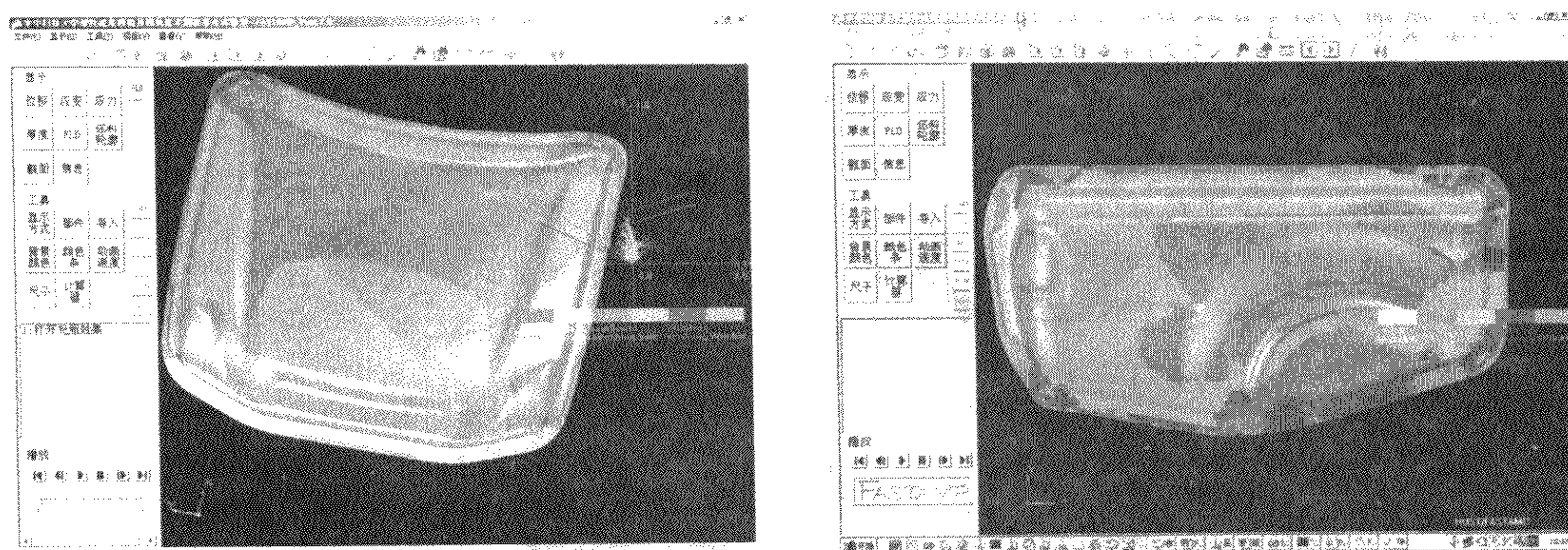
(5) 汽车覆盖件冲压工艺优化与成形性分析（图 3.7-5）

FASTAMP 软件针对汽车覆盖件作了专门处理，可以快速模拟覆盖件成形后的起皱、破裂、成形不足等缺陷；可以优化压边力、拉深肋、摩擦等工艺参数；可以校核压料面和工艺补充面是否合理；可以提供最优的毛坯形状。为汽车覆盖件工艺设计和工艺优化提供全面的解决方案。

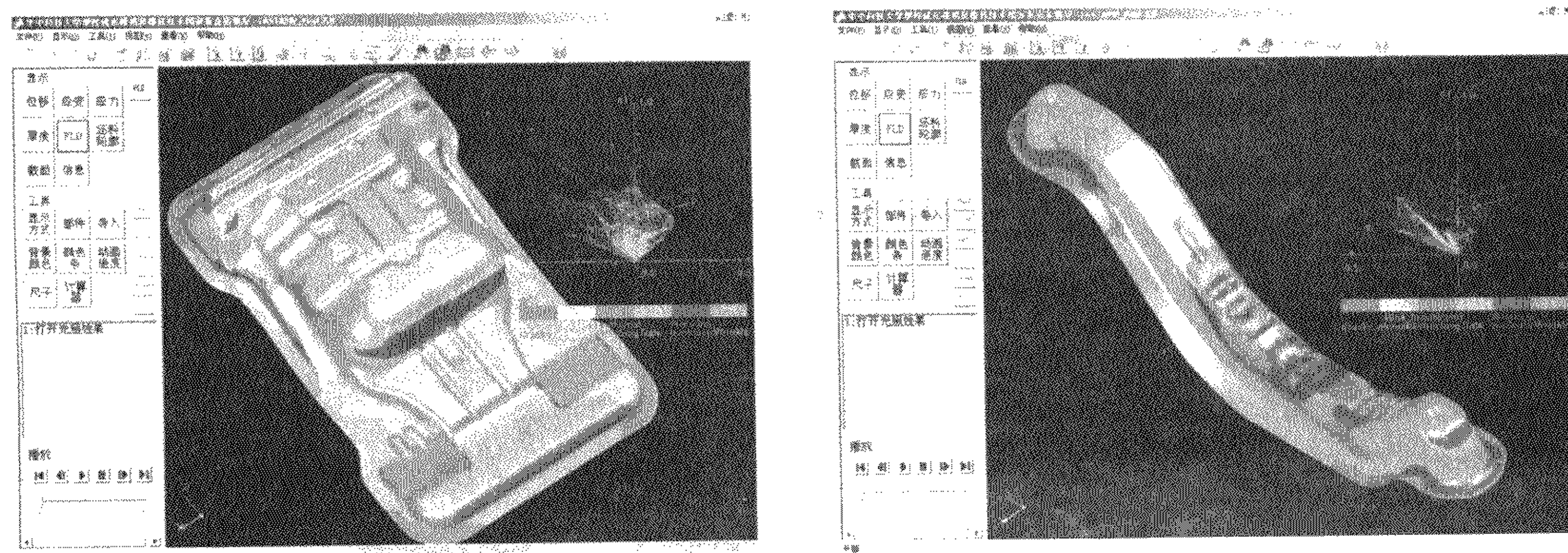


(a) 发动机罩内板的工艺开口优化　　(b) 开口后的成形性分析(FLD与WLD)

图 3.7-5　汽车覆盖件冲压工艺优化与成形性分析



(c) 拉深筋优化(FLD与WLD)



(d) 成形性分析(FLD与WLD)

续图 3.7-5

1.2　DYNAFORM 软件

1.2.1　简介

eta/DYNAFORM 是由美国 ETA 公司开发的用于板料成形模拟的专用软件包，可以帮助模具设计人员显著减少模具开发设计时间及试模周期，不但具有良好的易用性，而且包括大量的智能化自动工具，可方便地求解各类板成形问题。DYNAFORM 包括板成形分析所需的与 CAD 软件的接口、前后处理、分析求解等所有功能。专门用于工艺及模具设计涉及的复杂板成形问题，可以预测成形过程中板料的破裂、起皱、减薄、划痕、回弹，评估板料的成形性能，从而为板料成形工艺及模具设计提供帮助。

目前，除了传统的增量法模块以外，DYNAFORM 中还增加了 BSE（板料尺寸计算）模块，该模块利用一步法（onestep）求解器，可以方便地将产品展开，从而得到合理的坯料尺寸，并可以用于后续的成形性校核。此外，DYNAFORM还包含了用于模具设计的 DFE 模块，可以从零件的几何形状进行模具设计，包括压料面与工艺补充。DFE 模块中包含了一系列基于曲面的自动工具，如冲裁填补功能、冲压方向调整功能以及压料面与工艺补充生成功能等，可以帮助模具设计工程师进行模具设计。

目前，eta/DYNAFORM 已在世界各大汽车、航空、钢铁公司，以及众多的大学和科研单位得到了广泛的应用，自进入中国以来，DYNAFORM 已在长安汽车、南京汽车、上海宝钢、中国一汽、上海汇众汽车公司、洛阳一拖等知名企业得到成功应用。

1.2.2　应用实例

（1）板料尺寸计算（图 3.7-6）

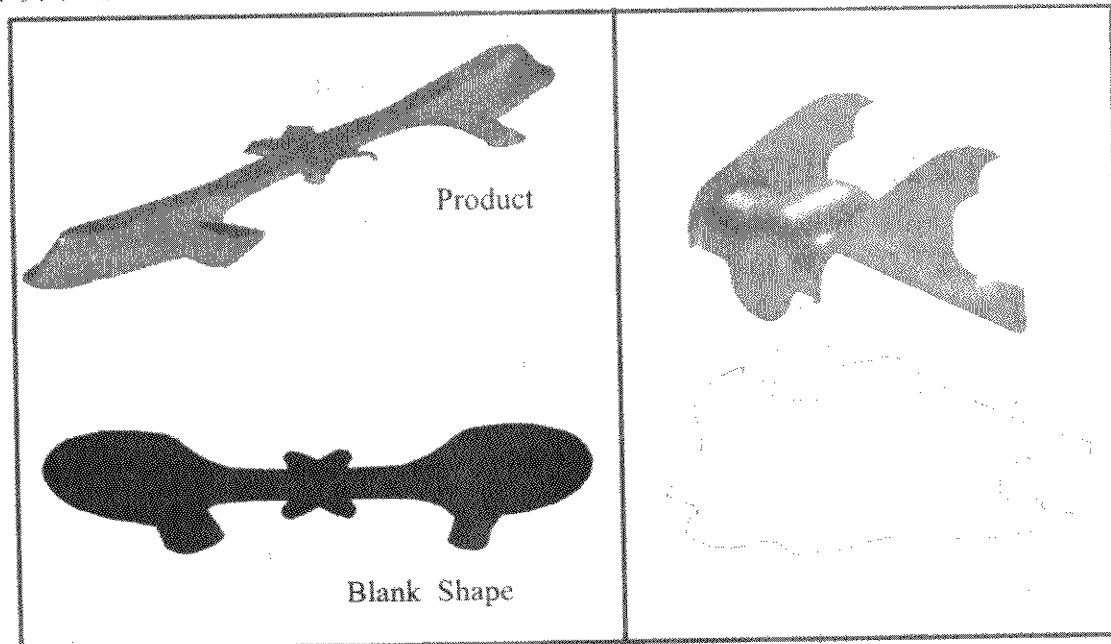


图 3.7-6　毛坯尺寸粗略估算

（2）多工序成形模拟（图3.7-7）

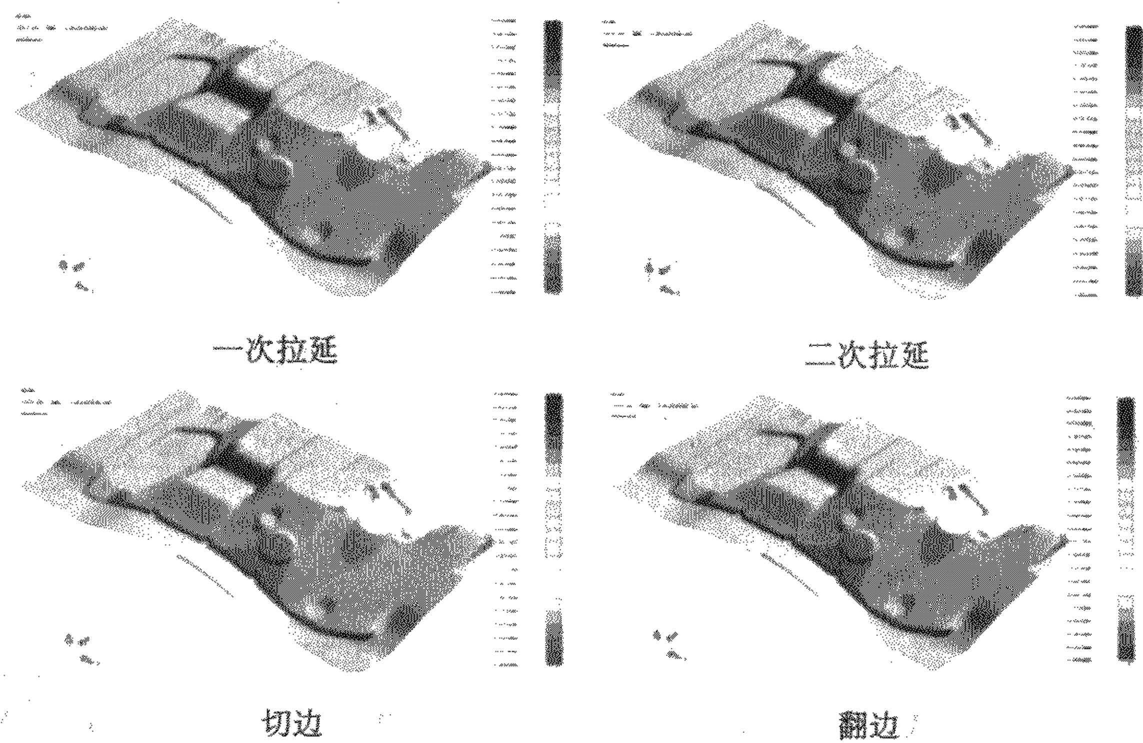


图3.7-7 多工序成形模拟

（3）模面辅助设计（图3.7-8）

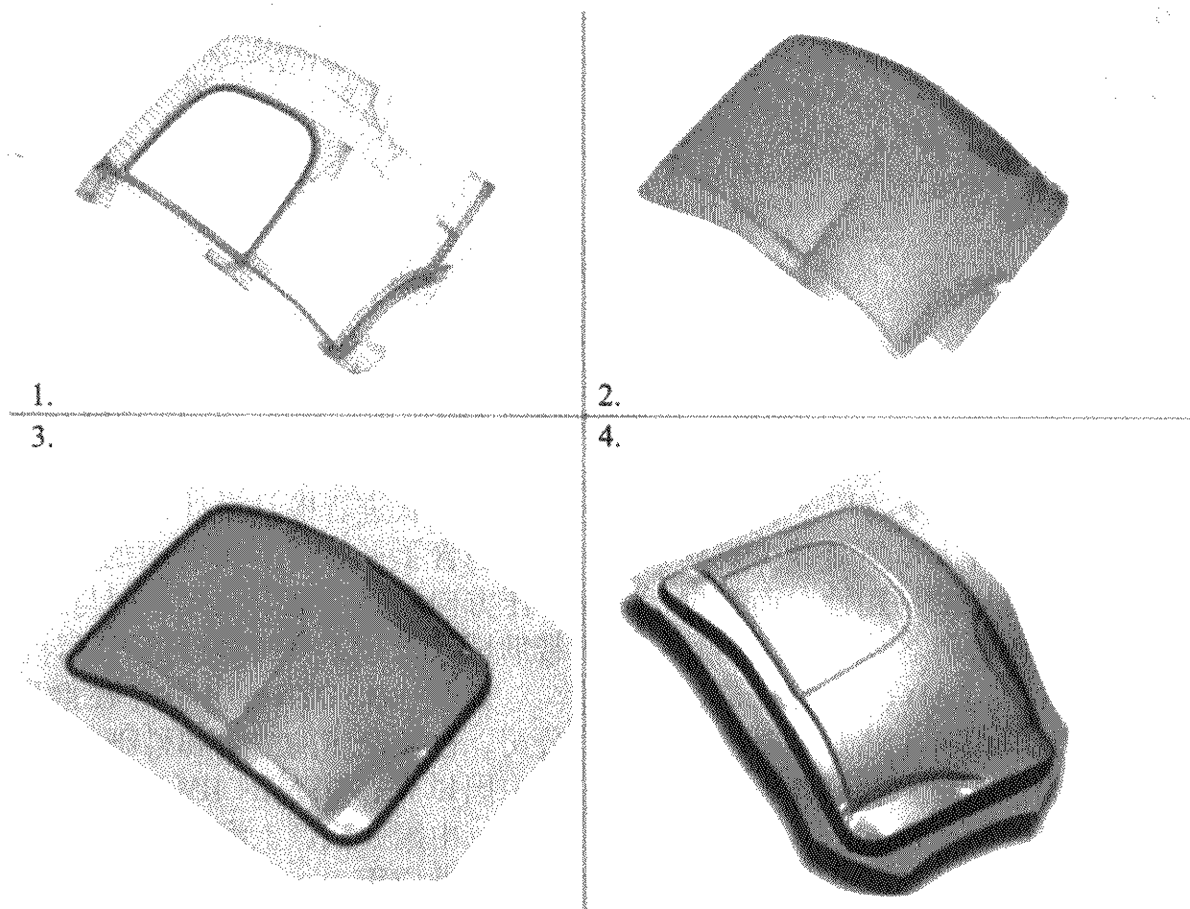


图3.7-8 模面辅助设计

1.3 AUTOFORM软件

1.3.1 简介

AUTOFORM是瑞士联邦工学院开发的板材成形模拟专用软件，目前最新的版本是AUTOFORM 3.2，在欧洲各大汽车企业有着广泛的应用，积累了大量的工程应用经验。它采用了静力隐式算法进行求解，采用了全拉格朗日理论，由于对壳单元面内和横向刚度都进行了解耦，消除了刚度矩阵的病态，保证了计算的收敛性，求解速度很快。

该软件的设计思想突出了易用性和针对性，作为冲压成形模拟的专业软件，它尽可能地简化用户的操作，使软件的功能尽可能自动执行。AUTOFORM可自动进行网格剖分，自动生成和交互修改压料面、工艺补充部分、拉深肋、凸模入口线、板坯材料等，可以自由选择调整冲压方向，产生工艺切口，定义重力作用、压边、成形、修边、回弹等工序或工艺过程。

AUTOFORM提供了众多用于金属板料零件和冲压模具设计的功能模块，贯穿于整个产品发展周期。AUTOFORM同样内置了一步法（onestep）求解器，用户仅需要用户输入零件几何模型就可以进行坯料展开，快速地得到冲压成形模拟结果，并可用于后续成性形分析。增量法模块（increment）则可以精确地模拟冲压成形过程，评估模具设计工艺方案。模具设计模块（DieDesigner）则可以帮助设计人员快速地生成模具型面。此外，AUTOFORM还提供了对工艺参数和几何参数进行优化计算的模块（Optimizer）和应用于液压胀形模拟的模块（Hydro&HydroDesigner）。在最新版的AUTOFORM 3.2中，还加入了用于计算坯料排布的排样模块（Nest）。

1.3.2 应用实例

（1）一步法（Onestep）（图3.7-9）

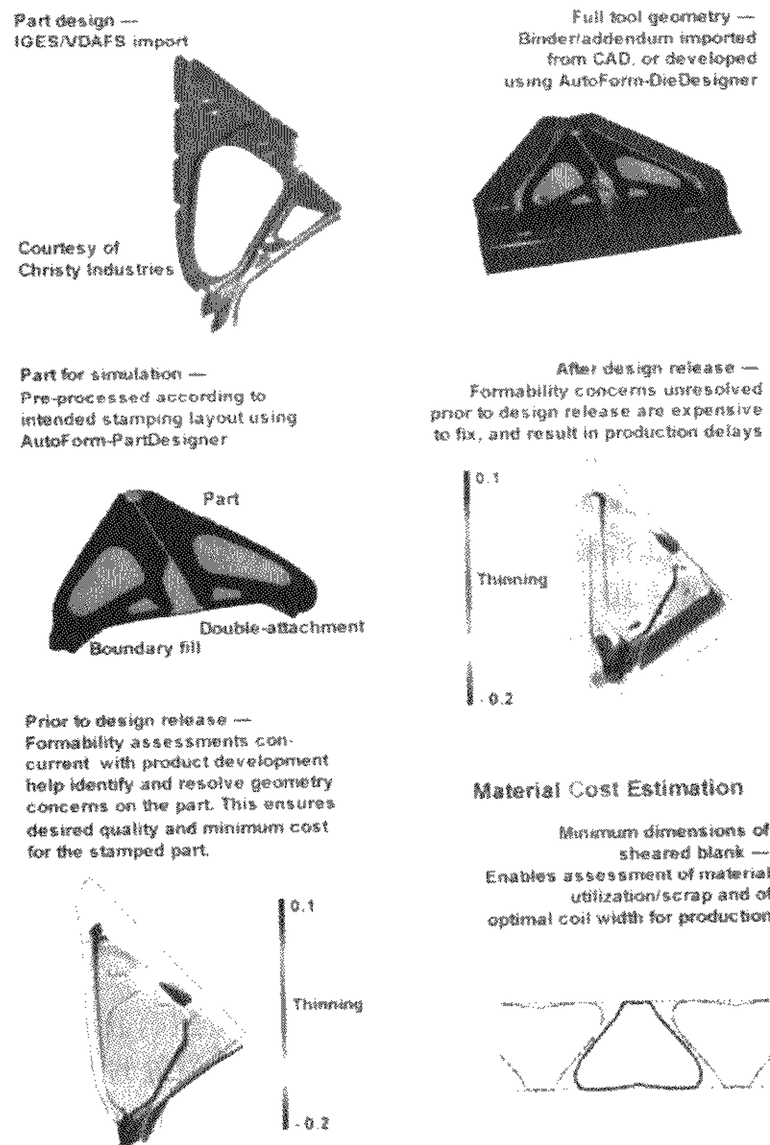


图 3.7-9 一步成形模拟

(2) 增量法（increment）（图 3.7-10）

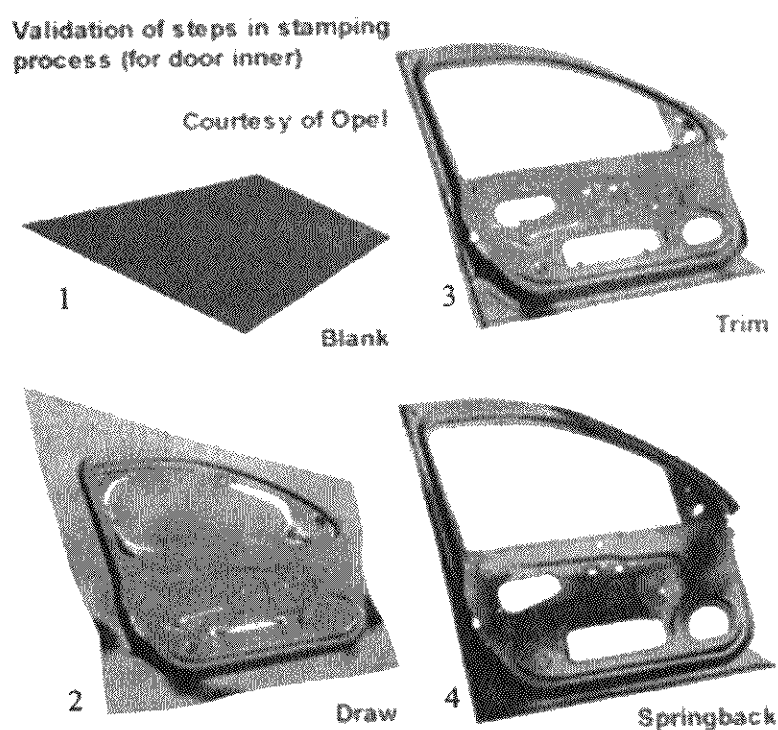


图 3.7-10 成形与回弹过程模拟

(3) 模具设计（DieDesigner）（图 3.7-11）
(4) 液压胀形模拟（hydro）（图 3.7-12）
(5) 排样（Nest）（图 3.7-13）
(6) 优化（Optimizer）（图 3.7-14）



图 3.7-11 压料面与工艺补充设计

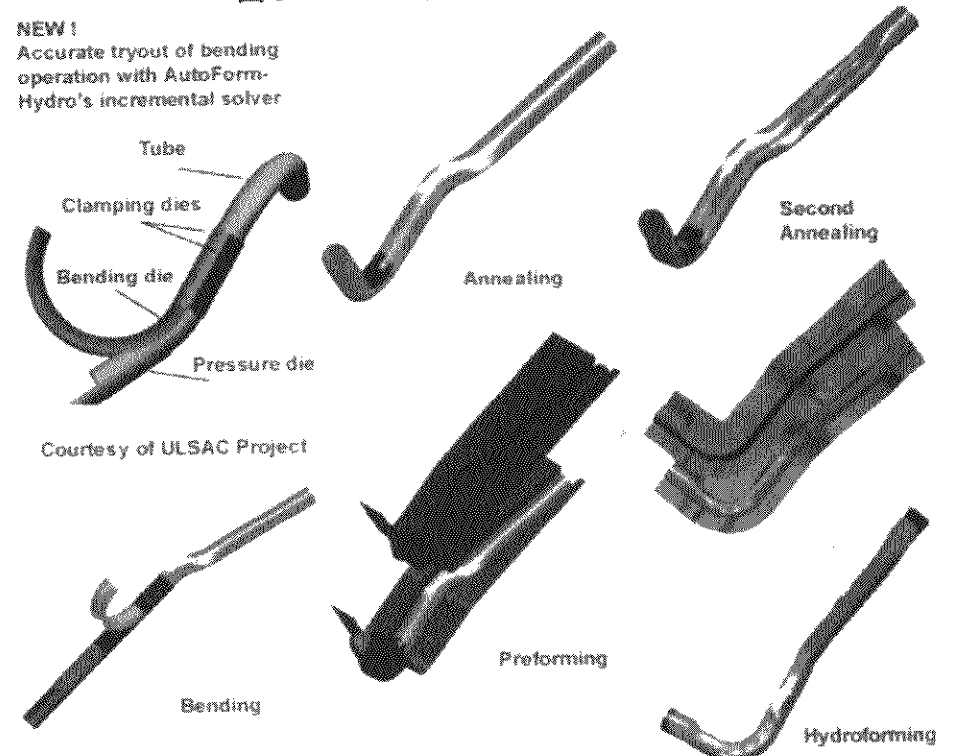


图 3.7-12 液压胀形过程模拟

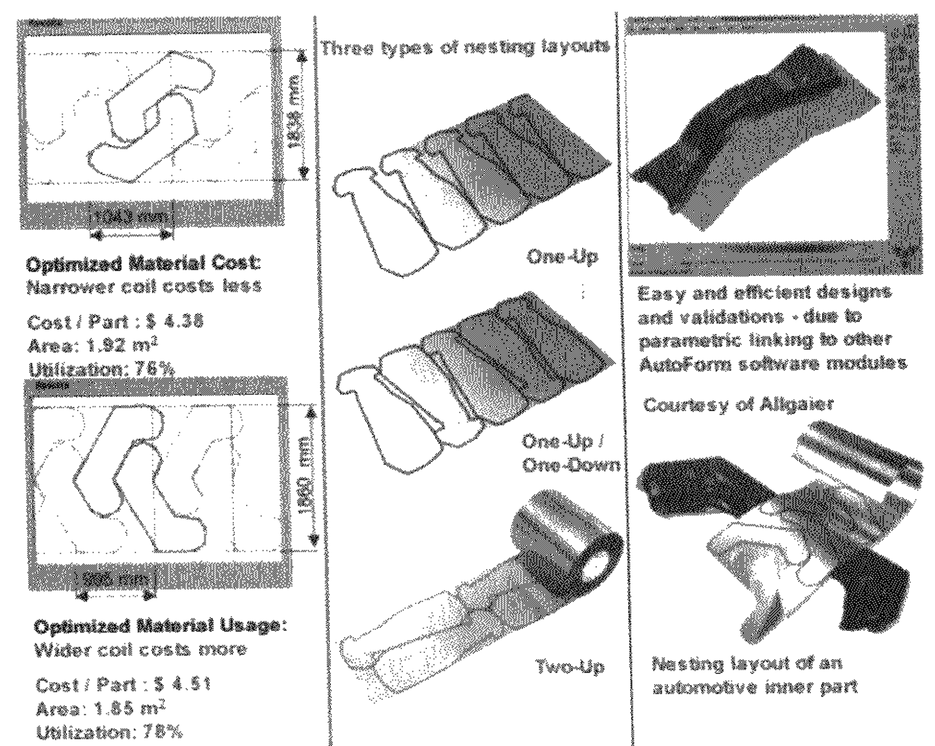


图 3.7-13 毛坯优化排样

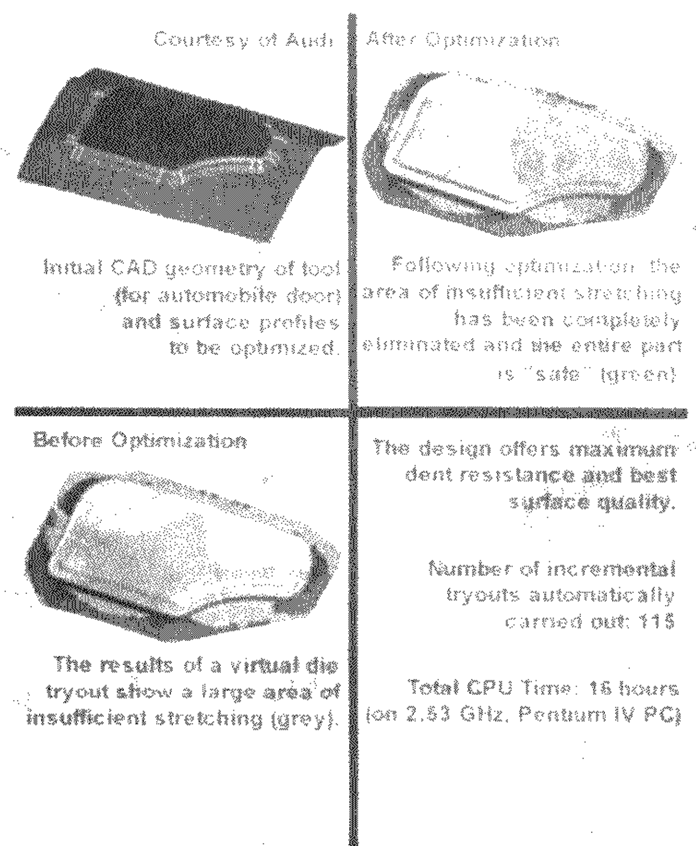


图 3.7-14 成形过程优化模拟

2 体积成形模拟软件

2.1 概述

随着塑性成形有限元数值模拟技术的不断发展，出现了许多商业化有限元分析软件，这些软件在工业生产中得到了广泛应用。常用的有限元分析软件包括：通用非线性有限元软件有美国的 MARC、ABAQUS、NIKE3D 和德国的 INDEED；用于体积成形过程的分析软件有美国的 DEFORM 、AutoForge 和 SuperForge，法国的 FORGE 等等。另外像 LS－DYNA 这种最初主要应用于板壳成形的软件，现在也可以进行如挤压、锻造等体积成形过程的分析。本章主要介绍应用比较广泛的 DEFOEM 和 AutoForge 两种软件。

2.2 DEFORM 软件

2.2.1 简介

体积成形有限元分析软件 DEFORM（Design Environment for Forming）是目前世界上公认的应用较为广泛、功能较强的模拟软件。它建立在 S.Kobayashi 等人研究工作的基础上，并由美国 Battele Columbus 实验室和俄亥俄州立大学精密成形工程中心共同开发，应用了计算机图形技术，并且加入功能丰富的前后置处系统。1990 年开发了自动网格生成（Automatic Mesh Generation）模块，1991 年完成了网格的自动重划分（Automatic remeshing）功能。目前广泛地应用在世界各国的研究机构和实际生产中。

DEFORM 软件是在一个集成环境内综合建模和对成形、热传导和成形设备特性进行模拟仿真分析的过程，SFTC 在 DEFORM 中预先提供了 146 种材料的数据，方便选取。适用于热、冷、温成形，给出成形过程中材料的流动规律、模具填充情况、行程载荷曲线、模具和材料内部的应力应变场、金属微结构和缺陷的产生与发展，为指导试验提供依据。DEFORM 软件可以分析金属成形过程中多个材料特性、不同的关联对象耦合作用下的大变形和热特性，使得分析模型和模拟环境和实际生产环境保持一致。具有的自动网格重划分功能可以根据变形情况自动生成与重生成优化的网格系统，以此能够保证金属成形过程中的模拟精度。该软件分为二维分析模块（DEFORM－2D）和三维分析模块（DEFORM－3D）以及热处理模块。

DEFORM－2D 用来分析平面应变或轴对称等金属成形过程。DEFORM －2D 具有计算平面应变和轴对称问题的功能；能够全自动网格划分和重划分功能，不需要人工干预；可以在前处理中自动生成边界条件，确保数据准备快速可靠；计算时的材料模型可以选择刚性、弹性、刚塑性、热弹塑性、热刚黏塑性、粉末材料及自定义类型等；在成形过程模拟中可以进行热力耦合分析；后处理模块中能够进行点跟踪、应力应变的云图和矢量图、行程载荷曲线、镜像和旋转等处理；能够分析多个塑性工件和组合模具应力；通过选择合适的损伤和裂纹生长模型，分析剪断、冲孔、机加工和冲裁模拟；也具有自接触边界状态处理能力能够自我修正接触边界。

DEFORM－3D 功能与 2D 类似，但它处理的对象为复杂的三维零件、模具等。DEFORM－3D 图形界面既强大又灵活，为用户准备输入数据和观察结果数据提供了有效工具。在 DEFORM－3D 中用于仿真分析的模型几何体，来自 CAD 系统（如 Ideas、Pro/E、UG 或 PATRAN 等）的面或实体造型（UNV、STL/SLA 等）格式。DEFORM－3D 还提供了三维几何操纵修正工具，这对于成形过程模拟极为重要。在 DEFORM－3D 的后处理中，具有 2D 切片功能，这样可以显示复杂工件和模具的剖面，方便观测金属流动和填充情况。同二维一样，输出的结果包括图形、原始数据及动画等。DEFORM－3D 在相当复杂的工业零件，如连杆、曲轴，以及具有复杂筋－翼的结构零件中，都有令人满意的模拟结果。

热处理模块能够模拟的热处理工艺类型有正火、退火、淬火、回火、时效处理、渗碳、蠕变、高温处理以及相变、晶粒变化和时效沉积等。能够精确预测硬度，金相组织体积比值（如马氏体、残余奥氏体含量等），热处理工艺引起的挠曲和扭转变形，残余应力以及含碳量等热处理工艺评价参数。能够基于 Johnson－Mehl 方程和 T－T－T 数据准确预测与扩散相关的相变。用 Magee 方程所描述的剪切过程相关的非弥散性相变，可以作为温度、应力和含碳量的函数来进行计算模拟。根据相硬度或 Jominy 数据能够精确预测热处理工艺处理后的最终硬度分布。每个相变具有各自独立的弹性、塑性、温度和硬化等物理参数。相应综合材料性能则由某一时刻各金相组织类型及其所占比例等因素决定。局部加热和淬火窗口可用于选择部位的热处理工艺。

2.2.2 计算实例

钟形罩是轿车等速万向节部件上最复杂的零件，属于典型的杯杆型零件，杯部为球形内表面和六条弧形滚球道组成，采用多工序温热挤压及冷精整与冷缩径成形，要求球形内表面及六条弧形滚珠球道仅留 0.3～0.4 mm 磨削量。目前，国内已有少数企业开始等速万向节钟形罩的生产，这些企业虽已具备一定的机加工和热处理能力，但因加工所用毛坯一般都是普通锻件，材料利用率低，生产效率低，成本高且年产量也很小。因此研究等速万向节钟形罩精化毛坯的成形技术，已成为我国汽车工业的重要问题。

图 3.7-15 为等速万向节部件结构图，它是由钟形罩、钢珠、保持架、星形套和接合套等组成。在等速万向节的零件中，以钟形罩的制造技术难度最大，主要是形状复杂，加工精度要求高。

图 3.7-16 为钟形罩零件图和预成形件图。由图可知，预成形件仅在零件图上对应部分加上一层机加工余量，以锻件公差代替精密加工公差。

本文在对钟形罩预成形件多种温热成形过程进行有限元模拟的基础上，对成形工艺方案进行分析比较，并对所选工艺方案提出了优化措施。

将钟形罩的三维变形简化为轴对称变形；对钟形罩多工序温热成形过程模拟，采用基于热力耦合的刚粘塑性有限元法，将工件视为刚塑性体，将冲头与凹模视为板壳结构，并假设为刚性体。钟形罩材料为CF53，泊松比 $\nu=0.3$，密度

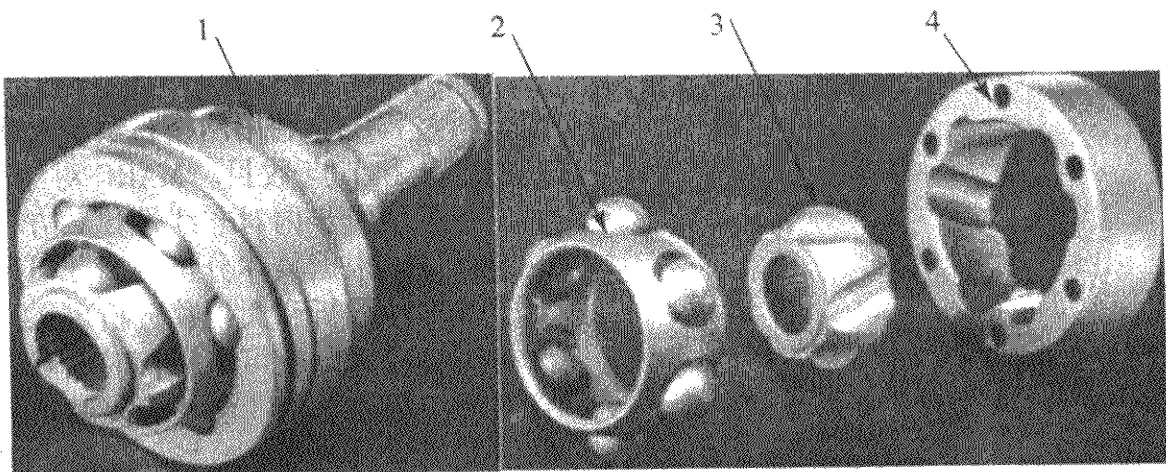
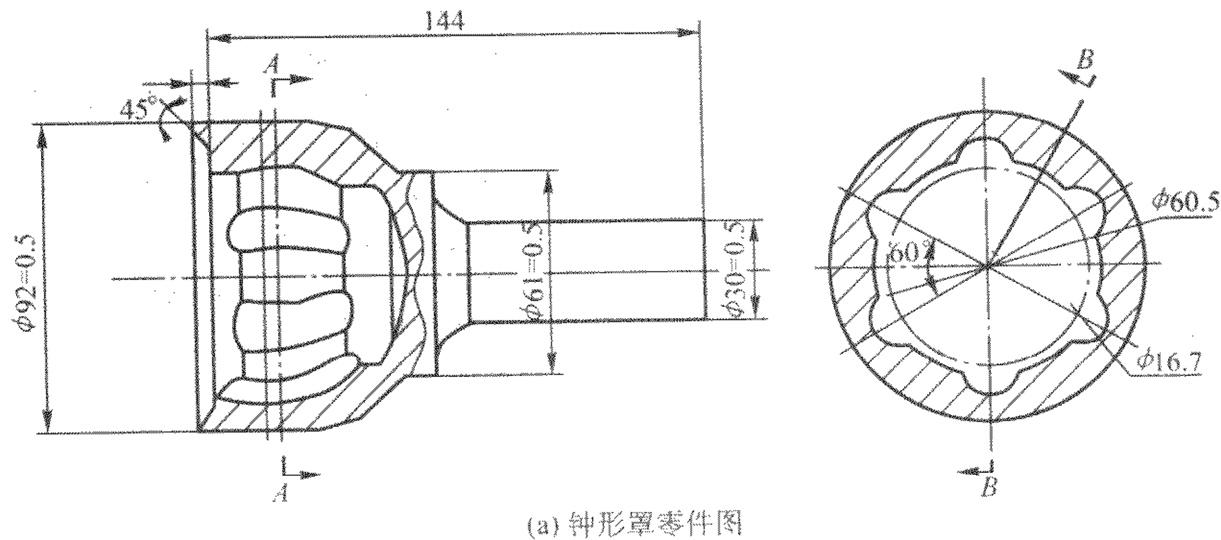
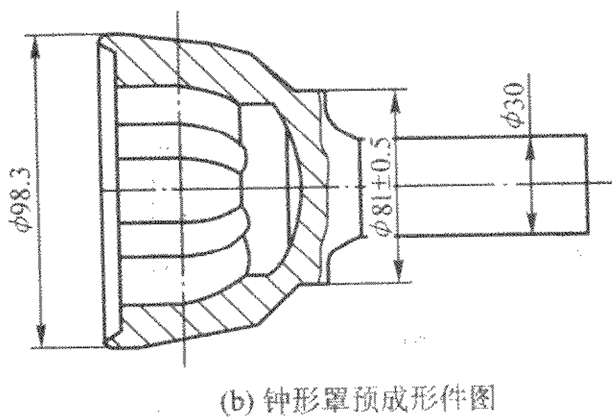


图 3.7-15　等速万向节部件结构简图

1—钟形罩；2—钢珠和保持架；3—星形套；4—接合套



(a) 钟形罩零件图



(b) 钟形罩预成形件图

图 3.7-16　钟形罩零件和预成形件图

$\rho = 7.85\ g/cm^3$，应力－应变曲线如图 3.7-17 所示。

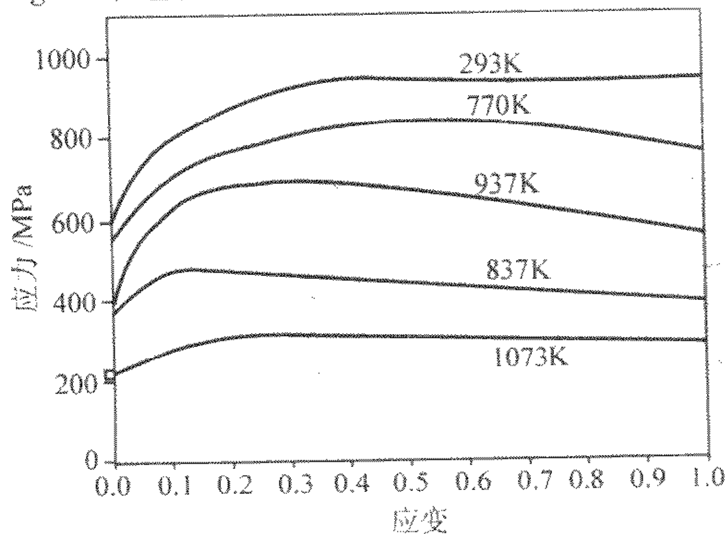


图 3.7-17　CF53 应力－应变曲线

与温度相关的物性指标见表 3.7-1；工件初始温度初选为 850℃，凹模与冲头温度分别选为 300℃和 250℃。工件热交换定义如下：与周围环境的热交换系数为 0.17；与模具的热交换系数为 20。塑性变形功转化为热能系数为 0.9。

摩擦模型为常切应力摩擦模型，取摩擦因子为 0.3；冲头速度为 25 mm/s；设备选为液压机。反挤成形过程如图 3.7-18 所示，成形过程等效应变如图 3.7-19 所示，温度分布情况见图 3.7-20。

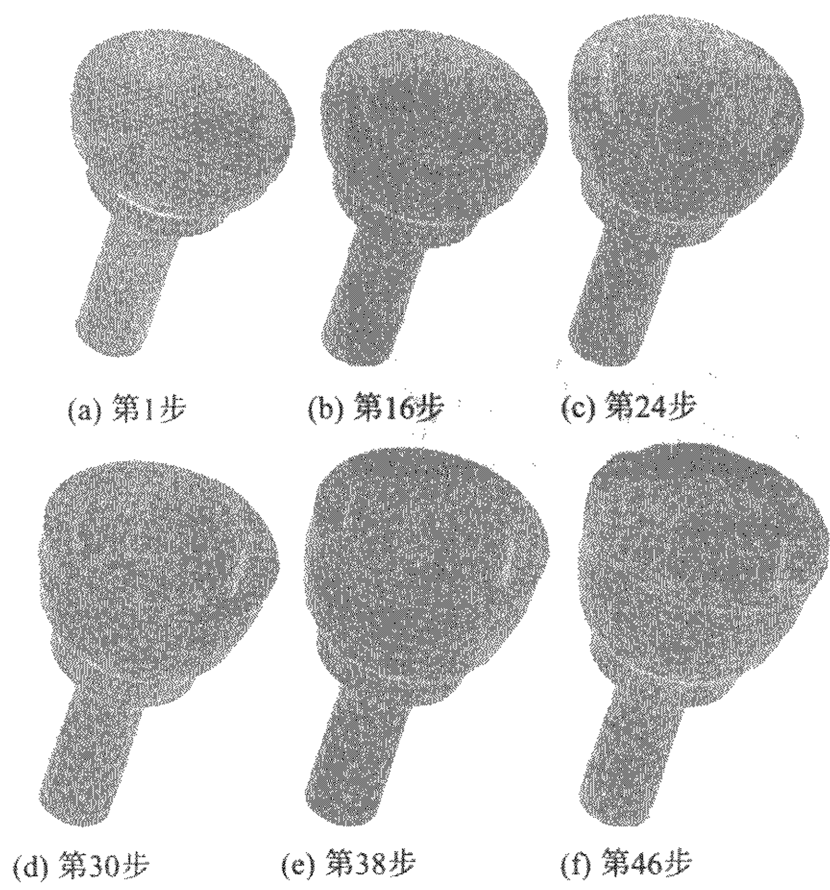
表 3.7-1　与温度相关的物性指标

温度/℃	20	100	200	300	400	500	600	1 500
弹性模量 E/GPa	213	207	199	192	184	175	164	69.44
热收缩率 $/10^{-6}\cdot K^{-1}$	11.9	12.5	13	13.6	14.1	14.5	14.9	14.9
热导率 λ $/W\cdot m^{-1}\cdot K^{-1}$	41.7	43.4	43.2	41.4	39.1	36.7	34.1	34.1
比热容 $/J\cdot kg^{-1}\cdot K^{-1}$	461	496	533	568	611	677	778	778

2.3　Marc/AutoForge 软件

2.3.1　简介

MARC 是功能齐全的高级非线性有限元软件，具有极强的结构分析能力。可以处理各种线性和非线性结构分析，包括线性/非线性静力分析、模态分析、简谐响应分析、频谱分析、随机振动分析、动力响应分析、自动的静/动力接触、屈曲/失稳、失效和破坏分析等。为满足工业界和学术界的各种需求，提供了层次丰富、适应性强、能够在多种硬件平台上运行的系列产品。MARC 软件按模块化编程，工作空间可根据计算机内存大小自动进行调整。用户如果对精度要求


(a) 第1步　(b) 第16步　(c) 第24步

(d) 第30步　(e) 第38步　(f) 第46步

图 3.7-18　反挤成形过程图

较高，可选用双精度进行运算。当单元数、节点数太多，内存不能满足需要时，程序能够自动利用硬盘空间进行分析。在分析过程中，利用网格自适应和重划分技术，能够变更单元的划分和节点数目。MARC 对于非线性问题采用增量解法，在各增量步内对非线性代数方程组进行迭代以满足收敛判定条件。根据具体分析的问题可采用不同的分析方法，如对于弹塑性分析和大位移分析可采用切线刚度法，对于蠕变分析或热应力分析可采用初应变法。对非结构的场问题如包含对流、辐射、相变潜热等复杂边界条件的非线性传热问题的温度场以及流场、电场、磁场，也提供了相应的分析求解能力；并具有模拟流－热－固、土壤渗流、声－结构、耦合电－磁、电－热、电－热－结构以及热－结构等多种耦合场的分析能力。

Marc/AutoForge 是采用 20 世纪 90 年代最先进有限元网格和求解技术，快速模拟各种冷热锻造、挤压、轧制以及多步锻造等体积成形过程的工艺制造专用软件。它综合了 MSC. Marc/MENTAT 通用分析软件求解器和前后处理器的精髓，以及全自动二维四边形网格和三维六面体网格自适应和重划分技术，实现对具有高度组合的非线性体积成形过程的全自动数值模拟。其图形界面采用工艺工程师的常用术语，容易理解，便于运用。Marc/AutoForge 提供了大量实用材料

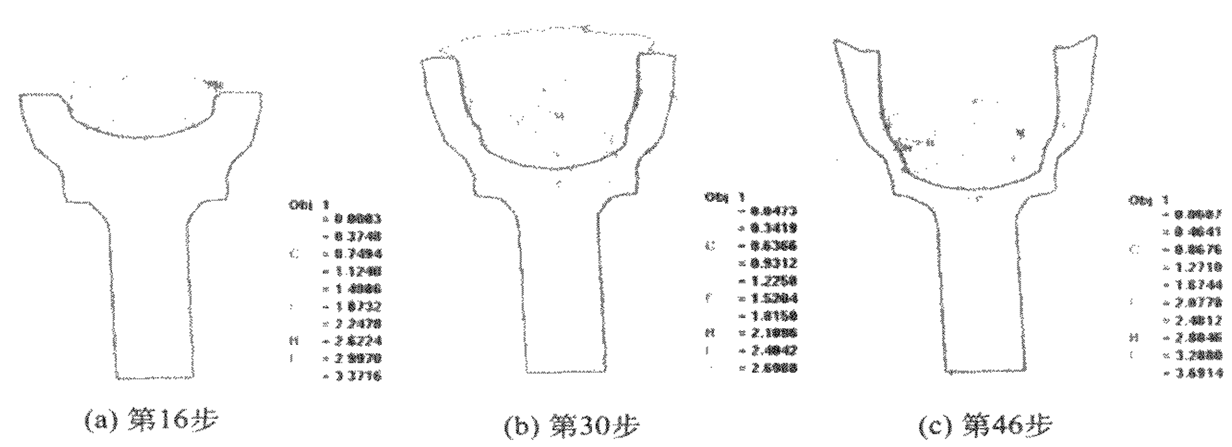

(a) 第16步　(b) 第30步　(c) 第46步

图 3.7-19　等效应变分布

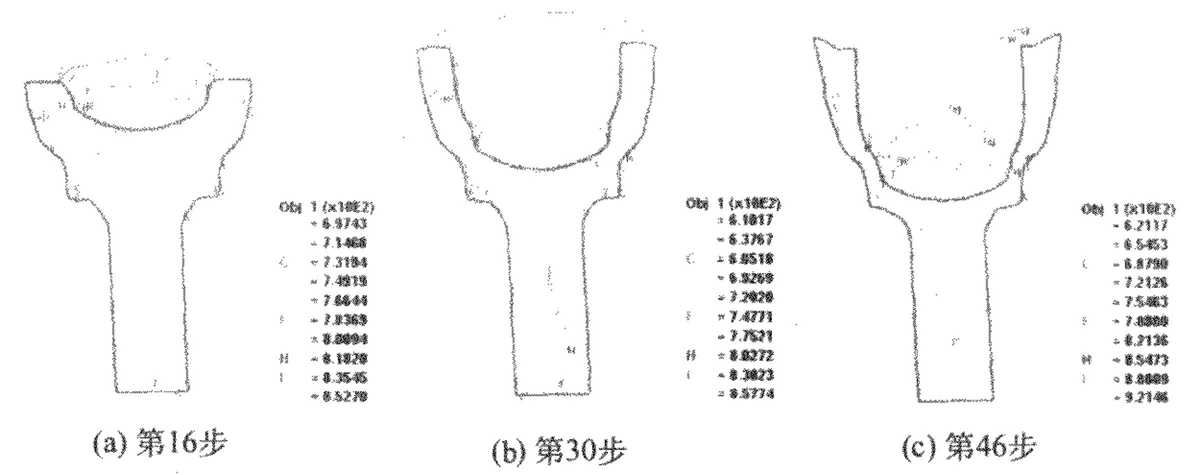

(a) 第16步　(b) 第30步　(c) 第46步

图 3.7-20　温度分布

数据以供选用，用户也能够自行创建材料数据库备用。Marc/AutoForge 除了可完成全 2D 或全 3D 的成形分析外，还可自动将 2D 分析与 3D 分析无缝连接，大大提高对先 2D 后 3D 的多步加工过程的分析效率。利用 Marc/AutoForge 提供的结构分析功能，可对加工后的包含残余应力的工件进行进一步的结构分析，模拟加工产品在后续的运行过程中的性能，有助于改进产品加工工艺或其未来的运行环境。此外，作为体积成形分析的专用软件，Marc/AutoForge 为满足特殊用户的二次开发需求，提供了友好的用户开发环境。

2.3.2　计算实例

同样以前述钟形罩预成形件反挤工序为例，模拟过程参数也同前，采用 Autoforge 进行数值分析的模型如图 3.7-21 所示。成形过程网格变化见图 3.7-22，成形工件的应力应变及温度分布如图 3.7-23 所示。

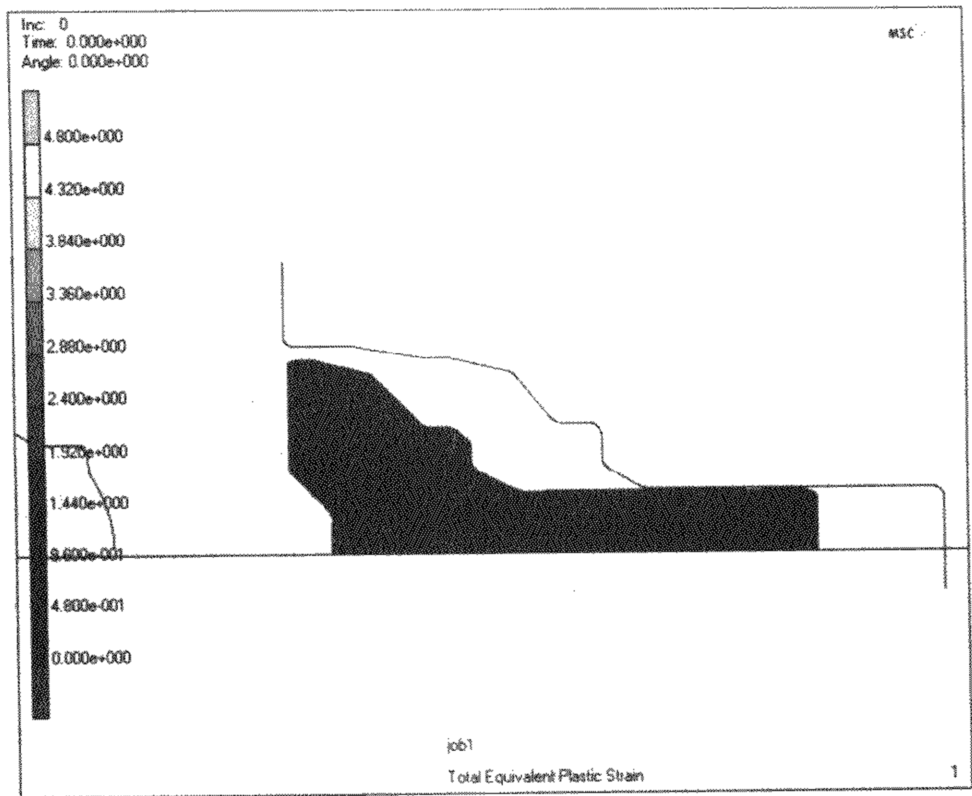
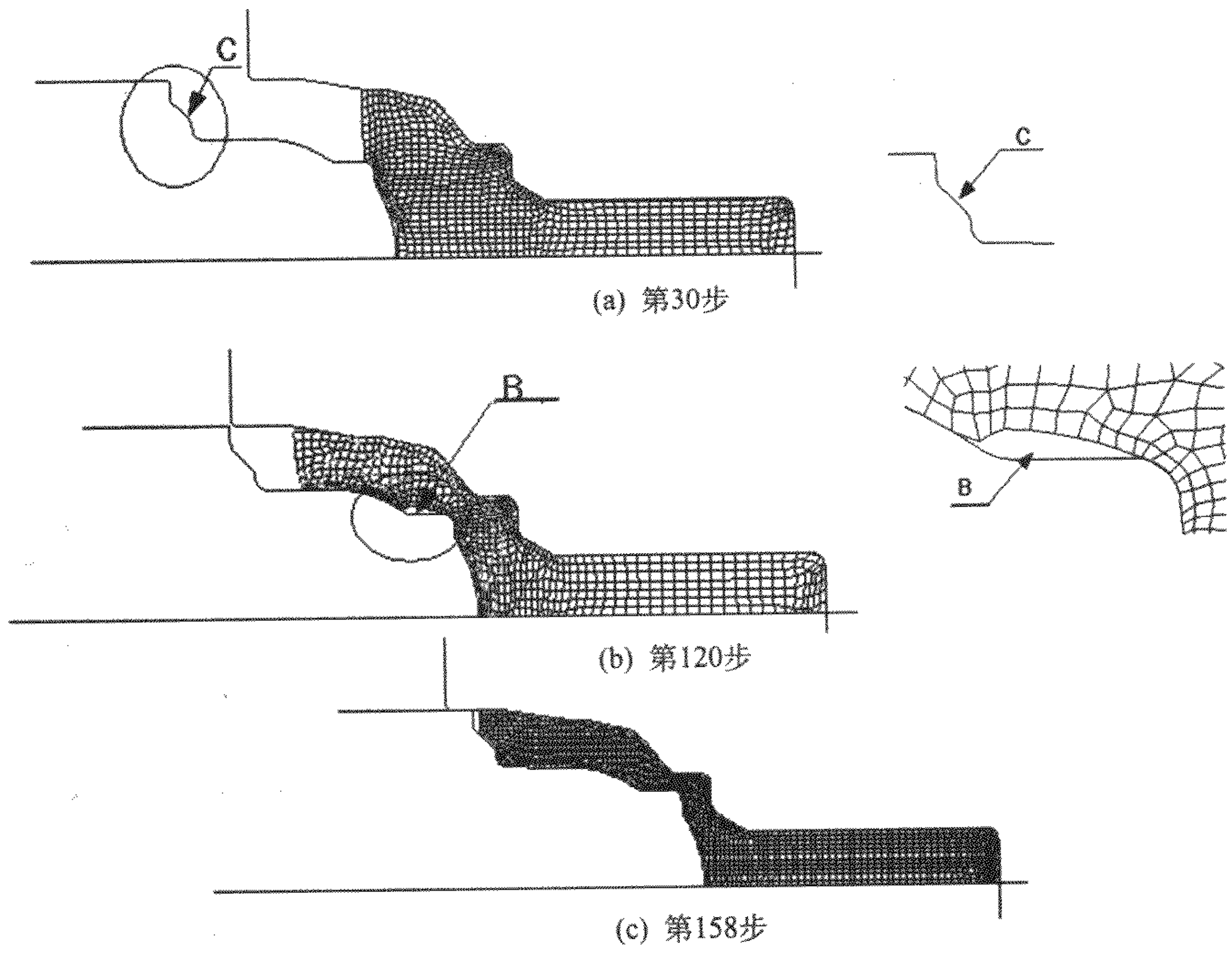


图 3.7-21　采用 Autoforge 进行反挤数值分析的模型

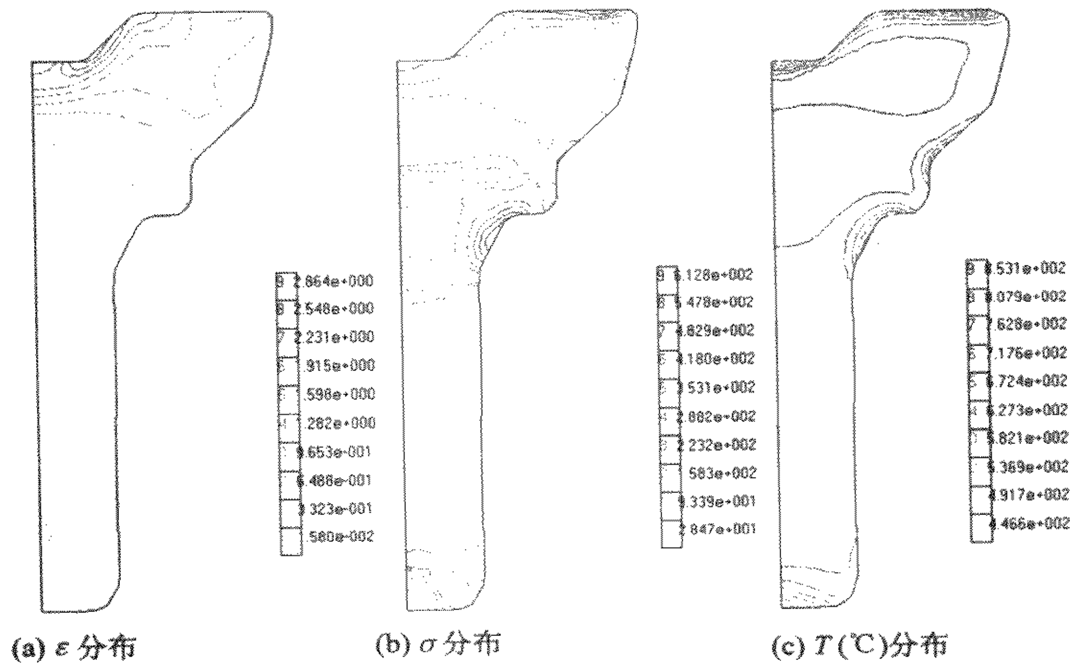


(a) 第30步

(b) 第120步

(c) 第158步

图 3.7-22　钟形罩反挤压模拟充填过程图



(a) ε 分布　(b) σ 分布　(c) T(℃)分布

图 3.7-23　杯部反挤压变形终了时刻场量分布

3　铸造成形模拟软件

经过四十多年的发展，目前铸造领域商品化数值模拟软件较多，国内外著名的软件系统有德国的MAGMASOFT、美国的PROCAST、英国的SOLSTAR、挪威的NOVACAST、日本的JSCAST、国内华中科技大学的“华铸CAE”等。这些软件在铸造领域得到了较为广泛的应用，取得了良好的应用效果。下面以国内铸造行业最著名的“华铸CAE”系统为例，介绍一下铸造成形模拟软件的特点及应用。

“华铸CAE铸造工艺分析系统”软件是分析和优化铸造工艺的重要工具，是华中科技大学经数十年研究开发，并在长期的生产实践中不断改进、完善起来的集成化软件系统。

它以铸件充型过程、凝固过程数值模拟技术为核心对铸件进行铸造工艺分析。可以完成多种合金材质（包括铸钢、球铁、灰铁、铸造铝合金等）、多种铸造方法（砂型铸造、金属型铸造、压铸、低压铸造、熔模铸造等）下铸件的凝固分析、流动分析以及流动和传热耦合计算分析。在预测铸件缩孔缩松缺陷的倾向，改进和优化工艺，提高产品质量，降低废品率，减少浇冒口消耗，提高工艺出品率，缩短产品试制周期，降低生产成本，减少工艺设计对经验对人员的依赖，保持工艺设计水平稳定等诸多方面都有明显的效果。曾在众多不同材质复杂铸件的工艺改进、工艺优化中圆满地完成增收降废的任务，创造了显著的经济效益和社会效益，博得众多生产厂家和同行专家的好评。系统总体性能居国内同类软件的领先地位，达到国际先进水平。

3.1　使用范围

材质包括：铸钢、灰铁、球铁、铸铝、铸铜及其他合金。

铸造方式：普通重力铸造、金属型、压力、低压、熔模等铸造方式。

3.2　分析内容

充型过程、流动与传热耦合过程、结晶凝固过程、应力应变形成过程。

3.3　缺陷预测

卷气、卷渣、冲砂、浇不足、冷隔、缩孔、缩松、裂纹、变形。

3.4　软件特点

1）自主版权，一次开发，贴近用户，贴近实际。

2）充分商品化，市场化，实用化。

3）铸件适应面宽，实用效果好，成果业绩丰硕。

4）界面傻瓜化，功能智能化，操作简捷轻松。

5）计算速度快，容量无限制，运行稳定可靠。

6）易学易用易掌握，学起来有滋有味，用起来爱不释手。

7）图形动画，透彻明了，直观动感，独领同类产品之最。

8）自学示范，向导帮助，服务周到，方便及时，舒适温馨。

3.5　软件功能

1）广泛的三维建模接口，从低档的AutoCAD，到中档的SolidEdge、SolidWorks，到高档的UG、Pro－E，等等，任凭用户选择搭配。

2）自动网格剖分，自动生成和维护迭代计算环境。

3）连续计算，间歇保护，自动恢复。

4）多铸造材质，多铸造方法，多工艺参数模拟分析。

5）充型模拟，耦合计算，确保精确的温度场模拟，辅助浇注系统优化设计。

6）温度计算，凝固模拟，为小到公斤级、大到百吨级铸件预测缺陷，优化工艺。

7）计算结果图形生成自动化，批处理化，动画服务菜单化。

8）为流动、温度、液相分布构造二维、三维二值图、色标图，为速度、温度梯度构造矢量分布图，为由用户任选的多个部位构造多点温度曲线。

9）结果显示，实时任意比例、任意旋转、任意剖切，显示可局部、可断面、可透视。

10）画面各颜色任意选用，任意搭配。

11）动画合成、动画分解、动画播放，完全菜单操作，功能自动完成。

12）充型过程、凝固过程动态演示，为缺陷预测，为工艺优化提供详尽的细节观察。

13）在线帮助、操作向导、自学教材，呼之即出，随时听用。

14）可以与流行的三维CAD、CAM、RPM（快速原型制造）进行集成。

图3.7-24～图3.7-29是该软件系统的三大模块的界面和一些图形表示。

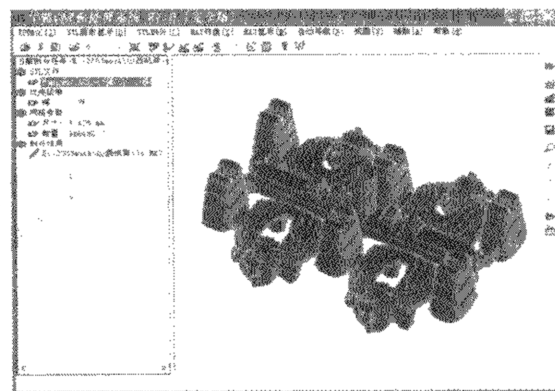


图3.7-24　华铸CAE系统前置处理界面

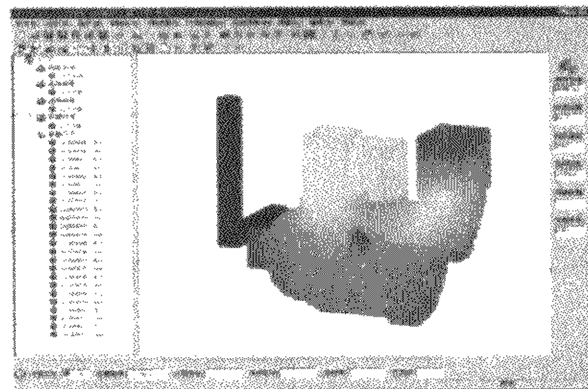


图3.7-25　华铸CAE系统计算部分界面

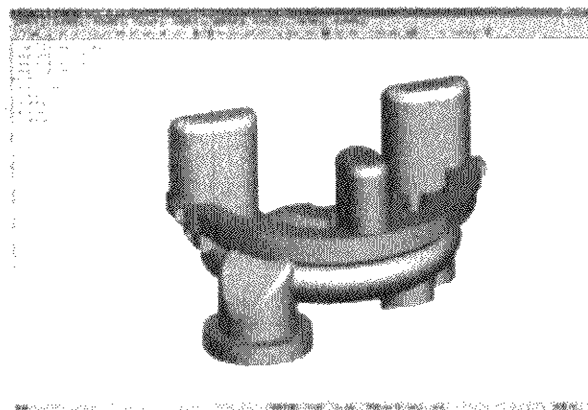


图3.7-26　华铸CAE系统后置处理界面

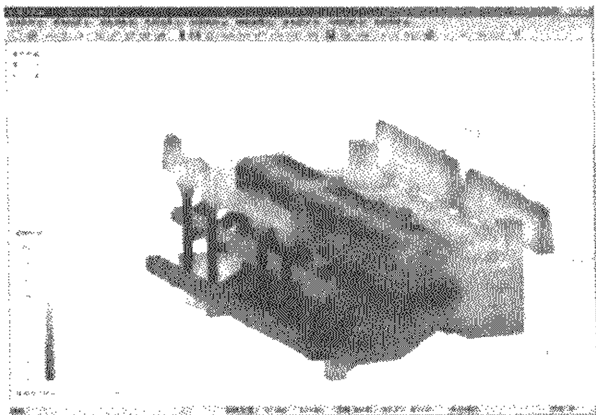


图 3.7-27 六缸柴油机机体色温分布图

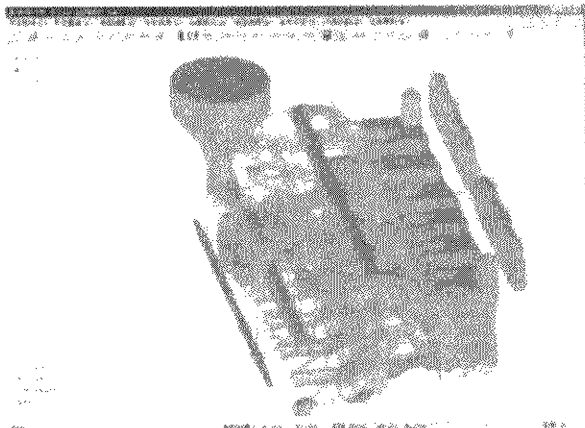


图 3.7-28 六缸柴油机机体凝固液相分布图

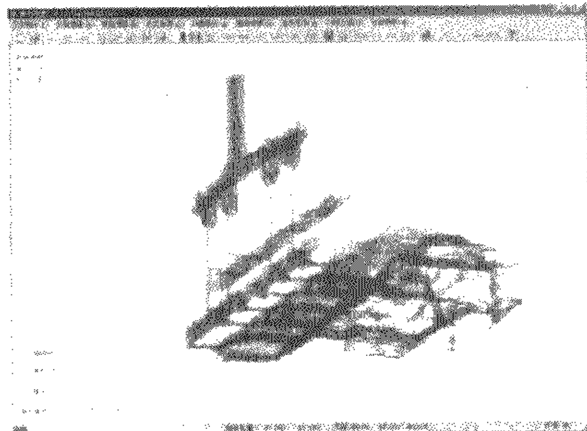


图 3.7-29 六缸机体充型流速图

3.6 华铸 CAE 应用实例

华铸 CAE 软件在生产应用中取得了巨大的成功，适用于铸钢、球铁、灰铁、铸铝各类铸件，大到一、二百吨，小到几公斤，无论是解决缩孔、缩松，还是优化浇冒口结构，提高工艺出品率，改进浮渣夹渣等等方面，都有非常成功的实例。

1）汽车后桥壳缩孔问题的改进　图 3.7-30 是国内某厂生产的球铁汽车后桥壳原生产工艺剖视图，经“华铸 CAE”软件模拟分析，发现有两处在凝固过程中出现较大断面的孤立液相，无法得到充分的收缩补充，因此会存在严重的缩孔缩松倾向，如图 3.7-31 所示。经实物解剖验证，证实确有较大孔洞存在。

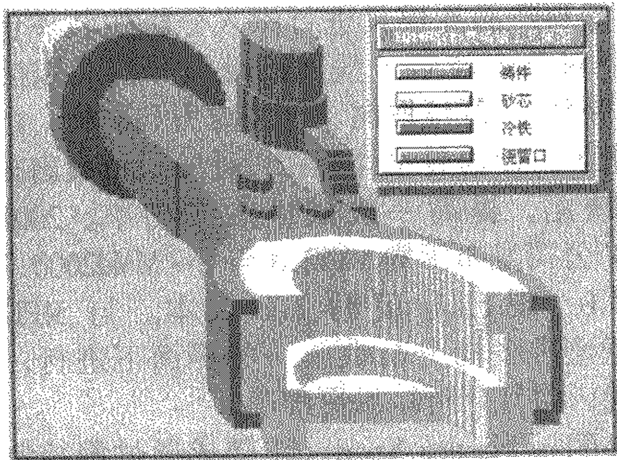


图 3.7-30 汽车后桥壳原工艺剖视图

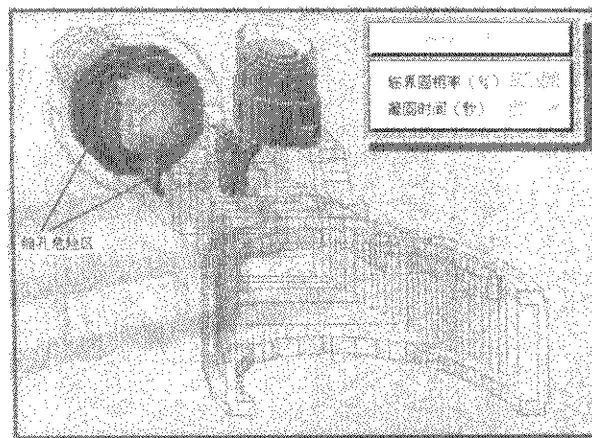


图 3.7-31 模拟显示的缩孔危险区

改进工艺，首先在端部突缘处加一环形冷铁，通过软件进行模拟，确认该冷铁可以消除端部突缘内的缩孔，但颈部侧面的缩松倾向仍较明显，再进行二次改进，在侧面再加一冷铁，如图 3.7-32。经模拟，此方案较好地解决了两处出现孤立液相的问题，见图 3.7-33。将此方案交付生产，并解剖所生产的铸件，证明缺陷确已消除，解决了长期隐藏的缩孔缩松质量隐患。

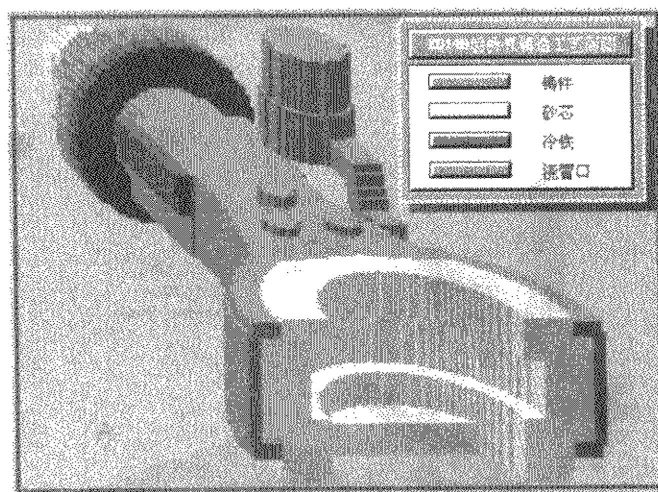


图 3.7-32 二次改进方案

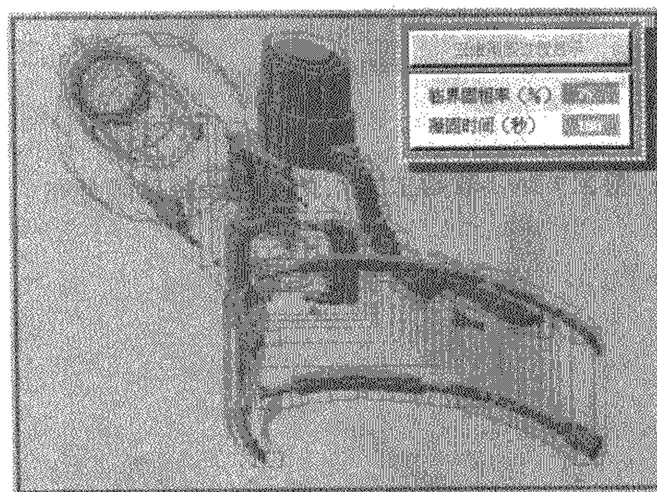


图 3.7-33 改进方案模拟结果

附带说明一点，由于球铁在凝固过程中有一个石墨化膨胀过程，在铸型刚度较好时，一些小断面的孤立液相会因此而得到所谓“自补缩”。工艺上就不必为这些分布很广的小断面孤立液相做许多琐碎的工作。实际上，相对于铸钢件，球铁件必然形状更复杂，更容易存在这种小断面的孤立液相。而有限的自补缩能力又恰好能够自我消化掉这些毛病。另外，始终与冒口连通的液相显然也不会引起缩孔缩松的危险，见图 3.7-33。

2）柴油机机体夹渣问题的改进　图 3.7-34 是某厂生产康明斯柴油机灰铸铁机体的工艺简图，该铸件形状复杂，为表达清晰起见，特将浇注系统单独绘出。其横浇道有上下两层，上层横浇道位于机体高度中心处，下层位于高度底部，见图 3.7-34。图 3.7-35 为柴油机机体装配图（左）与色温图（右）。

利用华铸 CAE 系统对上述缸体的工厂试制工艺方案进行了充型模拟分析，并对缺陷进行了预测。计算分析发现，该工艺方案充型的前期及后期都比较顺畅、平稳，但在中间

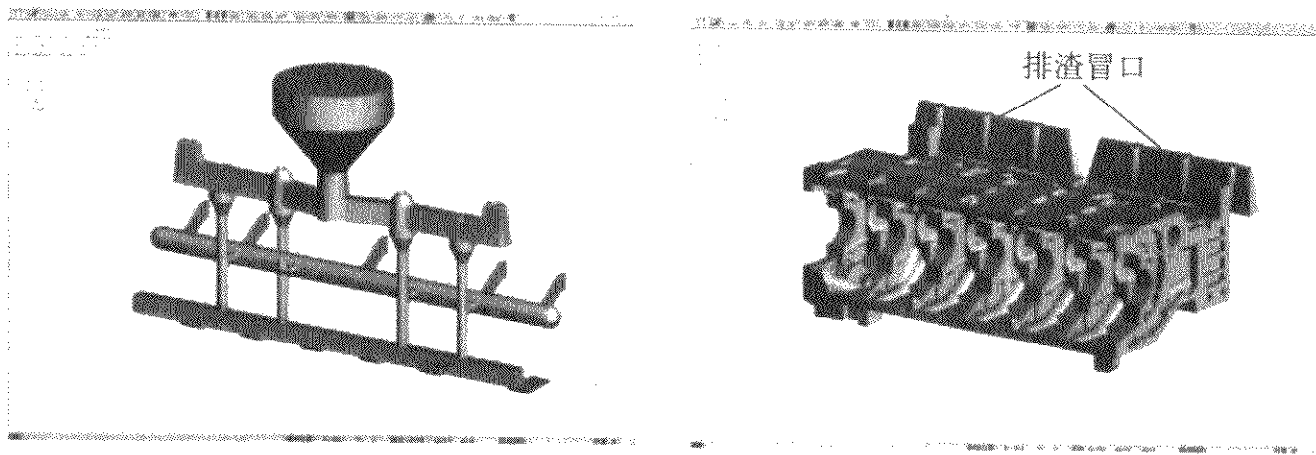


图 3.7-34　柴油机机体（右）及浇道（左）

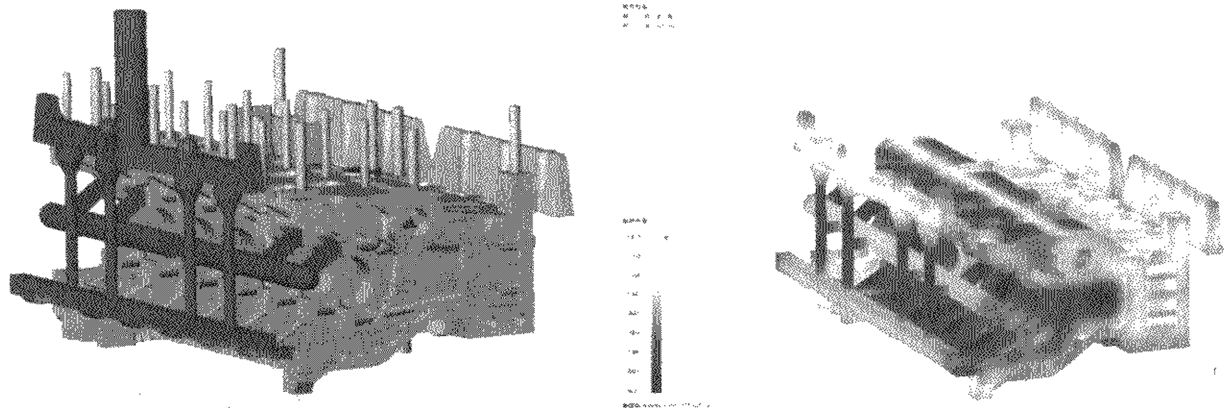


图 3.7-35　柴油机机体装配图（左）与色温图（右）

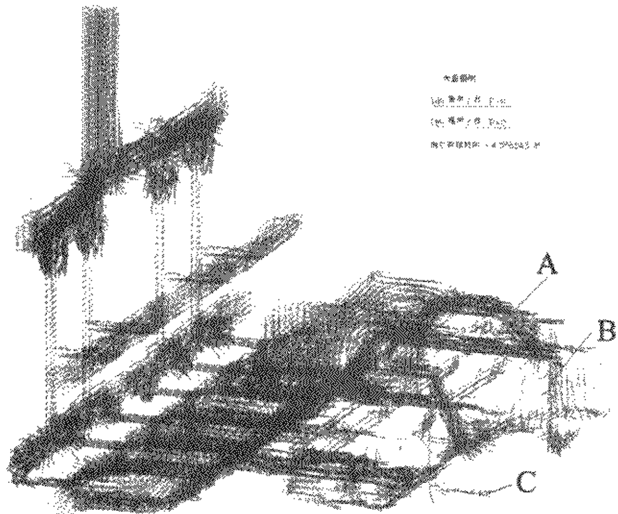


图 3.7-36　原始工艺 4.55 s 时的速度场分布

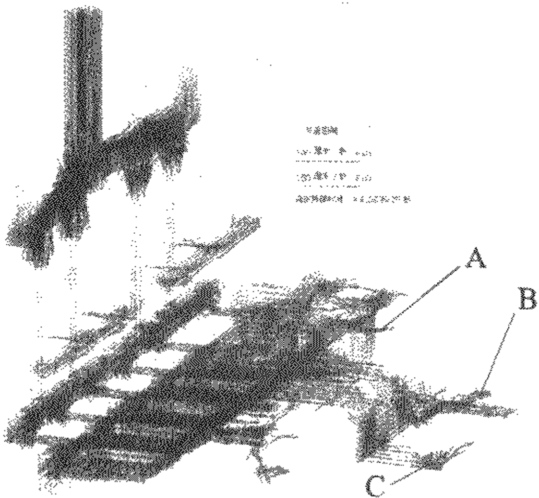


图 3.7-37　改进工艺 4.10 s 时的速度场分布

一阶段（大约为 3.8～6.0 s）充型顺序不好，出现明显的紊流。图 3.7-36 所示为原工艺方案在 4.55 s 时的速度场分布，可以看出金属液是从高处（A 点）向低处（B、C 点）流动，类似瀑布一样，在 B 点出现了明显的负压带。也就是说该工艺方案在充型过程中 B、C 两处会有较多的气体和渣搅入。而通过流动与传热的耦合模拟计算得知，此时 B、C 两处的温度较低（流动前沿），搅入的气和渣难以及时上浮，易造成卷气、夹渣缺陷。

为此，对上述原始工艺进行了改进，在相关部位增加导流槽，试图改善中间阶段的充型状况。图 3.7-37 即为改进方案在 4.1 s 时的速度场分布，可以看出，液态金属，是从底部向上充填，即先充到 C 点，再为 B 点、A 点，上述各点没有明显出现负压带，充型顺序得到显著改善，工艺得到了优化。

在上述模拟工作之前，某厂采用原始工艺方案生产了几十件，经解剖发现几乎每一个铸件在 B、C 两个位置都有卷气、夹渣缺陷发生，废品率甚高。后采用经模拟优化的改进工艺方案后，B、C 两处卷气、夹渣缺陷得以解决，废品率下降了一倍。这说明这一改进是切中要害的，也说明 CAE 软件的流动分析功能对于改进流动方式，克服夹渣缺陷是非常有效。

4　塑料注射成形模拟软件

4.1　常用软件简介

4.1.1　MoldFlow 软件

MoldFlow 软件是美国 MOLDFLOW 公司的产品，该公司自 1976 年发行了世界上第一套塑料注射成形流动分析软件以来，一直主导塑料成形 CAE 软件市场。2000 年 4 月，MOLDFLOW 公司收购了另一个世界著名的塑料成形分析软件 C－MOLD。

MoldFlow 软件主要包括如下三部分。

1）MoldFlow Plastics Advisers（产品优化顾问，简称 MPA）　塑料产品设计师在设计完产品后，运用 MPA 软件模拟分析，在很短的时间内，就可以得到优化的产品设计方案，并确认产品表面质量。

2）MoldFlow Plastics Insight（注塑成形模拟分析，简称 MPI）　对塑料产品和模具进行深入分析的软件包，它可以在计算机上对整个注塑过程进行模拟分析，包括填充、保压、冷却、翘曲、纤维取向、结构应力和收缩，以及气体辅助成形分析等，使模具设计师在设计阶段就找出未来产品可能出现的缺陷，提高一次试模的成功率。2001 年年底，MOLDFLOW 公司发布了集 MPI2.0 与 C－Mold2000.7 功能于一体的 MPI3.0，现已升级到 3.1 版本。与 MPI2.0 相比，MPI3.1 的模型准备与分析在统一的界面下进行，大大简化了操作步骤，而且分析结果更准确。

3）MoldFlow Plastics Xpert（注塑成形过程控制专家，简称 MPX）　集软硬件为一体的注塑成形品质控制专家，可以直接与注塑机控制器相连，可进行工艺优化和质量监控，自

动优化注塑周期、降低废品率及监控整个生产过程。

MPI是MoldFlow软件的主打产品，所包括的主要模块如下。

1）模型输入与修复 MPI3.1有三种分析方法：基于中心面的分析、基于表面的分析与三维分析。中心面既可运用MPI软件的造型功能完成，也可从其他CAD模型中抽取，再编辑；表面分析模型与三维分析模型直接读取其他CAD模型，如快速成形格式（STL）、IGES、STEP、Pro/E模型、UG模型等。模型输入后，软件提供了多种修复工具，以生成既能得到准确结果，又能减少分析时间的网格。

2）塑料材料与注塑机数据库 材料数据库包含了超过4 000种塑料材料的详细数据，注塑机数据库包含了290种商用注塑机的运行参数，而且这两个数据库对用户是完全开放的。

3）流动分析 分析塑料在模具中的流动，并且优化模腔的布局、材料的选择、填充和保压的工艺参数。

4）冷却分析 分析冷却系统对流动过程的影响，优化冷却管道的布局和工作条件，与流动分析相结合，可以得到完美的动态注塑过程。

5）翘曲分析 分析整个塑件的翘曲变形，包括线性变形和非线性变形，同时指出产生翘曲的主要原因以及相应的改进措施。

6）纤维填充取向分析 塑件纤维取向对采用纤维化塑料的塑件性能（如拉伸强度）有重要影响。MPI软件使用一系列集成的分析工具来优化和预测整个注塑过程的纤维取向，使其分布合理，从而有效地提高该类塑件的性能。

7）优化注塑工艺参数 根据给定的模具、注塑机、塑件材料等参数以及流动分析结果自动产生控制注塑机的填充保压曲线，从而免除了在试模时对注塑机参数的反复调试。

8）结构应力分析 分析塑件在受外界载荷情况下的力学性能，在考虑注塑工艺的条件下，优化塑件的强度和刚度。

9）确定合理的塑料收缩率 MPI3.1通过流动分析结果确定合理的塑料收缩率，保证模腔的尺寸在允许的公差范围内，从而减少塑件废品率，提高产品质量。

10）气体辅助成形分析 模拟气体辅助注射成形过程，对整个成形过程进行优化。

11）特殊注塑成形过程分析 MPI3.1可以模拟共注射、反应注射、微芯片封装等特殊的注射成形过程，并对其进行优化。

4.1.2 HSCAE软件

华中科技大学模具技术国家重点实验室是国内最早从事注塑成形CAE研究开发的单位，并于1988年推出了国内第一个注塑成形CAE软件HSCAE1.0，并在随后的10年间相继推出了多个升级版本。尤其是在“九五”期间，由传统的基于制品中心层模型的二维半分析软件发展了基于三维实体模型的三维模拟软件，并于1998年推出了新一代的基于制品实体模型的注塑成形模拟系统HSCAE5.0，突破了注塑成形CAE推广应用的“建模瓶颈”。目前该软件已在70多家公司、企业和学校推广应用，装机达400余套。

HSCAE 6.1是HSCAE软件的最新版本，该系统具有如下特色。

1）支持通用的STL数据交换文件，使得用目前国内外流行的造型软件（如Pro/E、UG、Solid Edge、InteSolid、金银花MDA等）所生成的制品模型通过STL均可以输入并转换到华塑HS CAE系统中，直接进行分析显示。

2）支持国产塑料材料数据库，分析结果准确可靠，同时还可以为用户提供塑料的流变参数测定，并将数据添加到软件的材料数据库中。

3）将人工智能技术与数值模拟技术紧密结合，在以等值线、矢量、阴影、动画、多模型、多窗口、图表、表格等方式充分表现分析结果的基础上，通过人工智能技术自动为用户提供较优的初始参数，由专家系统来处理CAE分析产生的大量数据，依据专家知识进行推理，将工程化的分析结论提供给用户，并为用户提供参考的改进方案，将仿真软件由传统的“被动式”计算工具提升为“主动式”的优化系统。

HSACE软件在注塑模设计中的作用可以体现在以下几方面。

1）优化塑料制品 运用模拟软件，可以得到制品所需的最小壁厚，优化制品结构，降低材料成本，缩短生产周期，保证制品能全部充满。

2）优化模具结构 运用模拟软件，可以得到最佳的浇口数量与位置，合理的流道系统与冷却系统，并对型腔尺寸、浇口尺寸、流道尺寸和冷却系统尺寸进行优化，在计算机上进行试模、修模，大大提高模具质量，减少修模次数。

3）优化注塑工艺参数 运用模拟软件，可以确定最佳的注射压力、保压压力、锁模力、模具温度、熔体温度、注射时间、保压时间和冷却时间，以注塑出最佳的塑料制品。

4.2 注射成形CAE分析结果的指导意义

注射成形CAE技术在当今的塑料工业领域得到了广泛应用。工程师们借助CAE技术可以全面地把握塑料注射成形过程，寻求一种增加产量、提高质量、节省时间和费用的最佳方案。CAE分析可以输出重要的设计数据，如压力分布、温度、剪切速率、剪切应力、速度等，设计者可由CAE获取诸如充填模式、熔接缝与气穴的位置、注射压力和锁模力大小、纤维取向、冷却时间、最终成形情况等信息。作为一种设计工具，CAE能够辅助模具设计师优化流道系统与模具结构，协助产品设计师从工艺的角度改进产品形状、选择最佳成形性能的塑料，帮助模具制造者选择合适的注射机，指导模塑工程师设置合理的工艺条件。使用CAE分析可以对不同的成形方案进行反复的评测对比，寻求最优设计。同时，CAE软件又是一种教学工具，通过对注射成形过程各阶段的定性与定量描述，CAE能够帮助设计者熟悉熔体在型腔内的流动行为，把握熔体流动、保压、凝固的基本原则，帮助设计新手克服经验的不足，帮助有经验的工程师注意那些也许会被忽视的细节。

虽然CAE技术具有节省时间和原材料、降低废品率、提高产品质量、缩短产品开发周期等优势，但它并不是解决所有成形问题的万能方法。实际上CAE只是一种分析工具，用来帮助或完善工程师的设计而不是取代他们。因此，和其他工具一样，CAE技术作用的大小取决于使用者的经验水平和熟练程度。确切地讲，CAE分析结果的精度在很大程度上取决于设计人员提供的输入数据，以及对输出结果的正确判断与解释。

下面以注射成形充模流动模拟为例介绍分析软件的指导作用。注射模流动模拟软件的指导意义十分广泛，它是一种设计工具，能够辅助模具设计者优化模具结构与工艺，指导产品设计者从工艺的角度改进产品形状，选择最佳成形性能的塑料，帮助模具制造者选择合适的注射机，当变更塑料品种时对现有模具的可行性做出判断，分析现有模具设计弊病。同时，流动软件又是一种教学软件工具，能够帮助模具工作者熟悉熔体在型腔内的流动行为，把握熔体流动的基本原则。下面逐项分析三维流动软件的主要输出结果是如何用来指导设计的。

1）熔体流动前沿动态显示 三维流动模拟软件能显示熔体从进料口逐渐充满型腔的动态过程，由此可判断熔体的

流动是否较理想的单向流形式（简单流动）（复杂流动成形不稳定，容易出现次品），各个流动分支是否同时充满型腔的各个角落（流动是否平衡）。若熔体的填充过程不理想，可以改变进料口的尺寸、数量和位置，反复运行流动模拟软件，一直到获得理想的流动形式为止。若仅仅是为了获得较好的流动形式而暂不考察详尽的温度场、应力场的变化，或是初调流道系统，最好是运行简易三维流动分析（等温流动分析），经过几次修改，得到较为满意的流道设计后，再运行非等温三维流动分析。

2）型腔压力　在填充过程中最大的型腔压力值能帮助判断在指定的注射机上熔体能否顺利充满型腔（是否短射），何处最可能产生飞边，在各个流动方向上单位长度的压力差（又称压力梯度）是否接近相等（因为最有效的流动形式是沿着每个流动分支熔体的压力梯度相等），是否存在局部过压（容易引起翘曲）。流动模拟软件还能给出熔体填充模具所需的最大锁模力，以便用户选择注射机。

3）熔体温度　流动模拟软件提供型腔内熔体在填充过程中的温度场。可鉴别在填充过程中熔体是否存在着因剪切发热而形成的局部热点（易产生表面黑点、条纹等并引起力学性能下降），判断熔体的温度分布是否均匀（温差太大是引起翘曲的主要原因），判断熔体的平均温度是否太低（引起注射压力增大）。熔体接合点的温度还可帮助判断熔合纹的相对强度。

4）剪切速率　剪切速率又称应变速率或者速度梯度。该值对熔体的流动过程影响甚大。实验表明，熔体在剪切速率为$10^3\ s^{-1}$左右成形，制品的质量最佳。流道处熔体剪切速率的推荐值约为$5\times10^2\sim5\times10^3 s^{-1}$，浇口处熔体剪切速率的推荐值约为$10^4\sim10^5\ s^{-1}$。流动软件能给出不同填充时刻型腔各处的熔体剪切速率，这就有助于用户判断在该设计方案下预测的剪切速率是否与推荐值接近，而且还能判断熔体的最大剪切速率是否超过该材料所允许的极限值。剪切速率过大将使熔体过热，导致聚合物降解或产生熔体破裂等弊病。剪切速率分布不均匀会使熔体各处分子产生不同程度的取向，因而收缩不同，导致制品翘曲。通过调整注射时间可以改变剪切速率。

5）切应力　切应力也是影响制品质量的一个重要因素，制品的残余应力值与熔体的切应力值有一定的对应关系，一般，切应力值大，残余应力值也大。因此总希望熔体的切应力值不要过大，以避免制品翘曲或开裂。根据经验，熔体在填充型腔时所承受的切应力不应超过该材料抗拉强度的1%。

6）熔合纹/气穴　两个流动前沿相遇时形成熔合纹，因而，在多浇口方案中熔合纹不可避免，在单浇口时，由于制品的几何形状以及熔体的流动情况，也会形成熔合纹。熔合纹不仅影响外观，而且为应力集中区，材料结构性能也受到削弱。改变流动条件（如浇口的数目与位置等）可以控制熔合纹的位置，使其处于制品低感光区和应力不敏感区（非“关键”部位）。而气穴为熔体流动推动空气最后聚集的部位，如果该部位排气不畅，就会引起局部过热、气泡、甚至充填不足等缺陷，此时就应该加设排气装置。流动模拟软件可以为用户准确地预测熔合纹和气穴的位置。

7）多浇口的平衡　当采用多浇口时，来自不同浇口的熔体相互汇合，可能造成流动的停滞和转向（潜流效应），这时各浇口的充填不平衡，影响制品的表面质量及结构的完整性，也得不到理想的简单流动。这种情况应调整浇口的位置。

注射模流动模拟软件只是一种辅助工具，它能否在生产中发挥作用并产生经济效益，在很大程度上取决于模具设计者的正确使用。流动软件的使用者必须熟悉注射成形工艺，具有一定的注射模设计经验。这样，用户才能针对性地利用流动软件解决模具结构设计或工艺问题。例如，如果浇口处剪切速率过高，是修正浇口尺寸，还是改变熔体温度，抑或更换注射材料，不具备注射成形工艺知识的人很难做出正确选择的。流动软件的输出结果涉及到塑料黏度、剪切速率、温度、压力以及它们的相互作用，即使是经验丰富的模具设计师也应学一点塑料流变学的知识，总结注射流动的基本规律，这样才能站在理论与实践结合的高度用好流动模拟软件。

4.3　应用实例

4.3.1　彩电前壳实例分析与讨论

图3.7-38所示为某公司所开发的彩色电视机前壳，其模具结构比较复杂，进行CAE分析的主要目的是确定模具最佳的浇注系统设计方案，以获得最为理想的熔料流动充填情况。所采用的模拟软件是华中科技大学模具技术国家重点实验室所开发的HSCAE。本制品的材料为ABS POLYLAC PA－707，对该种材料，先用流变实验测定其流变数据，再对流变数据进行处理与拟合，获得该材料的黏度参数：

$$B=2.868\,7E-06\ Pa*s,\quad T_b=1.069\,8E+0.4\ K$$

$$\tau^*=5.506\,4\ E+0.4\ Pa,\quad n=0.253\,2$$

考虑到制品比较复杂，初始方案包括四个主分流道，其中下面三个主分流道又各自分成两个分流道，各对应两个浇口（如图3.7-38所示）。充模时间设为7 s，注射温度为250℃。用所开发的软件分析后发现充填不平衡，制品的上部与两侧比下部明显地先充填，存在过压现象，同时制品充填结束时的温差大（133℃）。主要原因是浇口位置不合理，充模时间过长。

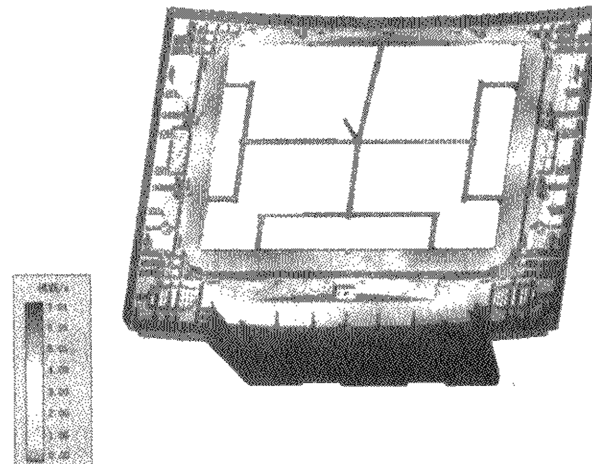


图3.7-38　初始方案的流动充填结果

考虑到初始方案分析得到的注射压力不是很高（55 MPa），因此将流道系统的设计改为如图3.7-39所示，两侧的浇口由两个减为一个，并将充模时间缩短为2.8 s。再次用CAE分析后发现，充填的平衡性变好，制品的各浇口基本上同时充填结束，同时减少了主要熔接缝的数目，注射压力变化不大（57 MPa），充填结束时的温差明显缩小（约71℃）。在注射时间显著减少的情况下，而所需注射压力并没有太大变化。

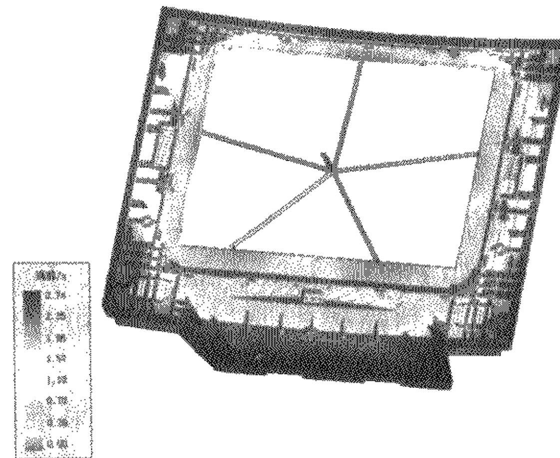


图3.7-39　改进方案的流动充填结果

现代注射机一般都提供了设置分级注射的功能，根据熔体流动前沿速度应保持恒定的原则（有利于制品性能的均匀一致），软件自动计算流动前沿面积（Melt Front Area，MFA）随制品已充填体积的变化，推荐如图 3.7-40 所示的分级注射曲线。因此在上述方案的基础上再采用分级注射，CAE 分析发现所需的注射压力进一步降低（48.9 MPa），这主要是因为分级注射中的注射速率在注射的后阶段有所下降。

图 3.7-41 为软件分析所需锁模力的曲线图，选择型号为 Vistal 1500 - B125 的注射机，其最大锁模力为 1 336 × 10^4 N，最大注射压力为 151 MPa，显然可以胜任该制品的成形加工，图 3.7-42 为最终方案充填结束时模拟的压力分布图。

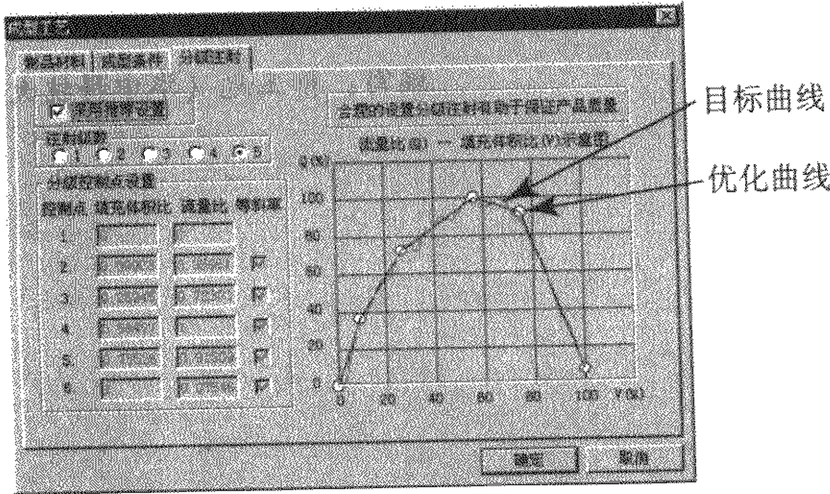


图 3.7-40　软件推荐的分级注射方案

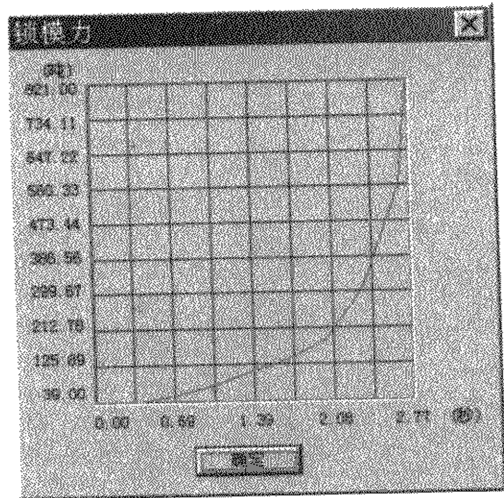


图 3.7-41　锁模力曲线

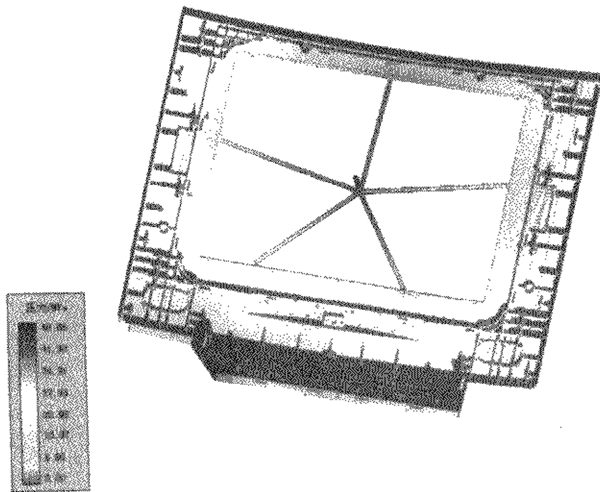


图 3.7-42　充填结束时的压力分布图

4.3.2　彩电后壳实例分析

与上述彩电前壳所配套的彩电后壳公称尺寸大、流动路径长，应用 CAE 分析的主要目的是在开模之前获得最佳的模具流道系统设计方案、合理的工艺条件，并确定制品熔接缝和气穴的位置。通过与结构设计师、工艺师协商，共设计了四个候选方案，其基本结构及工艺设计如表 3.7-2 所示。四个候选方案对应的流道结构及流动充填结果分别如图 3.7-43～图 3.7-45 所示。通过表 3.7-2 的对比分析，综合考虑可成形性、模具结构刚度强度、模具加工等因素，可以看出方案 3 比较合理，遂采用此方案试模，取得成功。图 3.7-46 为软件预测的制品熔接缝与气穴分布图（线条为熔接缝、小球为气穴位置）。

表 3.7-2　多方案的 CAE 模拟比较

方案	流道结构	注射时间 /s	注射温度 /℃	所需注射压力 /MPa	优缺点
1	一个主流道型浇口，见图 3.7-43	4.0	250	55.7	底板部分最后充填，流动不平衡
2	两条分流道，两点进胶，见图 3.7-44	4.0	250	60.8	流动基本平衡，但要在底板部分设置一浇口实现难度较大
3	三条分流道，三点进胶，见图 3.7-45	4.0	250	58.2	流动较平衡，流道系统易于实现
4	四条分流道，对应四点进胶，见图 3.7-46	4.0	250	49.8	流动平衡，压力较小，但很长的分流道将紧贴模壁，模具刚度、强度难以保证

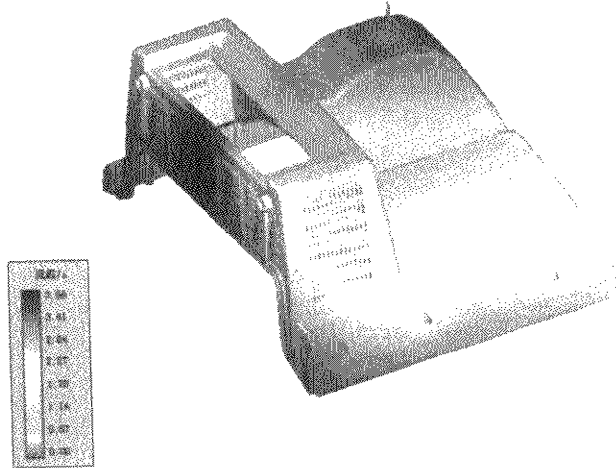


图 3.7-43　方案 1 的流道结构及充填结果

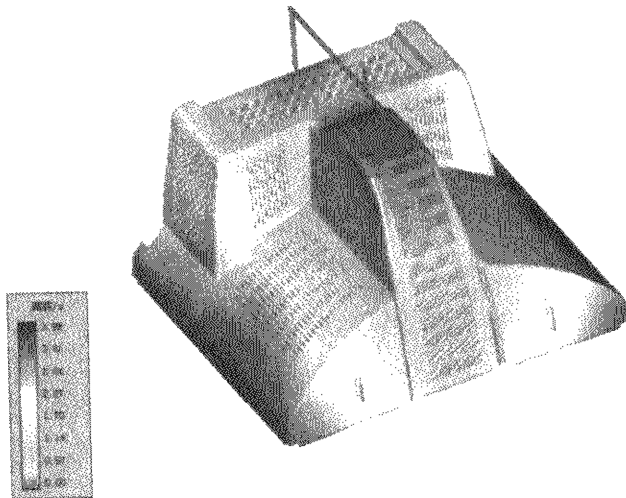


图 3.7-44　方案 2 的流道结构及充填结果

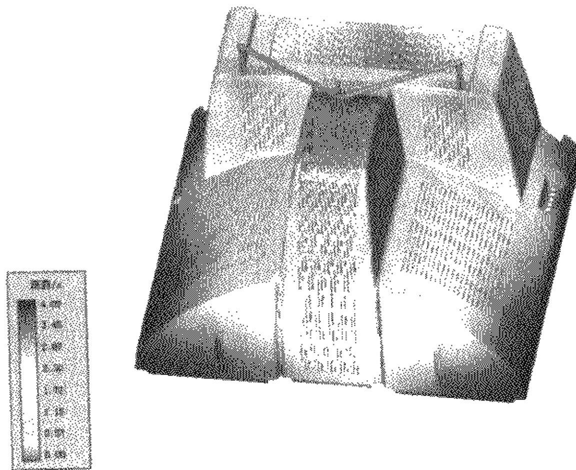


图 3.7-45　方案 3 的流道结构及充填结果

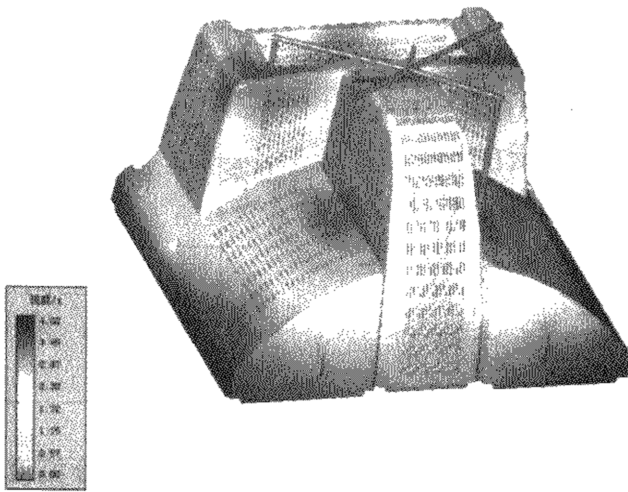


图 3.7-46 方案 4 的流道结构及充填结果

编写：柳玉起（华中科技大学）

参考文献

1 柳玉起，胡平，刘军化，陈塑寰．板材冲压成形变形局部化与动边界摩擦约束．固体力学学报，1997，（18）：204～212

2 王勖成，邵敏．有限单元法基本原理和数值方法．北京：清华大学出版社，1997

3 殷家驹，Combescure A．一般壳体结构大位移大应变弹塑性有限元分析与损伤力学的某些应用．应用力学学报，1990，7（1）：10～20

4 柳玉起，郭威，胡平．板料成形离散 Kirchhoff 理论单元模型．塑性工程学报，1995，（2）：4

5 Liu Yuqi，Hu ping，Liu Junhua．Flange earring and its control on deep-drawing of anisotropy circular sheets．ACTA MECHANICA SOLIDA SINICA，1999，（12）：4

6 Y．Q．Liu，J．C．Wang，P．Hu．The numerical analysis of anisotropic sheet metals in deep-drawing processes．J．Mates．Proc．Tech．，2002，（120）：1～3

7 Liu Yuqi，Wang Jingcheng，Hu ping．A finite element analysis of the flange earrings of strong anisotropic sheet metals in deep-drawing processes．ACTA MECHANICA SINICA．2002（18）：1

8 Liu Yuqi，Hu Ping，Wang Jingcheng．Springback Simulatron and Analysis of Strong Anisotropy Sheet Metals in U-channel Bending Process．ACTA MECHANICA SINICA，2002（18）：3

9 Liu Yuqi，Liu junhua，Hu Ping，Li Yunxing．Quantitative prediction for springback of unloading and trimming in sheet metal stamping forming．Chinese Journal of Mechanical Engineering．2003，（16）：2

10 李尚健主编．金属塑性成形过程模拟．北京：机械工业出版社，1999

11 中国机械工程学会，中国模具设计大典编委会．中国模具设计大典．第 1 卷．现代模具设计基础．南昌：江西科学技术出版社，2003

12 吴伯杰，陈森烂，王少林．网格再划分技术在金属成形过程有限元模拟中的应用．锻压技术，1992（2）：7～10

13 Wang Xinyun，Wang Zhongjin，Z．R．Wang．Numerical simulation of aluminum alloy ladder parts．Transaction of Nonferrous Metals Society of China．2002，13（2）：109～114

14 王新云，吴洪飞，王仲仁．外压作用下球壳阀芯失稳的有限元模拟和实验研究．锻压技术，2002，3：34～37

15 D．Y．Yang，J．H Yoon and N．K．Lee．A Practical Method of 3-D Remeshing in Forging of Complicated Parts．Adv．Tech．Plast．，1990，1：171～182

16 李赞，李尚键，肖景容．刚塑性变分原理的新形式及有限元法．机械工程学报，1992（3）：90～96

17 赵国群．锻造过程的正反向数值模拟：［博士学位论文］．上海：上海交通大学，1991

18 Cheng-Tang Wu，Mark E．Botkin and Hui-Ping Wang．7th International LS-DYNA Users Conference，2002，29～40

19 王华君．钟形罩多工序温冷联合成形过程的模拟分析：［博士学位论文］．武汉：华中科技大学，2002

20 K．Naksoo，S．Kobayashi．Preform design in H2 shaped cross sectional axisymmetric forging by the Int．J．M ach．Tools manifact，1990

第 4 篇

材料成形优化设计方法

主　编　赵国群

编　写　赵国群　王广春　管延锦
邹增大　陈　军　田学雷
周华民　王新洪　曲仕尧

审　稿　中国材料工程大典编委会

中国机械工程学会
中国材料研究学会
中国材料工程大典编委会

第1章 概 论

1 材料成形优化设计方法在现代制造业中的地位和作用

纵观世纪之交的世界科学技术发展趋势，可以发现，多学科的相互交叉、相互渗透、相互借鉴和多学科的系统集成已经成为各个学科领域进步和发展的巨大动力，也是当今全球科学技术研究的基本特点之一。对于制造业来讲，21世纪的竞争核心将是新产品的竞争。围绕新产品的竞争，一场以信息技术为特征的制造业革命正在全球波澜壮阔地展开。如何实现高质量、低成本、短周期的新产品、新工艺开发，正是赢得这场战争的关键。

材料成形或材料加工工艺较多，主要包括锻造、冲压、注塑、铸造、焊接等工艺，在机械、汽车、电器、航天航空、军工、建筑等诸多行业发挥着重要作用，是国民经济赖以发展的制造业中的重要制造工艺。就每种材料成形方法而言，主要包括材料、工装（模具）、装备。一般情况下，对于所给定的材料和制品的质量技术要求，其工艺和工装设计是新产品开发的主要工作。如何快速、低廉和准确地设计开发出合理的工艺和工装，是实现新产品开发与制造的关键。材料成形工艺设计涉及因素颇多，设计复杂。长期以来，工程设计人员一直沿用经验类比设计方法来处理工程实际问题，需要通过多次反复地“设计—分析—再设计”的过程，才可能得到一个较为满意的设计方案。显然，这个设计过程是人工试凑与类比分析的过程，不仅需要花费较长的设计时间，而且只限于从少数几个候选方案中进行分析比较。由于设计过程缺乏严格的、科学的定量分析计算，一般很难得到最优或近优的设计方案，特别是对于影响因素较多的复杂工程设计问题，更是如此。其设计周期、成本和质量受到很大影响，远远不能满足当今高质量、低成本、短周期的新产品的开发设计要求。如何综合考虑各工艺参数和过程参数的影响，设计出最佳的工艺和工装，一直是人们努力追求的目标。因此，近年来材料成形优化设计方法的研究成为材料成形领域重要的研究方向之一，对于提升我国材料成形水平和促进我国制造业的发展具有重要的意义。

材料成形优化设计方法的建立与材料成形过程的机理研究、数学建模方法、过程仿真或模拟技术、工程优化方法、计算机和软件技术等的发展密切相关。换言之，材料成形优化设计所研究的就是将上述理论、方法、技术综合应用于材料成形工艺和工装设计，以达到优化设计，提高产品的制造质量，降低产品开发成本和缩短产品开发周期的目的。

随着计算机技术的发展和应用，上世纪50年代发展起来的以线性规划与非线性规划为主要内容的新的数学分支——数学规划被应用于解决工程设计问题，形成了工程设计的新理论和新方法，即工程优化设计理论与方法。工程优化设计研究与实践的不断深入，使传统的工程设计方法发生了根本性的变革，从而把经验的、感性的、类比的传统设计方法转变为科学的、理性的、立足于计算分析的设计方法。特别是近年来，随着数值方法、可靠性设计、计算机辅助设计理论技术的发展及其与工程优化设计方法、人工神经网络、遗传算法等的结合应用，使整个工程设计过程逐步向自动化、集成化、智能化发展，其前景令人鼓舞。

计算机CAD技术和软件的发展使大量CAD/CAM软件如UGⅡ、CATIA、PRO/E等进入市场，并获得广泛应用，使人们完全实现了工程设计问题的三维形状建模。材料成形过程的建模与仿真研究于1962年开始用于铸造过程，进入70年代后，从铸造逐步扩展到锻压、焊接、热处理，在全世界形成了材料成形工艺模拟的研究热潮。经多年研究开发，目前已经出现了许多针对锻造、冲压、铸造、焊接等材料成形工艺模拟的商业软件，并已在生产中得到了广泛应用。材料成形过程的物理模拟与数值模拟实际上仍然是一个试错的验证过程，仍然依据设计者的经验知识对试错的方案进行修改。而建立在模拟和智能判断基础上的优化技术便可以自动地实现工程问题的最优化，计算机三维设计技术、工程问题建模与仿真技术、工程优化方法的发展以及人们对材料成形优化设计的迫切要求，使材料成形优化设计研究与应用成为可能，并得到了学术界和工程界的高度重视，其研究力度和应用范围越来越大。

2 材料成形优化设计方法概述

对工程问题进行优化设计，本质上是根据优化设计理论，采用优化设计算法，运用计算机高质量高速度地完成设计任务。为此，首先要把工程设计问题转化为数学模型，即用数学表达式描述工程设计问题；然后，按照数学模型的特点选择优化设计方法及其计算程序，运用计算机求得最优解，即最优设计方案。因此，工程优化设计主要包括建立数学模型和运用优化方法求解这两个方面的内容。根据工程优化设计问题的数学模型的某些特性，可将其分为不同类型。若在工程优化设计问题的数学模型中只含有目标函数而无约束条件，则称之为无约束优化问题；若在数学模型中含有约束条件，则称之为约束优化问题。无约束优化问题在工程设计中并不多见，但相当一部分约束优化问题的解法需建立在无约束优化问题的解法基础之上。

在材料成形领域，成形工艺繁多，每种工艺所涉及的工艺参数、过程参数和设计内容不同，其优化设计的目标函数、优化设计变量、优化步骤以及所采用的优化算法也不尽相同。以下将分别论述不同材料成形过程优化设计方法的研究对象和应用范围。

2.1 金属体积塑性成形优化设计方法

锻造、挤压、轧制、拉拔等金属体积塑性成形技术是一种少、无切削的金属加工方法，广泛应用于工业生产部门。金属体积塑性成形过程作为一个系统，其研究应包括对成形金属、模具、成形金属与模具作用的界面条件、加工设备以及最终产品特性的研究等。由于塑性成形过程伴随着很大的塑性变形，而且既存在着几何非线性特征，又有物理非线性，加之初始边界条件的复杂性及其数学处理上的困难，因而长期以来人们只能通过采用简化、假设和利用实验、经验数据、图表、模型等方法来进行理论分析和应用设计。金属体积塑性成形理论就其内容来分有金属塑性变形的物理-化学理论和数学-力学理论。其中，金属塑性成形的数学-力学理论主要是用宏观数学-力学分析方法求解塑性成形过程的力能参数、变形参数以及应力、应变、温度的分布规律和金属的塑性流动模式等，而这些理论问题正是金属塑性成形工艺制订、模具设计、产品质量控制与提高的重要基础。金属塑性成形的数学-力学分析方法可分为以平衡微分方程为基础的分析方法，例如主应力法、滑移线法、Neviea-Strokes方法和以能量守恒定律为基础的方法，例如上限法/上限单元法、Hill的一般解法、有限元法等。其中，主应力

法、滑移线法、Neviea－Strokes方法、上限法、Hill的一般解法目前主要对二维成形过程的分析具有一定的适用性。

以有限元方法为主的金属体积塑性成形过程的数值模拟技术与模具CAD/CAM技术的结合已成为金属塑性成形模具与工艺设计的核心技术，其商业软件DEFORM、FORGE3、QFORM、CASFORM等在工业生产中已经得到了广泛应用。体积塑性成形模拟技术能够真实地描述金属塑性成形过程材料的流动全貌，可方便地确定塑性成形过程各个阶段所需的变形功和载荷，获得工件的内部应力、应变、温度分布和金属流动规律，获得模具的应力、应变、温度分布和合理形状，预测工件的成形状况、残余应力、缺陷、晶粒的粒度和取向分布，从而在计算机上形象地显示出金属体积塑性成形过程的虚拟制造过程。

影响金属体积塑性成形过程的因素很多，掌握塑性成形的最佳路径对于模具与工艺设计至关重要。但实现金属塑性成形过程的优化设计具有相当大的难度。金属体积成形过程通常为非稳态大变形问题，其产品的质量主要取决于模具设计的水平。对于形状复杂的锻件，绝大多数采用多工序加工，工序数目的确定和各工序模具形状的设计，即预成形设计，是金属体积塑性成形工艺与模具设计的主要工作。近年来，基于数值模拟的计算机辅助设计，由于利用数值模拟方法，为塑性成形过程的优化设计提供了有力工具，国内外许多研究者开展了有关体积塑性成形过程优化设计方法的研究，特别是预成形设计这一重要问题的研究，并取得了较大进展。

金属体积塑性成形优化设计方法主要通过优化工艺参数和预成形模具形状达到所要求的最终成形件的形状和组织性能。目前体积塑性成形优化设计方法主要有基于数值模拟技术的预成形设计方法、基于灵敏度分析的模具优化设计方法和遗传算法等。

基于数值模拟的预成形设计方法主要有反向模拟法和正向模拟法。反向模拟是从具有假定的应变分布和速度分布的终锻件开始，反向跟踪到给定工件形状和应变分布所经历的变形路径，以确定满足产品技术要求的预成形件形状和初始坯料，故这种模拟方式又称为反向跟踪。反向模拟技术中广泛使用的方法是有限元法和上限元法。反向模拟本身不是一种结果惟一的技术，由于变形过程和顺序的多样性，根据边界条件的不同选择可以得到不同的预成形形状。因此，边界条件的处理就成为反向模拟技术中的关键问题。目前二维问题的边界还没有通用的准则，对于三维问题，处理节点脱模的准则问题更为复杂。因此，如何解决好节点的边界条件问题是决定反向模拟技术未来应用的关键所在。正向模拟是从坯料开始，模拟工件在模具作用下的整个成形过程，获得工件的变形状况、各种场量分布以及成形件的成形状况等，设计者由模拟结果验证设计结果的可行性，如果成形过程存在问题，则设计者根据模拟结果修正工艺和模具设计方案，再次进行设计方案的模拟验证，直至获得较理想的设计结果。正向模拟中广泛使用的是有限元法，这种模拟设计可归结为利用数值模拟技术进行设计结果验证的试错法。其基本思路仍与传统的试错法思路一样，只不过所用的验证手段不同，对不合理设计的修改方案还是需要由设计者根据模拟结果和对影响模拟结果的因素的经验提出，设计过程的自动化问题还有待于进一步提高。

在锻造行业，在努力降低设计与制造成本的同时，提高锻件的性能和质量已成为大势所趋。因而必须对成形工艺进行优化设计，以获得少无飞边高性能的锻件。同时，优化过程应实现自动化，从而提高设计效率，减少设计过程的盲目性。目前具有代表性的优化设计方法有基于灵敏度分析的正向模拟优化法、直接微分法、拟合优化法等。

基于灵敏度分析的优化方法属于梯度型优化方法。这种方法首先确定目标函数和设计变量，然后找出目标函数和设计变量之间的函数关系，推导目标函数对设计变量的灵敏度方程，根据设计变量的现有值求解出这些灵敏度信息，再利用有效的优化方法确定设计变量的最优搜索方向，得到更优的设计变量，再求解灵敏度信息，如此反复，直至优化迭代收敛为止。优化设计目标一般为通过优化设计预锻模具形状或预成形件形状和坯料形状获得尺寸合格、少无飞边的终锻件，也包括通过优化工艺参数和模具形状参数，达到降低载荷和改善终锻件组织与力学性能的目的，以及包括多种目标要求的多目标优化设计等。这种方法涉及的金属塑性成形过程的灵敏度方程推导较复杂。虽然节点坐标对设计变量和节点速度对设计变量的灵敏度是通用的，但目标函数对设计变量的灵敏度是针对特定的目标函数和特定的设计变量推导得出，并固化在优化程序中。因此，如要改变目标函数和设计变量，由目标函数对设计变量的灵敏度到节点坐标对设计变量或节点速度对设计变量的灵敏度之间的转换关系就需要重新推导，并重新编制一部分优化程序。为此，可将常用的灵敏度信息固化在优化程序中，根据不同的目标函数调用不同的灵敏度信息。这部分工作还有待于完善。

直接微分法是一种基于求导技术的方法。在该方法中通过求解由有限元等式的摄动产生的微分形式的矩阵方程来得到状态变量对设计变量的微分。然后应用这些微分值计算当前设计准则基于其上的设计灵敏度。该方法通过优化塑性成形的预成形件和模具形状，达到获得合理终锻件的目的，且可以处理多目标函数的问题，例如，可以通过控制工件变形均匀性、温度分布均匀性和工件轮廓与预定直线偏移量的目标函数，对成形工艺进行净成形优化设计、均匀变形优化设计和均匀温度的优化设计。但优化迭代的次数较多，一般在20次以上，说明直接微分法的收敛速度不够理想，这在一定程度上限制了它在预成形设计上的发展。

拟合优化方法思想来源于外推法。这种方法用简单的函数来近似逼近目标函数和设计变量之间的函数关系，通过求解这个简单函数的极值点来推测真实函数的极值点。拟合优化法属于非梯度型直接优化法，它不需要求解灵敏度信息，而只需要根据目标函数值确定优化方向。人们用这种方法对轴对称闭式模锻等工艺的预成形进行了优化设计，设计目标是使锻件中奥氏体颗粒分布尽量一致。模具形状用三次多项式表示，设计变量为三次多项式的系数。这种方法的优点在于不需要触及有限元程序代码，只要有数值模拟软件，就可以进行优化设计。但对设计变量个数有一定的限制。当设计变量个数大于3时，有可能会使拟合过程失效。

遗传算法是一种模拟生命进化机制的搜索和优化算法，它利用“优胜劣汰”的生命机制进行优选，是一种直接优化法。主要用于自动控制、机器人学、神经网络中。与其他算法相比，遗传算法具有全局优化和隐含并行性的优点，尤其适用于设计灵敏度可能是不连续的或者目标函数具有多个极值点，设计迭代只依赖于目标函数的情况。根据遗传算法的群体规模，遗传算法又分为简单遗传算法和微观遗传算法。简单遗传算法的群体规模一般为30～200，且规模越大越有利于防止过早收敛问题的出现，但由于金属成形问题是一个几何非线性和物理非线性耦合的复杂过程，此过程模拟比较费时，而遗传算法必须通过数值模拟来获得目标函数，计算耗时较大，简单遗传算法不适应于分析金属成形过程。而微观遗传算法可克服此弊端，群体规模可少至6，且具有自适应性，效率更高。微观遗传算法可用于工艺参数和模具形状的优化设计。该方法具有较强的专业性，群体、染色体、适应值、模式等概念较多，编码、复制、杂交、变异等基本操作较多，且复制、杂交和变异等操作都会对模式产生影响，所有这些操作有可能会影响到它在材料成形优化设计中的作用。

在金属体积塑性成形过程预成形优化设计方面，还有一致变换方法及理论流动模型、神经网络、形状内插值法、改进的约束变尺度法等。

2.2 金属板料冲压成形优化设计方法

金属板料冲压工艺优化设计，就是采用现代计算机技术、现代数值计算和优化方法，对板料冲压成形工艺安排（如冲裁、弯曲和拉深工艺等）及其各工序的模具设计进行最佳合理方案搜索，获得最佳设计方案，从而减少实际的物理试模次数，实现少无缺陷冲压成形，达到提高材料利用率、提高生产率和降低产品开发成本的目的。

板料冲压工艺是一种先进的金属制品少无切削制造方法，以金属板料塑性变形为基础，利用模具和冲压设备对金属板料进行加工，以获得所需要的零件形状和尺寸。冲压工艺与切削加工相比，具有生产率高、加工成本低、材料利用率高、产品尺寸精度稳定和容易实现机械操作的自动化等一系列优点，尤其适用于大批量生产。

冲压工艺大致可以分为分离工序和成形工序两大类。分离工序又分为落料、冲孔和切割等；成形工序则分为弯曲、拉深、翻边、缩口、胀形和旋压等。根据零件的形状、尺寸精度和技术要求，可以采用以上工序的组合对板料毛坯进行加工，从而获得所要求的制品。冲压工艺和模具设计是一个反复试错的过程，传统的冲模设计采用实际试模的方法，导致模具开发周期长，成本高。随着现代计算机技术、数值计算方法和有限元数值模拟技术的发展，模具设计的合理性检验可以通过计算机分析快速实现，从而使得设计人员可以及时地对不合理的模具设计进行修改。通过工艺优化，可以利用计算机代劳大量的工程计算工作，在设计的合理空间内进行高效的搜索，最终获得比较合理的模具设计。因此，板料冲压工艺的设计优化，可以大大降低试模次数，显著缩短冲模开发时间，降低开发成本。

冲裁工艺优化的核心是在获得断面质量符合要求的冲裁件的前提下进一步提高材料利用率；弯曲工艺优化的核心主要是有效地控制下料尺寸，同时，有效控制和补偿回弹角度；拉深成形工艺优化的核心主要是获得合理的下料尺寸、合理的变形力控制，避免起皱和拉裂，对于复杂的成形过程，还要控制和补偿回弹。

板料冲压工艺及其过程优化所采用的优化方法除了第2章所列出的工程优化设计方法外，还经常使用正交实验法、人工神经网络法、基因遗传算法和响应面法等。

正交实验法，即常规的正交设计方法，以数值模拟为工具，进行设计实验，从而可以发现冲压质量关于各工艺参数的灵敏度，为工艺优化奠定基础，从而解决各种板料成形工艺中各工艺参数与冲压质量之间函数关系难以建立，甚至不清楚各工艺参数影响冲压质量的灵敏度的问题。

人工神经网络是一种大规模的分布式并行处理系统，它可以模拟人脑的记忆和联想功能，不需要了解过程的输入和输出参数之间的变化规律。通过对给定的样本数据进行学习，从大量的数据中提取规则，以一组权重系数的形式形成一种网络的稳定状态，然后通过联想记忆和推广能力来获取所需数据。

基于基因遗传算法的板料成形过程优化设计的原理、步骤与体积塑性成形过程的类似。

响应面法是以试验设计为基础的用于处理多变量问题建模和分析的统计处理技术，该方法用于优化设计一般包括因素的筛分、寻优、寻域和优化。

2.3 注塑成形优化设计方法

如何以更低的价格、更好的质量，及时地将产品推向市场、赢得竞争，是注射成形企业的主要目标。要使产品具有市场竞争力，产品质量起着关键作用。注射成形零件质量可以用力学性能、尺寸精度及外观等来评价。影响注射成形零件质量的因素很多，包括：成形方案（与成形有关的模具结构，主要包括浇注系统和冷却系统）、材料性能（主要包括聚合物材料的流变性能、物理性能、力学性能等）、成形过程中的工艺参数（主要包括注射温度、注射速度、保压压力、保压时间、冷却介质温度、冷却介质流速等众多变量）等。

成形工艺参数是影响注射成形零件质量非常重要的一个方面，注射成形的工艺参数通常有10多个，所有参数都可以影响到零件的质量。针对工艺参数对成形缺陷及产品机械性能的影响，人们已经进行了许多研究工作。实验表明成形参数对产品质量有很大的影响，如：增加保压压力和保压时间可以减少凹陷、对无定型材料降低注射速度有利于消除流痕等。

传统上注射成形的工艺参数由有经验的试模人员来确定。在实践中，试模师根据以前工作中类似的零件来设定当前零件的初始工艺方案，然后不断地调整、试验，直到获得满意的零件，调整过程的长短依赖于工艺师个人的经验。然而，由于工业企业对有经验工艺师的需求不断增长，导致有经验的工艺师远远供不应求，通常一个工艺师需要10~20年的经验才能成为试模专家。

面对注射成形企业的全球化竞争，采用传统的试错法（trial - and - error）来确定成形工艺参数已不可行，大量的研究工作都在尝试采用不同的方法来优化注射成形工艺参数，以便缩短市场响应时间和获得满意的产品质量。注射成形工艺参数优化方法及其相关技术主要包括数学模型、数值模拟、工艺窗口、实验设计方法、专家系统、人工神经网络、基于实例的推理、遗传算法等。优化工艺内容主要包括：浇注系统优化设计、冷却系统优化设计、注射成形工艺参数优化和充填模式优化等。

2.4 铸造成形优化设计方法

铸造过程是一个庞大而复杂的系统，在长期的生产实际中积累了许多经验，这些经验对铸造过程的工艺优化起到了积极的促进作用。随着科学技术的发展，一些优化方法在铸造过程中得到了初步应用。计算机技术的发展，使铸件凝固模拟技术成功地应用于铸造生产的工艺设计，为铸造工艺的优化提供了可靠的依据。

铸造工艺的影响因素较多且复杂，例如，浇注位置、分型面、浇注系统及其形式、冒口设计、铸件结构、熔炼过程配料优化等，上述工艺参数均为铸造过程的主要参数，在设计过程中都应进行优化设计。由于问题的复杂性，在优化方面往往就单一问题进行处理，且某些参数的优化尚无较科学的方法。目前在铸造成形工艺优化方面，主要有冒口优化设计、熔炼过程配料优化以及基于数值模拟技术的虚拟试错实验方法等。

和其他成形工艺优化一样，铸造工艺优化一般也需要建立优化目标，确定优化变量和约束条件，选择优化算法等。例如，如果将铸件的缩松达到最小值为目标函数，其值大小可视铸件生产的执行标准而定；冒口的高度和直径作为设计变量，同时由于其尺寸要小于铸型的尺寸，故也是约束条件。此外，冒口的模数由于要求大于铸件的模数，也是约束条件。采用随机方向法进行最优解的搜索，其求解过程为：模拟→检查目标函数→修正设计变量→再模拟……直至满足所要求的优化收敛目标。

在合金的熔炼过程中，配料是很重要的一个生产环节。配料的结果对合金生产的成本、合金产品的质量以及生产工

艺的设计都有很大影响。不同的熔炼工艺和不同的原材料有不同的配料计算方法，当前最常用的配料优化方法是线性规划中的单纯形法，以及在此基础上发展的大M法、对偶单纯形法等，还有专门对应于块状原材料的整数规划方法。随着现代数学在工程中的不断应用，模糊算法、模拟退火算法、神经网络优化方法等也在优化配料中得到了应用。计算机技术的发展使解决复杂的计算问题越来越简单，传统配料过程中许多计算简单的方法逐渐淘汰，而基于科学计算的配料优化方法得到了越来越广泛的应用。

铸造工艺计算机辅助设计与计算和铸件凝固过程的数值模拟技术已经达到了比较实用化的程度，利用铸造工艺CAD/CAE虚拟试错实验可达到对工艺优化的目的。在铸造工艺CAD过程中，利用计算机根据确定的工艺参数可以完成三维图形下的质量和模数计算、冒口设计计算、浇注系统设计计算和冷铁等的设计计算，铸造工艺CAD完成后得到铸件的铸造工艺三维图形。之后利用铸造工艺CAE对这种工艺下的铸件凝固过程进行数值模拟，并预测在凝固过程中产生的缺陷（如缩孔、缩松等）是否在铸件上。设计者根据模拟结果和设计要求，调整或修改工艺参数，或者采用具有智能化的数值模拟分析软件根据模拟的结果对工艺参数进行调整，然后进行下一次模拟，最终达到合理的工艺参数组合，实现铸造工艺优化。

2.5 焊接成形优化设计方法

优化设计不仅为工程设计提供了一种重要的科学设计方法，使得在解决复杂的设计问题时，能从众多的设计方案中找到尽可能完善的、最适宜的设计方案，而且能够大大提高设计效率和设计质量，获得明显的经济效益和社会效益。

目前最优化技术在焊接领域中主要应用于焊接冶金和材料优化设计、焊接产品优化设计、焊接工艺参数优化设计、焊接自动化优化设计、焊接生产规划设计、焊接生产工艺过程的优化设计、焊接生产备料的优化设计等。

对于焊接冶金和材料配方优化设计，建立数学模型是设计和改进药皮配方的关键环节之一。数学模型必须正确反映药皮组分对焊条性能的影响规律。焊条优化设计通常采用回归设计方法建模，为了得到准确的数学模型，必须合理设计试验方案，严格试验条件，科学地处理试验数据。焊条的优化问题一般属于非线性规划问题，应选用有约束的优化设计方法求解。惩罚函数法是求解这类问题的有效方法，并且在焊条的优化设计中已得到成功的应用。优化方法确定后，编制计算程序和上机计算。计算时应尽量多选几个初始点，以获得全局最优解。

药芯焊丝配方设计是决定药芯焊丝内在质量的关键环节，而配方设计是一个多因子试验设计问题，各因子之间存在着复杂的交互作用，并没有固定的数学模型，因此，在配方设计时常采用经验法或单因子轮换法寻找理想配方。这种设计方法在设计新焊丝时往往带有盲目性。根据药芯焊丝性能的要求，首先确定熔敷金属中元素数量、质量分数及过渡系数，然后采用优化技术设计开发这类产品，不仅可以提高产品质量、缩短生产周期，而且可以降低成本。

焊剂配方优化设计时，要确定混料中各成分因子最佳含量的百分比，首先要选择恰当的试验设计方案进行试验分析；其次要根据目标要求及试验分析结果，选用科学的数学模型及参数估计方法建立目标函数，最后要在一定的约束条件下对目标函数进行寻优，确定混料中各成分因子最佳含量的百分比。

焊接结构设计是整个焊接结构生产中的首道工序，它决定了焊接生产的制造方向及技术要求。焊接结构的形式有多种，在保证经济和安全的条件下，根据使用中不同的受力性质、载荷和环境选用最优的结构材料，根据不同的结构形式选用最优的焊接方法，根据母材及相应的焊接方法选用最优的焊接材料、焊接工艺、焊接前后处理、质量检测方法以及减少残余引力、应力集中的方法。一般的设计目标是使结构具有较高的性能，最有效地利用材料，尽可能地降低成本，优化设计的特点是将实际问题数学模型化，然后利用数学规划的方法进行优化求解。

焊接电源的优化设计过程大致包括建立优化设计的数学模型、确定计算方法、编制计算程序、上机计算、计算结果分析和实用化处理。其优化设计是一个有约束的优化问题，常用方法有罚函数法、网格法及随机试验法等。其中罚函数法适用于中型优化设计问题；网格法和随机试验法适用于小型优化设计问题。

选择焊接方法时必须符合以下要求：能保证焊接产品的质量优良可靠，生产率高，生产费用低，能获得较好的经济效益。影响焊接方法选择的因素主要有产品结构类型、工件厚度、接头型式和焊接位置、母材性能等。上述因素的合理搭配可采用优化设计方法获得，从而根据具体情况确定出最优的焊接方法。焊接方法可采用评价函数法进行优化选择，评价函数的基本原理是先将各优化指标依一定关系进行组合，从中选择一个单目标问题的目标函数，被称为评价函数。利用该评价函数对各种方案进行评价，最终确定最优方案。评价函数法的关键是合理地构造一个能恰当地反映所研究问题的评价函数，工程上常用的评价函数构造方法有平均加权法、线性加权法、虚拟目标法等。

对于焊接工程多阶段决策问题优化设计，其特点是，可将整个问题按时间或空间的顺序分成若干阶段，且只有对其中任意一个阶段做出决策后，才能确定下一阶段的状态，依次对每个阶段做出决策，整个问题的状态即可完全确定。对于这种多阶段决策过程的最优化问题，动态规划是一种有效的求解方法。

3 21世纪材料成形优化设计方法发展展望

近10多年来，国内外对材料成形优化设计方法进行了诸多研究与应用，并取得了较大进展。随着相关科学技术的发展、新型材料加工模式的出现和实际生产对优化设计、优化制造、优化生产的更大需求，材料成形优化设计方法必将得到更大的发展和应用。

在建模技术研究方面，数值分析建模与仿真理论和方法研究将得到高度重视，重点是建模与仿真关键技术的研究和仿真系统的开发，不断拓展材料成形的领域，例如特种成形、特种材料成形等，不断提高仿真结果的精度、解决问题的范围和与实际情况的接近程度，从而实现材料加工过程的全过程精确仿真，为材料成形过程设计方案的验证提供快捷、低廉的手段，进而为材料成形工艺与工模具设计方案的优化与选择提供强有力的工具。

在优化目标函数方面，目前多数情况下，大都针对具有主要影响作用的因素，建立单目标函数，并或多或少地附加一定的约束条件，这种优化将其他较次要因素的取值和影响固定在某一水平上。对于多目标优化设计，如何建立合理的多目标函数是实现全局优化的关键之一，目前人们往往通过不同的加权系数将多个子目标构成一个综合目标，例如，在注塑成形工艺参数优化中，可以将较低的注射压力、较小的温差、较低的剪切速率和切应力等子目标函数综合成一个目标函数，通过不同的加权系数将上述子目标联系起来，同时将设备许可的最大注射压力、熔体降解温度、熔体不流动温度、防止熔体降解和流动失稳的临界值等作为参变量许可的变化范围施加约束。多目标函数中的加权系数对优化结果具有明显影响，如何确定各子目标函数的加权系数是取得合理

优化结果的关键。因此，开展对加权系数的研究有助于构造更科学合理的多目标函数。

对于多数材料成形优化设计，目前的研究大多集中在二维问题上，对于复杂三维问题的研究还有待于进一步加强。例如，对于基于灵敏度分析的锻造过程优化设计，目前的研究多是针对两个工序的成形过程进行模具优化设计，对于更为复杂的多工序成形过程优化设计的研究还较少，目前的研究集中于金属塑性成形过程的单目标优化设计，其多目标优化设计的研究不多；目前的优化方法的收敛性和收敛速度还有待于提高，需要进一步完善或随着研究的发展寻求更高效的适合于体积成形变形状况的优化方法。

对于锻造、挤压、冲压、注塑、铸造等材料成形过程，如何充分考虑模具设计参数与成形工艺参数的耦合，进行工艺参数与模具形状参数的综合优化，是值得今后努力研究的问题。目前材料成形过程的数值模拟软件已经达到了实用化的程度，在实施成形过程模拟时，必须进行模型定义与建立、求解与解析结果，如果结果表明有必要修改设计，则必须改变模型的几何形状，并重复上述步骤，当模型较复杂和修改次数较多时，其过程繁杂费时，因此，对于材料成形工艺参数优化，通过引入CAD中参数化以及参数化设计的概念，可以大大节约时间，避免重复性劳动，减少出错时间，提高分析效率，为此，近年来，又出现了参数化数值模拟方法，参数化数值模拟方法与优化软件结合，可提高优化效率和简化优化步骤。

遗传算法将在材料成形优化设计中继续发挥重要作用。结合材料成形工艺改进遗传算法，解决容易陷入局部最优和收敛速度较慢的问题，提高优化过程中的全局搜索能力和进化效率。将具有实现函数快速非线性映射能力的人工神经网络技术与具有并行计算及全局优化能力的遗传算法结合，利用各自的优势，充分利用数值模拟平台，进行材料成形过程的优化设计，也将成为今后研究的重点和热点之一。

目前，理论上的优化结果与工程实际中的应用还存在一些差异，这是由于工程实际中不仅要考虑产品质量，还要考虑工作效率、生产的经济合理性、设备及工艺条件的限制等诸多因素。因此，有必要研究和探索更加完善的优化目标函数，使优化包含的内容更加准确，得到的优化结果更能符合实际情况。新产品开发过程的环节较多，其中，成形工艺与工模具设计只是其中的重要组成部分。因此，在产品设计与开发过程中，应当综合考虑开发过程的各个环节，以达到产品开发过程的整体优化，技术集成与系统集成将在今后产品开发过程中发挥重要作用。新产品的开发从三维造型设计开始，通过结构受力或运动分析与优化设计，达到结构设计要求，并同时考虑实际生产时材料成形的工艺性，从成形工艺的难度、模具形状的复杂程度及其工序数目的多少等方面综合考虑，从而在外观、结构、运动或受力状态、成形工艺、模具开发成本等多方面达到综合平衡。采用快速成形与制造技术，快速制造出样品，进行结构、装配、外观等的综合评估或试验，在满足要求的前提下，进行小批量产品制造，以配合整套产品的评价和市场试营销等，当上述设计开发环节结束后，进行批量生产用模具和加工工艺的设计与优化，进行生产用工模具的加工制造、试模、生产，进而完成产品的整个设计、开发和生产过程。在上述设计开发过程中，数字化设计与制造贯彻于产品的整个开发过程，并且与成形工艺仿真及优化技术集成，可快速完成产品设计，并有效地确保所设计的产品结构最优和工艺可行，从而有效地提高产品的开发速度、开发质量和减少开发成本与风险。

在材料成形领域，结构优化设计方法、数值模拟技术、成形过程优化设计理论与方法、模具计算机辅助设计与制造、快速成形技术、反求技术、专家系统、人工智能的不断发展和应用，必将使材料成形工艺与工模具设计最终走向数字化、智能化、快速化和最优化。

编写：赵国群（山东大学）

第 2 章　工程优化设计方法

材料成形工艺模拟研究于 1962 年开始于铸造工程，进入 20 世纪 70 年代后，从铸造逐步扩展到锻压、焊接、热处理，在全世界形成了材料成形工艺模拟的研究热潮。经多年研究开发，目前已经出现许多针对常规锻造、冲压、铸造等材料成形工艺模拟的商业软件并已在生产中得到了广泛应用。

材料成形过程的物理与数值模拟实际上仍然是一个试错的验证过程，仍然强烈地依据设计者的经验和知识。而建立在模拟基础上的优化技术便可以自动地实现工程问题的最优化。因此，工程优化方法和技术在材料成形数值模拟技术发展到一定成熟阶段后便得到了学术界和工程界的高度重视。

长期以来，工程设计者一直沿用经验类比设计方法来处理工程实际问题，需要通过多次反复地“设计—分析—再设计”过程，才可能得到一个较为满意的设计方案。显然，这个设计过程是人工试凑与类比分析的过程，不仅需要花费较长的设计时间，而且只限于在少数几个候选方案中进行分析比较。由于设计过程中缺乏严格与科学的定量分析计算，一般很难得到近乎最优的设计方案，特别是对于影响因素很多的复杂工程设计问题，更是如此。

随着计算机技术的发展和应用，上世纪 50 年代发展起来的以线性规划与非线性规划为主要内容的新的数学分支—数学规划被应用于解决工程设计问题，形成了工程设计的新理论和新方法，即工程优化设计理论与方法。工程优化设计研究与实践的不断深入，使传统的工程设计方法发生了根本性的变革，从而把经验的、感性的、类比的传统设计方法转变为科学的、理性的、立足于计算分析的设计方法。特别是近年来，随着有限元法、可靠性设计、计算机辅助设计的理论与方法的发展以及它们与优化设计方法的结合应用，使整个工程设计过程逐步向自动化、集成化、智能化发展，其前景是令人鼓舞的。

对工程问题进行优化设计，本质上是根据优化设计理论，采用优化设计算法，运用计算机高质量高速度地完成设计任务。为此，首先要把工程设计问题转化为数学模型，即用数学表达式描述工程设计问题；然后，按照数学模型的特点选择优化设计方法及其计算程序，运用计算机求得最优解，即最优设计方案。因此，工程优化设计包括建立数学模型和运用优化方法求解这两个方面的重要内容。

根据工程优化设计问题的数学模型的某些特性，可将其分为不同类型。若工程优化设计问题的数学模型中只含有目标函数而无约束条件，则称之为无约束优化问题；若数学模型中含有约束条件，则称之为约束优化问题。虽然无约束优化问题在工程设计中并不多见，但由于有相当一部分约束优化问题的解法是建立在无约束优化问题的解法基础之上的，因而无约束优化问题及其解法在本章中也要重点介绍。

无约束优化问题的目标函数如果是一元函数，则称之为一维优化问题，其求解方法称之为一维搜索方法。对于约束优化问题，可按其目标函数与约束函数的特性，分为线性规划问题和非线性规划问题。如果目标函数和所有的约束函数都是线性函数，则称之为线性规划问题；否则，则称之为非线性规划问题。对于目标函数是二次函数而约束函数都是线性函数这一类的问题，一般称之为二次规划问题。

本章仅简要介绍几种在材料成形领域中常用的工程优化设计方法，为后面几章的内容提供一定的理论基础，有关更深广方面的工程优化设计方法，请参考有关文献。

1　工程优化设计的数学模型

工程设计问题一般都可以用数学表达式来描述，即转化为数学模型。一个优化设计的数学模型通常包含设计变量、目标函数和约束条件三个基本要素，由它们所组成的工程优化设计数学模型所表达的含义是：在满足一定的约束条件下，寻求一组设计变量值，使得目标函数达到极小值（或极大值）。

(1) 设计变量

在优化设计中不断变化的一组参数（也构成为一个向量），称为设计变量或设计向量。一组设计变量代表一个设计方案，设计变量有连续变量和离散变量两种不同形式。若设计变量的值可在上限与下限之间连续变化，则称之为连续变量；若设计变量的数值是按照某种离散数列变化的，则称为离散变量。设计变量的数目称为优化设计问题的维数。如有 n 个设计变量，则称为 n 维设计问题。若将 n 维设计问题的 n 个设计变量，按一定顺序排列组成一个 n 维向量，即

$$x=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}} \tag{4.2-1}$$

以 n 个设计变量 x_i（$i=1, 2, \cdots, n$）为坐标轴就构成 n 维空间。一般设计变量均为实变量，故称为实欧氏空间，用 R^n 表示。在工程优化设计中，称此实欧氏空间为设计空间。

(2) 目标函数

在确定的设计思想指导下，由设计变量表达的用来评价所追求指标的函数，称为目标函数。目标函数当然是设计变量的函数，一般表示为

$$f(x)=f(x_1,x_2,\cdots,x_n) \tag{4.2-2}$$

目标函数是根据设计要求，按照设计准则建立起来的设计变量和已定参数的非线性多元函数。目标函数在求解时一般都将其求解过程转化为在各种约束条件下求解某一多元非线性函数的极小值问题，即当目标函数在设定的约束条件下具有极小值时，我们就认为它所代表的设计变量是最优的，使 $f(x)$ 达到极小值的解 x^* 就是最优的设计方案。

选择和确定目标函数是优化设计中最重要决策之一。在探索和解决实际工程问题时，应尽量把要解决的目标集中，即集中解决主要目标。当然，在实际的工程设计和工程问题中，也常常存在多目标问题，即在考虑某一种目标时又必须顾及其他方面的要求，所以工程实际的优化问题也常常要处理多目标问题。这种在优化设计中要求几项指标同时达到最优值的问题，称为多目标优化问题。一个多目标优化问题的目标函数，表达式将变为

$$\min f(x) \qquad x\in R^n$$

$$f(x)=[f_1(x),f_2(x),\cdots,f_k(x)]^{\mathrm{T}} \tag{4.2-3}$$

式中，$f_1(x),f_2(x),\cdots,f_k(x)$ 称为分目标问题，它们具有不同的含义和完全独立的关系。

一般来说，处理多目标优化问题的主要方法是线性加权组合法，其基本原理是总目标函数 $f(x)$ 为各分目标函数加权后的线性叠加

$$f(x)=\sum_{i=1}^{k} w_i f_i(x) \tag{4.2-4}$$

式中，k 为分目标个数；w_i 加权因子，为大于 0 小于 1 的

数，其选值是整个多目标优化问题的关键。

(3) 约束条件

在优化设计过程中，设计变量不断改变其取值，以期达到目标函数的最小值，但设计变量的改变和取值要受到一系列的限制和约束，问题本身对于设计变量的一些限制条件，构成了对设计变量的约束条件。对设计变量的约束条件，均以等式和不等式两种形式表示为一系列的约束方程

不等式约束：$g_u(x) \geqslant 0 \quad (u=1,2,\cdots,m)$

或者　$g_u(x) \leqslant 0$

亦可将其一律转化为 $g_u(x) \geqslant 0$ 形式，只须前面冠以负号。

等式约束：　$h_v(x)=0 \quad (v=1,2,\cdots,p<n)$

等式约束方程的总数 p，不应等于和超过设计变量的个数，否则未知量将全部有定解而无所谓有什么极小值问题，优化问题变得毫无意义。对于不等式约束方程的个数 m，并无限制。

优化数学模型的建立过程实际上就是确定设计变量、目标函数和约束条件的过程，一个优化问题的命题是否正确，应看其设计变量及其约束条件的选择和设定是否合理，目标函数的表达式是否协调，这三者在逻辑结构上是否严谨无误。可以说解决了优化数学模型问题，也就解决了优化问题。

根据目标函数和约束条件的表达形式可将工程优化问题分为以下三类。

1) 无约束优化问题　无约束优化问题的一般表达式为

$$\min f(x) \qquad x \in R^n$$

虽然无约束优化问题在工程设计中并不多见,但由于有相当一部分约束优化问题都是建立在无约束优化问题解法基础上的,因而此类优化问题及其解法应给予重视。

2)等约束优化问题　等约束优化问题的一般表达形式为

$$\min f(x) \qquad x \in R^n$$
$$\text{s.t. } h_v(x)=0$$

3) 等约束和不等约束优化问题，即一般形式的优化问题该类一般形式优化问题的表达形式为

$$\min f(x) \qquad x \in R^n$$
$$\text{s.t. } h_v(x)=0$$
$$g_u(x) \geqslant 0$$

2　工程优化问题的迭代算法

根据无约束工程优化问题的极值必要条件，可得到求解极值点的联立方程组，似乎求解这些联立方程组，就可以完全解决无约束工程优化问题。但实际上，能用这种方法求解的工程优化问题并不多，它只适用于设计变量少且低次的函数，对于高次多元目标函数，根据极值条件所得到的方程组一般仍是高次多元方程组。要联立求解这种高次非线性方程组是很困难的，甚至是不可能的。至于约束工程优化问题，利用约束极值条件来求解，情况更为复杂。因而实用工程优化方法大多不采用这种解析方法，而代之以数值迭代算法。

(1) 数值迭代算法的迭代格式

工程优化方法的数值迭代算法是基于计算机高速运算特点构造而成的，它不是解析方法，而是按同一格式进行反复迭代计算，逐步逼近工程优化问题最优解的方法。其基本迭代格式是：从某一点 $x^{(k)}$ 出发，根据目标函数与约束函数在该点的某些信息（如函数值、一阶导数、二阶导数等），确定本次迭代方向 $s^{(k)}$ 和适当步长 α_k，按式 (4.2-5) 求得一个新点 $x^{(k+1)}$，然后判断此新点是否为问题的最优解。如此反复迭代，直到求得问题的最优解为止。

$$x^{(k+1)}=x^{(k)}+\alpha_k s^{(k)} \qquad (k=1,\ 2,\ \cdots) \tag{4.2-5}$$

由此可知，工程优化问题的迭代算法中要解决的主要问题是确定迭代方向 $s^{(k)}$ 和适当的步长因子 α_k。

确定迭代方向 $s^{(k)}$ 的基本原则应当是使上述迭代过程向最优解逼近的过程。对于求解极小化的问题，当然应使方向 $s^{(k)}$ 是函数值下降的方向。我们知道，负梯度方向是最速下降的方向，因而凡与负梯度方向 $-\nabla f(x^{(k)})$ 成锐角，或者与梯度方向成钝角的方向，即满足 $\|\nabla f(x^{(k)})\| \cdot \|s^{(k)}\| \cos(\nabla f(x^{(k)}), s^{(k)})<0$ 的方向 $s^{(k)}$，均可作为迭代方向。

当下降的迭代方向 $s^{(k)}$ 确定以后，一般希望由迭代点 $x^{(k)}$ 出发，沿该方向前进一个适当的步长，使之经过此迭代求得的迭代点 $x^{(k+1)}$ 是此方向具有最小函数值的点。因此，步长因子 α_k 一般应根据此原则来确定，并称这样确定的步长因子为最优步长因子。不过，有时确定最优步长因子并不方便，故也可以取一个适当步长因子。所谓适当步长因子是指按此步长因子沿所确定的迭代方向前进一步，可得到函数值有一定程度下降的迭代点。

(2) 数值迭代算法的评价准则

由前述工程优化问题的数值迭代算法的迭代格式可知，迭代算法的关键步骤是确定迭代方向和步长因子。因而根据迭代方向的确定方法的不同，可形成不同的优化算法。一般将可靠性、有效性和收敛性作为优化迭代算法的评价准则。

算法的可靠性是指在给定的合理精度要求下，经过一定的计算时间和一定的迭代步骤，求解优化问题的计算成功率。而算法的成功率一般是通过考查所选定的一组考题的求解过程进行解题测试获得的。

算法的有效性是指算法的解题效率，一般可用解题的平均计算时间或函数（含目标函数、约束函数及其梯度）的求值次数来衡量。但由于采用不同算法计算一次函数值所花费的计算时间可能相差很大，因此将解题的平均计算时间作为衡量解题效率的主要指标，而将函数求值次数作为辅助指标。

算法的收敛性是评价算法优劣的主要指标。所谓算法收敛，就是随着迭代次数的增加，最终求得的迭代点 $x^{(k)}$ 与优化问题的极值点 x^* 在设计空间中的距离可以达到任意小，即

$$\lim_{k\to\infty}\|x^{(k)}-x^*\|=0$$

收敛速度的快慢也是衡量算法收敛性的重要指标，一般根据算法收敛速度的快慢将算法分为线性收敛（即一阶收敛）、超线性收敛和二阶收敛算法。

若此算法的迭代点序列 $\{x^{(k)}\}$ $(k=1,2,\cdots)$ 到极小点的距离依次按等比序列减小，即对于某一常数 β，有不等式 (4.2-6) 成立，则称此算法为线性收敛算法。推而广之，可将满足不等式 (4.2-7) 的算法称之为 α 阶收敛算法。当 $\alpha=1$，即式 (4.2-6)，此算法为线性（一阶）收敛；当 $\alpha=2$，则称此算法为二阶收敛；当 $1<\alpha<2$ 时，则称此算法为超线性收敛。显然，二阶收敛速度最高，线性收敛速度最低，而超线性收敛速度介于两者之间。一种算法如果具有超线性收敛速度，便可以认为是一个好的算法。

$$\frac{\|x^{(k)}-x^*\|}{\|x^{(k-1)}-x^*\|} \leqslant \beta \ (k=1,2,\cdots) \tag{4.2-6}$$

$$\frac{\|x^{(k)}-x^*\|}{\|x^{(k-1)}-x^*\|^{\alpha}} \leqslant \beta \ (k=1,2,\cdots) \tag{4.2-7}$$

(3) 数值迭代算法的终止准则

任何一种下降迭代算法的迭代过程都是逐步逼近极小点的过程，要收敛于极小点，往往需要很多的迭代次数，计算工作量和时间较多。因此，一般在迭代点相当接近于极小点时，就可认为已求得近似最优点而终止迭代。为此，必须规定迭代点近似最优点的程度准则，即终止准则。由于函数的

理论极小点 x^* 是事先不知道的，因而无法用 x^* 点来确定迭代点与其接近的程度，即无法计算 $\| x^{(k)} - x^* \|$。因此常采用以下三种准则作为迭代算法的终止准则。

1）前后两个迭代点间的距离达到充分小，即对于给定的足够小量 ε_1（$\varepsilon_1 > 0$），有不等式（4.2-8）成立，则一般可认为 $x^{(k+1)}$［或 $x^{(k+1)}$ 和 $x^{(k)}$ 的平均值］是近似最优点，而终止迭代。

$$\| x^{(k+1)} - x^{(k)} \| \leqslant \varepsilon_1 \qquad (4.2\text{-}8)$$

2）前后两个迭代点的目标函数值充分接近，即对于给定的足够小量 ε_2（$\varepsilon_2 > 0$），有不等式（4.2-9）成立，则可终止迭代。

$$\| f(x^{(k+1)}) - f(x^{(k)}) \| \leqslant \varepsilon_2 \qquad (4.2\text{-}9)$$

3）根据极值存在的必要条件，可规定式（4.2-10）作为迭代终止准则。其中 ε_3 为充分小的正数。

$$\| \nabla f(x^{(k)}) \| \leqslant \varepsilon_3 \qquad (4.2\text{-}10)$$

应当指出的是，对于 $f(x^{(k)})$ 或 $x^{(k)}$ 分量之绝对值很大的情况，若采用式（4.2-8）或式（4.2-9）作为终止准则，因为准则过分严格，会使迭代过程大为加长。为此可取相对值，分别用以下两式取代：

$$\| x^{(k+1)} - x^{(k)} \| / \| x^{(k)} \| \leqslant \varepsilon_1 \qquad (4.2\text{-}11)$$

$$\| f(x^{(k+1)}) - f(x^{(k)}) \| / \| f(x^{(k)}) \| \leqslant \varepsilon_2 \qquad (4.2\text{-}12)$$

此外，上述三种终止准则都有可能失效的时候，如果要完全可靠地判断迭代点是否充分靠近极小点，则应将上述三个终止准则结合起来使用。

3　无约束优化方法

无约束优化问题的一般数学表达式为

$$\min f(x) \quad x \in R^n \qquad (4.2\text{-}13)$$

求解这类问题的方法，称为无约束优化方法。

虽然工程优化问题几乎都是有约束条件的，但是，有约束条件工程优化问题的优化方法的基本求解策略或者是将约束优化问题转化为无约束优化问题来求解，或者是来源于无约束优化方法。因而可以认为，无约束优化方法是约束优化方法的基础。所以，无约束优化方法在工程优化设计中有着十分重要的作用。

若函数 $f(x)$ 为一元函数，则求解问题的无约束优化方法称之为一维搜索方法。一维搜索方法有：分数法、黄金分割法、二次插值法和三次插值法等。工程优化问题中的目标函数为一元函数的情况并不多见，一般都为多元函数。多元函数的无约束优化方法，可按其确定的搜索方向所使用的信息和方法的不同分为两大类：一类方法是需要利用函数的一阶偏导数甚至二阶偏导数构造搜索方向，如梯度法、牛顿法、变尺度法和共轭梯度法等。由于需要计算偏导数，这类方法计算量大，但收敛快，一般称之为解析法。另一类方法是仅利用迭代点的函数值来构造搜索方向，如坐标轮换法、模式搜索法、Powell 共轭梯度法和单纯形法等。由于只需要计算函数值，对于无法求导或求导困难的函数，这类方法就有突出的优越性，但是其收敛速度较慢，一般称之为直接法。

3.1　一维搜索方法

目标函数为一元函数的一维搜索方法是按照迭代格式（4.2-14）求解的，即从已知点 $x^{(k)}$ 出发，沿给定的方向 $s^{(k)}$ 搜索，以得到目标函数沿 $s^{(k)}$ 方向的极小点，其实质是求 α 的一个最优步长因子 α_k，使式（4.2-15）成立。

$$x^{(k+1)} = x^{(k)} + \alpha s^{(k)} \qquad (4.2\text{-}14)$$

$$f(x^{(k+1)}) = f(x^{(k)} + \alpha_k s^{(k)}) = \min_{\alpha} f(x^{(k)} + \alpha s^{(k)}) \qquad (4.2\text{-}15)$$

由于式（4.2-15）中的 $x^{(k)}$ 和 $s^{(k)}$ 都已确定，所以式（4.2-15）所表达的问题就是以 α 为设计变量的一维无约束优化问题。

一维搜索方法的求解一般分两步进行。第一步是确定函数值最小点的所在区间，即搜索区间；第二步是求出该区间内的最优步长因子 α_k 的值。

（1）搜索区间的确定

根据函数的变化情况，可将区间分为单峰区间和多峰区间。对于单峰区间$[a,b]$的函数值，必然存在高—低—高的变化特征。如果在一维搜索之前，可估计极小点的大致位置，则可以直接给出搜索区间；否则，需采用试算法确定之。常用的确定搜索区间的方法是进退试算法，也称为跨步法，其求解方法如下：

设给定初始点 a 及初始步长 T_0，待定的搜索区间为$[a,b]$。首先计算函数值 $f(a)$ 及 $f(a+T_0)$ 并比较它们大小。若 $f(a) > f(a+T_0)$，则表明极小点在试算点的右侧，特征为高—低，说明跨步方向是正确的，需再跨一步，步长为 $T_1 = 2T_0$，向前试算，直至 $f(a+T_{i-1}) > f(a+T_i)$ 为止，则边界为$[a, a+T_i]$；若 $f(a) < f(a+T_0)$，则表明极小点在试算点的左侧，特征为低—高，说明跨步方向是错的，应采用 $T = -T_0$，后退试算来确定搜索区间。

有时，当确定正确的搜索方向后，可采用倍增步长法，即将步长顺次增大一倍来加速区间的搜索速度。

（2）黄金分割法

黄金分割法的基本思想是：通过比较单峰区间内两点函数值，不断舍弃单峰区间左端或右端一部分，使区间按固定区间缩小率 0.618 逐步缩小，直至极小点所在的区间缩小到给定的误差范围内而得到近似最优解。

黄金分割法的计算步骤如下：

1）给定区间$[a,b]$及允许误差 ε_1 及 ε_2；

2）在区间$[a,b]$内取两个试算点 $\alpha_0^{(1)} = b - \lambda(b-a) = b - 0.618(b-a)$，$\alpha_0^{(2)} = a + \lambda(b-a) = a + 0.618(b-a)$ 并计算函数值 $f_1 = f(\alpha_0^{(1)})$，$f_2 = f(\alpha_0^{(2)})$；

3）若满足收敛判断准则 $\left| \dfrac{f_2 - f_1}{f_1} \right| \leqslant \varepsilon_1$ 则转至第 5）步，否则转至第 4）步；

4）比较函数值 f_1 和 f_2 的大小。若 $f_2 \leqslant f_1$，则取$[\alpha_0^{(1)}, b]$为新区间，而 $\alpha_0^{(2)}$ 则作为新区间内的第一个试算点，即令 $a \Leftarrow \alpha_0^{(1)}$，$\alpha_0^{(1)} \Leftarrow \alpha_0^{(2)}$，$f_1 \Leftarrow f_2$，而另一试算点可按下式计算出：

$$\alpha_0^{(2)} = a + \lambda(b-a) = a + 0.618(b-a)$$

并计算函数值 $f_2 = f(\alpha_0^{(2)})$，然后转第 3）步；若 $f_2 > f_1$，则取$[a, \alpha_0^{(2)}]$为新区间，而 $\alpha_0^{(1)}$ 作为新区间的一个试算点，即令 $b \Leftarrow \alpha_0^{(2)}$，$\alpha_0^{(2)} \Leftarrow \alpha_0^{(1)}$，$f_2 \Leftarrow f_1$，而另一个试算点可按下式计算出：

$$\alpha_0^{(1)} = b - \lambda(b-a) = b - 0.618(b-a)$$

并计算函数值 $f_1 = f(\alpha_0^{(1)})$，然后转第 3）步；

5）若满足收敛判断准则 $\left| \dfrac{\alpha_0^{(2)} - \alpha_0^{(1)}}{\alpha_0^{(1)}} \right| \leqslant \varepsilon_2$ 则转至第 6）步；否则，令 $a \Leftarrow \alpha_0^{(1)}$，$b \Leftarrow \alpha_0^{(2)}$ 并转第 2）步；

6）可令极小点为 $x^* = \alpha^* = \dfrac{1}{2}(\alpha_0^{(2)} + \alpha_0^{(1)})$，$f^* = f(x^*)$ 并输出结果，停止迭代。

与上述步骤一致的计算框图如图 4.2-1 所示。

（3）二次插值法

二次插值法的基本思想是：在给定的单峰区间中，利用函数上的三个点构造一个二次插值多项式来近似代替原来复

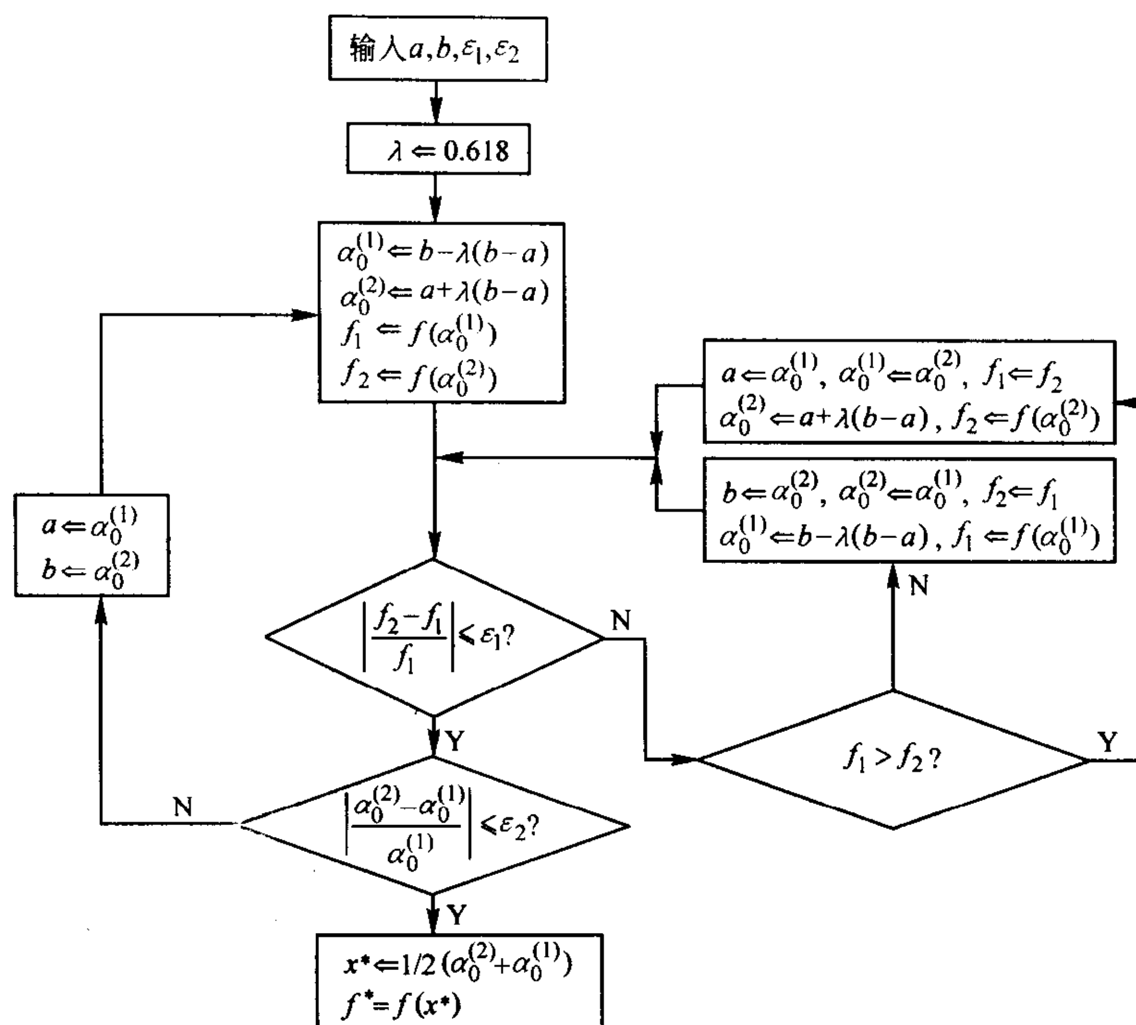


图 4.2-1　黄金分割法计算框图

杂的一元函数，并求出此二次插值多项式的极小点，作为原函数的极小点的近似值。

二次插值法的原理及求解过程如下。

设某一元函数$f(x)$，在单峰区间$[\alpha_1, \alpha_3]$内取一点α_2，且$\alpha_1 < \alpha_2 < \alpha_3$，这三点对应的函数值分别为$f_1 = f(\alpha_1)$，$f_2 = f(\alpha_2)$，$f_3 = f(\alpha_3)$。于是可通过原函数曲线的三个点$(\alpha_1, f_1)$，$(\alpha_2, f_2)$，$(\alpha_3, f_3)$作出一条抛物线（二次曲线），此抛物线方程可表示为

$$p(\alpha) = a + b\alpha + \alpha^2 \tag{4.2-16}$$

由此方程可以很容易求得它的极小点α_p^*，此时求得的α_p^*不一定与原函数的极值点x^*重合。为求得满足预定精度要求的原函数的近似极小点，一般要求进行多次迭代，在已有的四个点α_1、α_2、α_3及α_p^*中选择新的三个点，得到一个缩小了的单峰区间，并利用此单峰区间的三个点再一次进行插值，如此进行下去，直到达到给定精度为止。

3.2　梯度法

梯度法是求解多维无约束优化问题的解析方法。它的基本思想是：将迭代点$x^{(k)}$处的最速下降方向，即负梯度方向$-\nabla f(x^{(k)})$作为搜索方向，按照迭代格式（4.2-17）求得负梯度方向的一个极小点$x^{(k+1)}$，作为原问题的一个近似最优解；若此问题尚不满足精度要求，则再以$x^{(k+1)}$作为迭代起始点，以$x^{(k+1)}$处的负梯度方向$-\nabla f(x^{(k+1)})$作为搜索方向，求得该方向的极小点$x^{(k+2)}$；如此进行下去，直到所求得的解满足收敛条件为止。

$$x^{(k+1)} = x^{(k)} + \alpha s^{(k)} = x^{(k)} - \alpha \nabla f(x^{(k)}) \tag{4.2-17}$$

（1）梯度法的迭代步骤及计算框图

梯度法的迭代步骤如下：

1）给定初始点$x^{(0)}$，允许误差ε，并置$k=0$；

2）计算迭代点的梯度$\nabla f(x^{(k)})$和搜索方向$s^k = -\nabla f(x^{(k)}) / \| \nabla f(x^{(k)}) \|$；

3）如满足$\| \nabla f(x^{(k)}) \| < \varepsilon$，迭代结束并输出$x^{(k)}$，否则进行下一步；

4）进行一维搜索，求最优步长α_k：$\min\limits_{\alpha} f(x^{(k)} + \alpha s^{(k)}) = f(x^{(k)} + \alpha_k s^{(k)})$；

5）求新迭代点$x^{(k+1)} = x^{(k)} + \alpha_k s^{(k)}$；并令$k=k+1$，转第2）步。

程序框图见图4.2-2。

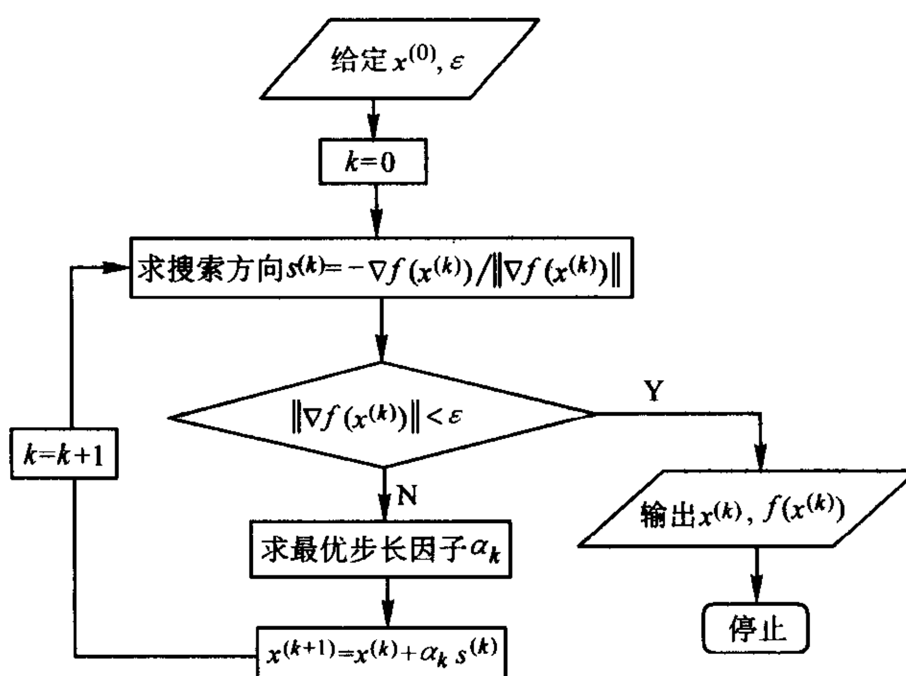


图 4.2-2　梯度法计算框图

（2）梯度法的特点

由于梯度法每次迭代的搜索方向是选用最速下降方向，因此又称为最速下降法。从这点来看，容易使人认为这种方法是一个使函数值下降最快的方法，但实际上并非如此。从任意迭代点$x^{(k)}$沿负梯度方向$-\nabla f(x^{(k)})$搜索极小点$x^{(k+1)}$，这就是求一元函数$f_1(\alpha) = f(x^{(k)} + \alpha s^{(k)})$的极小值问题。可令

$$\frac{\mathrm{d}f_1(\alpha)}{\mathrm{d}\alpha} = [\nabla f(x^{(k)} + \alpha s^{(k)})]^{\mathrm{T}} s^{(k)} = 0 \tag{4.2-18}$$

由于$s^{(k)} = -\nabla f(x^{(k)})$，$x^{(k+1)} = x^{(k)} + \alpha s^{(k)}$，故得$[\nabla f(x^{(k+1)})]^{\mathrm{T}} \nabla f(x^{(k)}) = 0$

上式表明，相邻两次迭代的极小点处梯度正交，即相邻两次迭代的搜索方向互相正交。因而搜索路线成锯齿状，并

且愈接近极小点，其步长愈小，收敛速度愈慢。由此可见，所谓最速下降是指函数在 $x^{(k)}$ 附近出现的局部性质，从全局看，负梯度方向并非是下降最快的方向。

应当指出的是，目标函数的性态对梯度法收敛速度影响极大。例如，当二元二次函数的等值线是比较扁的椭圆时，其梯度法逼近函数极小值的过程呈直角锯齿状；而当等值线是圆时，则可一步达到极小值点。

梯度法的优点是迭代过程简单，而且在迭代最初的步长还是较大的，即在远离极小值点时，函数下降还是比较快的。因此，常将梯度法与其他方法结合，即在迭代的开始几步用梯度法，当接近极小点时，再改用其他算法，以加快收敛速度。

3.3　牛顿法

牛顿法是梯度法的进一步发展，其基本思想是在求目标函数$f(x)$的极小值时，先将它在点 $x^{(k)}$ 附近作泰勒展开提取二次近似函数式，然后求出这个二次函数的极小值点，并以此点作为原目标函数的极小点 x 的一次近似值；若此值不满足精度要求，则以此近似点作为下一次迭代的初始点，求出二次近似值；照此方式迭代下去，直至所求出的近似值极小点满足精度要求为止。

牛顿法求解的具体过程如下。

考虑从 $x^{(k)}$ 到 $x^{(k+1)}$ 的迭代过程，在 $x^{(k)}$ 点用一个二次函数$q(x)$来密切该函数 $f(x)$，在 $x^{(k)}$ 点把 $f(x)$按泰勒公式展开（到第二项为止），见式（4.2-19）。

$$f(x)\approx q(x)=f(x^{(k)})+\nabla f(x^{(k)})^{\mathrm{T}}(x-x^{(k)})+\frac{1}{2}(x-x^{(k)})^{\mathrm{T}}\nabla^2 f(x^{(k)})(x-x^{(k)}) \tag{4.2-19}$$

对 $q(x)$ 求导数，并知 $\nabla q(x^{(k)})=0$，则 $\nabla q(x)=\nabla f(x^{(k)})+\nabla^2 f(x^{(k)})(x-x^{(k)})=0$，$(x-x^{(k)})=-\dfrac{\nabla f(x^{(k)})}{\nabla^2 f(x^{(k)})}$，于是 $x=x^{(k)}-\nabla^2 f(x^{(k)})^{-1}\cdot\nabla f(x^{(k)})$。在极值点附近，我们用($x=x^{(k+1)}$)来近似极值点，得到

$$x=x^{(k)}-\nabla^2 f(x^{(k)})^{-1}\cdot\nabla f(x^{(k)}) \tag{4.2-20}$$

此式即为牛顿法的迭代公式，其搜索方向为 $s^{(k)}=-\nabla^2 f(x^{(k)})^{-1}\cdot\nabla f(x^{(k)})$，此方向称为牛顿方向，步长因子 $\alpha_k=1$。

牛顿法利用 $x^{(k)}$ 点提供的两阶导数信息$\nabla f(x^{(k)})$、$\nabla^2 f(x^{(k)})$、$\nabla^2 f(x^{(k)})^{-1}$来构成它下一次迭代的修正量，而 $\nabla^2 f(x^{(k)})^{-1}\cdot\nabla f(x^{(k)})$比起只有$\nabla f(x^{(k)})$来，不只是考虑了局部的最速下降方向，而且还考虑了$\nabla f(x^{(k)})$本身的变化趋势。因此它是一个考虑了全局的迭代方向，必然具有较快的收敛速度。

3.4　变尺度法

变尺度法（又称 DFP 法）是在克服了梯度法收敛慢和牛顿法计算量大的缺点而发展起来的，是求解无约束优化问题最有效的算法，在工程优化设计中得到了广泛的应用。

变尺度法的基本思想是：利用牛顿法的迭代形式，然而并不直接计算$\nabla^2 f(x^{(k)})^{-1}$，而是利用一个对称正定阵 $\boldsymbol{A}^{(k)}$ 近似地代替$\nabla^2 f(x^{(k)})^{-1}$。$\boldsymbol{A}^{(k)}$ 在迭代过程中不断改进，最后逼近$\nabla^2 f(x^{(k)})^{-1}$。这种方法省去了$\nabla^2 f(x^{(k)})$的计算和求逆，计算量大为减少。

变尺度法的迭代公式为

$$x^{(k+1)}=x^{(k)}-\alpha_k\boldsymbol{A}^{(k)}\nabla f(x^{(k)}) \tag{4.2-21}$$

式中，$-\boldsymbol{A}^{(k)}\nabla f(x^{(k)})$为变尺度法的搜索方向，$\boldsymbol{A}^{(k)}$ 为变尺度矩阵，在迭代过程中逐次形成并不断修正。

当 $\boldsymbol{A}^{(k)}=\boldsymbol{I}$（单位矩阵）时，式（4.2-21）变为

$$x^{(k+1)}=x^{(k)}-\alpha_k\nabla f(x^{(k)}) \tag{4.2-22}$$

而当 $\boldsymbol{A}^{(k)}=\nabla^2 f(x^{(k)})^{-1}$时，式(4.2-21)变为

$$x^{(k+1)}=x^{(k)}-\alpha_k\nabla^2 f(x^{(k)})^{-1}\nabla f(x^{(k)}) \tag{4.2-23}$$

显然，式（4.2-22）就是梯度法的迭代公式，式（4.2-23）就是牛顿法的迭代公式。由此可知，梯度法及牛顿法可看作变尺度法的一种特殊形式。

由上述可知，在迭代开始时，一般令 $\boldsymbol{A}^{(0)}=1$，此时变尺度法的迭代公式为式（4.2-22），它就是梯度法的迭代公式；而当变尺度矩阵逼近$\nabla^2 f(x^{(k)})^{-1}$时，变尺度法迭代方向也逼近牛顿方向，其迭代公式也逼近牛顿法的迭代公式。因而，变尺度法最初的几步迭代与梯度法类似，函数值下降较快；而最后几步迭代与牛顿法相近，可较快收敛到极小点。可见，变尺度法克服了梯度法收敛慢的缺点，但却保留了梯度法在最初几步函数值下降快的优点；同时，变尺度法避免了计算$\nabla^2 f(x^{(k)})$及其逆矩阵，从而克服了牛顿法计算量大的缺点，但却有较快的收敛速度。因此，在目标函数的梯度容易计算的情况下，DFP 变尺度法是一种很有效的方法。

4　约束优化方法

在实际工程中没有约束条件的优化问题是很少见的，很难设想一个实际的工程设计在满足各种设计要求时会不受到一定的限制和各种外界因素及本身的制约。具有约束条件的优化问题的一般形式已为大家所熟知，即：

$$\begin{aligned}&\min f(x) && x\in R^n\\ &\text{s.t.}\ h_i(x)=0 && (i=1,2,\cdots,m)\\ &\quad g_i(x)\leqslant 0 && (i=m+1,m+2,\cdots,p)\end{aligned}$$

其中，$f(x)$、$h_i(x)$、$g_i(x)$都假定具有连续偏导数。

按求解这类约束最优化问题所采用的原始信息和对约束条件处理方法的不同，可将约束优化分为两类：一是直接从可行域中寻找出它的约束最优解，称之为约束优化问题的直接解法；二是将复杂的约束优化问题转化为一系列简单的容易解决的子问题，用这一系列子问题的解去逼近原问题的解，称之为约束优化问题的间接解法。直接解法是依靠直接迭代来求解，且每次迭代中的迭代点都必须落在可行域中。直接迭代法有网格法、随机方向法及复合形法等，这一类方法的优点是算法简单、直观性强、对函数无特殊要求，其缺点是计算量大、收敛慢、效率低，只适用于维数低、精度要求不高的问题。间接法是将有约束的优化问题转化为一系列无约束问题，然后用比较成熟的无约束问题求解方法求出最优化解。间接法的关键是如何根据目标函数和约束条件把两方面的因素巧妙地合适地组织起来构造一个新函数，用它来替代原函数，同时又满足优化条件，使其优化解能足够近似地代替原函数的优化解。

4.1　Lagrange 乘子法

Lagrange 乘子法是一种将约束问题转化为无约束问题的优化方法，它通过引入一些称为乘子的待定系数，把原有的约束条件与这些乘子结合，并添加到原目标函数中去，构成一个无约束的新目标函数，新目标函数的无约束最优解就是原约束问题的最优解。最优化方法的许多近代结果都可以看成是这种方法的推广。

在等式和不等式约束两种条件下的优化问题，虽然两者是相通的，但从较易观察的角度和较易解决的角度出发，应该从等约束问题入手，先研究在等约束条件下如何求解优化问题。

等约束条件优化问题的一般形式为

$$\begin{aligned}&\boldsymbol{x}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\\ &\min f(x)\\ &\text{s.t.}\ h_v=0 \quad (v=1,2,\cdots,p<n)\end{aligned}$$

无约束条件下极值存在的条件是：

必要条件　$\nabla f(x^*)=0$

充分条件　$\nabla^2 f(x^*)>0$

我们从一个二元问题入手进行分析，得出结论后再扩展与推广到 n 维问题。设定一个二元函数：

目标函数　$f(x_1,x_2)$

约束条件　$h(x_1,x_2)=0$

现从极值条件开始，如有极值点，必有

$$\frac{\partial f(x^*)}{\partial x_1}=\frac{\partial f(x^*)}{\partial x_2}=0$$

为了以后分析方便，将上式改写成全微分形式：

$$\mathrm{d}f=\left[\frac{\partial f(x^*)}{\partial x_1}\right]\mathrm{d}x_1+\left[\frac{\partial f(x^*)}{\partial x_2}\right]\mathrm{d}x_2=0$$

即

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1}=-\frac{\dfrac{\partial f(x^*)}{\partial x_1}}{\dfrac{\partial f(x^*)}{\partial x_2}} \qquad (4.2\text{-}24)$$

对于约束条件，同样有

$$\mathrm{d}h=\left[\frac{\partial h(x^*)}{\partial x_1}\right]\mathrm{d}x_1+\left[\frac{\partial h(x^*)}{\partial x_2}\right]\mathrm{d}x_2=0$$

满足约束条件 $h(x_1,x_2)=0$，同满足上式是等价的，做同样变换后，得到

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1}=-\frac{\dfrac{\partial h(x^*)}{\partial x_1}}{\dfrac{\partial h(x^*)}{\partial x_2}} \qquad (4.2\text{-}25)$$

同时满足式(4.2-24)、式(4.2-25) 两式，则得到等约束极值存在的必要条件，即

$$\frac{\dfrac{\partial f(x^*)}{\partial x_1}}{\dfrac{\partial f(x^*)}{\partial x_2}}=\frac{\dfrac{\partial h(x^*)}{\partial x_1}}{\dfrac{\partial h(x^*)}{\partial x_2}}=-\frac{\mathrm{d}x_2}{\mathrm{d}x_1}$$

简化后得到

$$\frac{\partial f(x^*)}{\partial x_1}\frac{\partial h(x^*)}{\partial x_2}=\frac{\partial h(x^*)}{\partial x_1}\frac{\partial f(x^*)}{\partial x_2}$$

即

$$\frac{\partial f(x^*)}{\partial x_1}\frac{\partial h(x^*)}{\partial x_2}-\frac{\partial h(x^*)}{\partial x_1}\frac{\partial f(x^*)}{\partial x_2}=0 \qquad (4.2\text{-}26)$$

式 (4.2-26) 本身是一个两维偏导数的混合式，不易分清问题，最好是能找出一种针对某一维（或分别对于两维）的偏导数方程，才能得出无约束条件下极值存在条件的形式。Lagrange 为了解决这一关键问题而引入了一个新的系数 λ，称为 Lagrange 乘子，又称拉格朗日乘子或拉氏乘子。

令 $\lambda=\dfrac{\partial f(x^*)}{\partial x_1}\Big/\dfrac{\partial h(x^*)}{\partial x_1}=\dfrac{\partial f(x^*)}{\partial x_2}\Big/\dfrac{\partial h(x^*)}{\partial x_2}$，用 $\dfrac{\partial h(x^*)}{\partial x_2}$ 除式 (4.2-26)，则得

$$\frac{\partial f(x^*)}{\partial x_1}-\lambda\frac{\partial h(x^*)}{\partial x_1}=0$$

同样，用 $\dfrac{\partial h(x^*)}{\partial x_1}$ 除式 (4.2-26) 并移位变号得

$$\frac{\partial f(x^*)}{\partial x_2}-\lambda\frac{\partial h(x^*)}{\partial x_2}=0$$

还有一个等约束条件 $h(x_1,x_2)=0$。

现在我们将上述几个有关的式子按顺序整理如下：

$$\frac{\partial f(x^*)}{\partial x_1}-\lambda\frac{\partial h(x^*)}{\partial x_1}=0$$

$$\frac{\partial f(x^*)}{\partial x_2}-\lambda\frac{\partial h(x^*)}{\partial x_2}=0$$

$$h(x_1,x_2)=0$$

按着极值条件的第一条——必要条件，应当是梯度矩阵为0，即该函数对各维的偏导数为0。从上列三式中的前二式可以看出这一被求偏导数的函数至少要有 $f(x)$ 和 $\lambda h(x)$ 两个部分，第二部分在对 x_1 及 x_2 的两次求偏导数中都有系数 λ。从第三个式子中看到所设想的函数既没有对 x_1，亦没有对 x_2 的任何偏导，只有 $h(x_1,x_2)$ 的原型，而且消掉了系数 λ，这就使我们自然想到这可能是对 λ 求偏导数，因此可将 λ 看作第三个自变量，问题就迎刃而解了。

根据这一构想，设想出这一新的函数必然具有如下形式：

$$l(x_1,x_2,\lambda)=f(x_1,x_2)-\lambda h(x_1,x_2)$$

把 $l(x_1,x_2,\lambda)$ 看成是无约束条件下的某一要寻找极值的新函数，则其极值存在的必要条件为

$$\nabla l(x_1,x_2,\lambda)=0$$

即

$$\frac{\partial l(x,\lambda)}{\partial x_1}=\frac{\partial f(x^*)}{\partial x_1}-\lambda\frac{\partial h(x^*)}{\partial x_1}=0$$

$$\frac{\partial l(x,\lambda)}{\partial x_2}=\frac{\partial f(x^*)}{\partial x_2}-\lambda\frac{\partial h(x^*)}{\partial x_2}=0$$

$$\frac{\partial l(x,\lambda)}{\partial \lambda}=-h(x_1,x_2)=0$$

这样我们在引入拉氏乘子 λ 之后，将一个在等式约束 $h(x_1,x_2)=0$ 下求 $f(x_1,x_2)$ 的极值问题转化为一个在无约束条件下求一个新函数 $l(x_1,x_2,\lambda)$ 的极值问题。当其梯度 $\nabla l=0$ 这一必要条件存在时，原函数极值存在的条件及约束条件也均可得到满足，这是由最后的条件向前追溯必然会得到的结论。法国著名数学家 Lagrange 首先提出来的这一重大成果是他对人类数学科学的重大贡献，我们称这样一种约束条件下函数极值的方法为 Lagrange 乘子法。解这一问题的关键是引入系数 λ，它表示在极值点附近目标函数的变化与约束函数的变化的比率关系。

Lagrange 定理可规整叙述如下：

若 x^* 是 $\min\limits_{x\in R^n} f(x)$

s.t. $h_v(x)=0\quad(v=0,1,2,\cdots,p<n)$

的最优解，则存在实数 $\lambda_1,\lambda_2,\cdots,\lambda_p$，使得

$$\nabla f(x^*)-\sum_{v=1}^{p}\lambda_v\nabla h_v(x^*)=0$$

在具有不等式约束的问题中，如果最优点处在容许集内部，则所有约束都是不起作用的，这时最优性条件便转化为无约束的简单情况，即 $\nabla f(x^*)=0$。如果最优点处在容许集边界上，那么在最优点处必有某些约束变为起作用的约束，某些不等式约束变成等式约束，既然是等式约束，Lagrange 定理当然适用。Kuhn 和 Tucker 两位学者根据 Lagrange 定理的基本原理，考虑了不等式约束在求优化解中的作用，将其推广应用于不等式约束条件下的优化问题，提出了相应的定理，称为 Kuhn－Tucker 定理，叙述如下。

若 x^* 是 $\min\limits_{x\in R^n} f(x)$

s.t. $g_u(x)\geqslant 0\quad(u=1,2,\cdots,m)$

的最优解，则存在非负实数 μ_i（其中 $i\in I$），使

$$\nabla f(x^*)-\sum_{i\in I}\mu_i(x^*)=0$$

式中，$I=\{i\mid g_i(x^*)=0,i=1,2,\cdots,m\}$；$\mu_i$ 为广义 Lagrange 乘子。

Lagrange 理论至少是在理论上较圆满地解决了在约束条件下求优化解问题。但正如以前一样，在理论上解决问题并

不等于实际解决问题。理论上的思路提供了解决实际技术问题的线索和途径，由此便产生了惩罚函数法。

4.2　外部惩罚函数法（外点法）

(1) 惩罚函数法的提出和构成

对等式约束极小化问题，在目标函数和约束函数满足一定条件时，是能够用Lagrange函数求出极小点的。对于不等式约束，采用概念及方法推广的类比方法，引出了Kuhn－Tucker定理，也在原理上得到了解决。但是，用拉氏函数法来求解约束条件的优化问题是非常麻烦的，实际工程上难以直接应用，因而必须寻找更实际的且又可靠的方法。惩罚函数法便是在拉氏定理的原理指导下发展起来的一类重要的算法。

同样以二元问题为例：

$$\min f(\boldsymbol{x}) = x_1^2 + x_2^2$$
$$\text{s.t.}\ x_1 + x_2 - 2 = 0$$

这是一个等式约束问题，容许集 D 为一直线 $x_1 + x_2 = 2$，优化点为 $\boldsymbol{x}^* = [1,1]^{\mathrm{T}}$。

我们按拉氏函数的原理组成一个新的函数，然后按无约束条件的优化问题来求解它的优化点。如果这一新函数不能完全代替它得到真正的优化点，可以一次次地逼近。根据拉氏定理，上述简单二元问题的新函数组成大致如下：

新函数＝原函数＋系数×约束函数

此外，还应该有一个基本要求：在满足约束条件时，应使这一新函数具有原型，即当满足 $x_1 + x_2 - 2 = 0$ 约束方程时，新函数应有 $\phi(x,\gamma) = f(x) = f(x_1,x_2)$；在不满足约束条件，即 $x_1 + x_2 - 2 \neq 0$ 时，必须使 $f(x)$ 不可能有极小点，而最简单的办法就是使新函数 $\phi(x,\gamma) = \infty$。但其中第一项为原函数 $f(x)$，不会有任何可能达到 ∞，所以只有使第二项实现趋于 ∞ 的想法。第二项中当 $x_1 + x_2 - 2 \neq 0$，也只有使系数 γ 趋于 ∞ 才可能实现，因为第一项 $x_1 + x_2 - 2 \neq 0$ 已有定值。

根据上述构想，令

$$f(x_1,x_2) = \begin{cases} x_1^2 + x_2^2 & 当\ x_1 + x_2 = 2\ 时 \\ \infty & 当\ x_1 + x_2 \neq 2\ 时 \end{cases}$$

就是说，如满足约束条件，使 $f(\boldsymbol{x})$ 保持原型；如不满足约束条件，使 $f(\boldsymbol{x})$ 不可能有极小点，而使它的值达到无穷大。为此，新函数构造如下：

$$\phi(x_1,x_2,\gamma) = x_1^2 + x_2^2 + \gamma(x_1 + x_2 - 2)^2$$

式中，γ 是一个数值很大的正数，而约束函数的平方 $(x_1 + x_2 - 2)^2$ 保证第二项恒为正数。当约束条件满足时，$x_1 + x_2 - 2 = 0$，第二项全部消失；当不满足时，由于 γ 很大，趋向于 ∞，整个函数 $\phi \to \infty$。可见，上述新函数能够比较理想地实现上述构想。迭代点 x 离约束边界愈远，约束函数值愈大，第二项便愈大，自然也就无法求出极值点。

试解一下上述函数：

$$\frac{\partial \phi}{\partial x_1} = 2x_1 + 2\gamma(x_1 + x_2 - 2) = 0$$

$$\frac{\partial \phi}{\partial x_2} = 2x_2 + 2\gamma(x_1 + x_2 - 2) = 0$$

解出 $x_1^* = x_2^* = \dfrac{2\gamma}{2\gamma + 1}$，当 $\gamma \to \infty$，可得 $x_1^* = x_2^* = 1$。

使 γ 由某一起始值开始，一次次地增大，直到趋于 ∞，ϕ 函数的极值点就逐次逼近 $f(x_1,x_2)$ 的极小点。

由上述惩罚函数法的提出和形成过程可以看到，它是建立在两个基础上：一是建立在Lagrange函数理论基础上，可以把约束优化问题转化为无约束优化问题及该无约束优化函数是如何组成的；二是建立在惩罚函数的策略思想基础之上的，如果满足约束条件便保持原函数，如果不满足约束条件则施以惩罚而使之不可能有优化解。惩罚函数法的产生是拉氏定理的基本理论和惩罚策略思想相结合的结果。

(2) 外部惩罚函数算法的构成

对于等式和不等式约束条件的处理，在同一原则指导下罚函数具有不同的形式。以下分三种情况分述如下。

1) 对于等式约束下的 n 元函数　把上述构成原理推广到 n 元函数后会出现一系列的系数 γ。为了简化问题，较快地看出解优化极值存在的方向，将各系数 γ_1，γ_2，…，γ_n 均取为同一系数，这样新函数将具有以下形式：

$$\min f(\boldsymbol{x}),\ \boldsymbol{x} = [x_1,x_2,\cdots,x_n]^{\mathrm{T}} \qquad x \in R^n$$
$$\text{s.t.}\ h_v(\boldsymbol{x}) = 0, v = 0,1,2,\cdots,p < n$$
$$\phi(\boldsymbol{x},\gamma) = f(\boldsymbol{x}) + \gamma \sum_{v=1}^{p} [h_v(\boldsymbol{x})]^2$$

式中，$\phi(\boldsymbol{x},\gamma)$ 为增广目标函数；$f(\boldsymbol{x})$ 为原函数；γ 为惩罚因子；$\gamma \sum_{v=1}^{p} [h_v(\boldsymbol{x})]^2$ 为惩罚项或惩罚函数。

由最后一式可见，惩罚项是由惩罚因子 γ 和各约束函数的平方项构成的，因而必须把所有的约束函数全部列入，惩罚项的个数与等约束方程的个数是一样的。

2) 对于不等式约束条件下的 n 元函数　推广到不等式约束条件下的 n 元函数求解优化点问题：

$$\min f(\boldsymbol{x}),\ \boldsymbol{x} = [x_1,x_2,\cdots,x_n]^{\mathrm{T}} \qquad x \in R^n$$
$$\text{s.t.}\ g_u(x) \geqslant 0 \qquad (u = 1,2,\cdots,m)$$

其增广目标函数将具有新的形式，引入阶跃函数的新概念。

构成增广目标函数：

$$\phi(\boldsymbol{x},\gamma) = f(\boldsymbol{x}) + \gamma \sum_{u=1}^{m} [g_u(\boldsymbol{x})]^2 u[g_u(\boldsymbol{x})]$$

式中，$u(t) = \begin{cases} 0 & 当\ t \geqslant 0 \\ 1 & 当\ t < 0 \end{cases}$

当自变量 $t \geqslant 0$ 时，$u(t) = 0$，即整个惩罚项均不存在，当满足约束条件 $g_u(x) \geqslant 0$ 时，增广目标函数中只保留原目标函数一项；当 $t < 0$ 时，$u(t) = 1$，全部惩罚项起作用，使增广目标函数值趋于 ∞，不可能有极小值。

3) 对于具有等式和不等式约束的 n 元函数　对于同时具有等式约束与不等式约束问题，可综合采取两种惩罚策略，新的增广目标函数为

$$\phi(\boldsymbol{x},\gamma) = f(\boldsymbol{x}) + \gamma\alpha(\boldsymbol{x})$$

式中右边第二项为惩罚函数，它体现所谓两种惩罚策略，由上述两种形式构成：

$$\alpha(\boldsymbol{x}) = \sum_{v=1}^{p} [h_v(\boldsymbol{x})]^2 + \sum_{u=1}^{m} [g_u(\boldsymbol{x})]^2 u[g_u(\boldsymbol{x})]$$

该式称为约束问题的惩罚函数。显然有

$$\alpha(\boldsymbol{x}) = \begin{cases} = 0 & 当\ \boldsymbol{x} \in \boldsymbol{D} \\ > 0 & 当\ \boldsymbol{x} \notin \boldsymbol{D} \end{cases}$$

$\boldsymbol{D} = \{(\boldsymbol{x}) \mid g_u(\boldsymbol{x}) \geqslant 0, u = 1,2,\cdots,m; h_v(\boldsymbol{x}) = 0, v = 1,2,\cdots,p\}$ 是容许集。

惩罚函数法的基本策略是根据约束特点构造某种形式的“惩罚”函数，然后把它迭加到目标函数上去组成增广目标函数，使得约束问题转化为一系列无约束问题来求解。“惩罚”是一种策略，它对于在求解过程中企图违反约束条件的那些迭代点施以极大的“惩罚”，赋以很大的目标函数值，迫使迭代点无限地向容许集靠近。惩罚函数法原理易懂，程序简单，使用有效，但在向最优点靠近时，增广目标函数的性质可能变“坏”，致使方法失效。但惩罚函数法仍不失为求解约束最优化问题的一个重要而基本的解法，具有重要优

点，在多数情况下都是有效的。

4.3 内部惩罚函数法（内点法）

所谓内部惩罚函数法，是指在迭代求解的整个过程中，初始点 $x^{(0)}$、迭代点 $x^{(k)}$ 及最终的优化点 x^* 都始终在容许集内，迭代点永远在可行域内移动。对接近容许集边界的点施加惩罚，直至在边界上给予达到∞的惩罚。这就好比在边界上筑起一道很高的围墙，阻碍迭代点穿越边界，把它们封死在容许集内。这样容许集内部就构成了非空集，否则失去惩罚意义。因此这种算法只适应于不等式约束，因为等式约束下的容许集空间中不存在内点。

内点法与外点法一样，完全建立在同一"惩罚"政策思想之上，差别只在于迭代点相对于边界的位置是在边界之外还是在边界之内。组成增广目标函数时各项的构成原则以及基本思路都与外点法完全一致。只是由于对不等式约束，迭代点始终应在域内，只有在极端情况下才能靠近边界，因而不可能在边界外移动。

对于一般不等式约束条件下的优化问题：

$$\min f(\boldsymbol{x}),\ \boldsymbol{x}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\quad x\in R^n$$

$$\text{s.t.}\quad g_u(\boldsymbol{x})\geqslant 0\quad(u=1,2,\cdots,m)$$

$$\text{容许集 } D=\{x\mid g_u(x)\geqslant 0,u=1,2,\cdots,m\}$$

为实现上述内部惩罚策略的构想，必须在约束函数接近边界即 $g_u(\boldsymbol{x})$ 趋近于零时使函数值趋向∞，加以惩罚，使之"不敢"靠近而只能在域内移动。

对于 $g_u(\boldsymbol{x})$ 趋于零，若需给予其值为∞的惩罚，增广目标函数的构成可具有如下形式：

$$\phi(\boldsymbol{x},\gamma)=f(\boldsymbol{x})+\gamma\sum_{u=1}^{m}\frac{1}{g_u(\boldsymbol{x})}$$

当 $g_u(\boldsymbol{x})=0$ 时，$\frac{1}{g_u(\boldsymbol{x})}=\infty,\phi(x,\gamma)\to\infty$。

令 $\beta(\boldsymbol{x})=\sum_{u=1}^{m}\frac{1}{g_u(x)}$，则增广目标函数简写为

$$\phi(\boldsymbol{x},\gamma)=f(\boldsymbol{x})+\gamma\beta(\boldsymbol{x})$$

式中，$\gamma\beta(x)$ 为惩罚项。

当迭代点 $x^{(k)}$ 从容许集 $\boldsymbol{D}$ 内部某一点沿着某一方向向边界接近时，至少有一个约束函数 $g_u(\boldsymbol{x})$ 将趋于零，因而在 $\sum_{u=1}^{m}\frac{1}{g_u(\boldsymbol{x})}$ 中至少有一项趋于∞，使 $\beta(\boldsymbol{x})$ 趋于∞，从而达到惩罚的目的。换言之，即迭代点不可能靠近边界。

根据上述思想，内点惩罚函数可以有以下三种构造形式：

1) $\beta(x)=\sum_{u=1}^{m}\frac{1}{[g_u(\boldsymbol{x})^2]}$

2) $\beta(x)=\sum_{u=1}^{m}\ln[g_u(\boldsymbol{x})]$

3) $\beta(x)=\sum_{n=1}^{m}\frac{1}{g_u(\boldsymbol{x})u[-g_u(\boldsymbol{x})]}$

内点法与外点法求解过程的差别较大。内点法是在集内，极小点由上向下移动到最后极限；外点法的迭代点则始终是在集外，极小点是从下部向上移动，最后达到极限点。两种不同方法的惩罚因子 γ 也完全不同。

5 多目标优化问题的解法

多目标优化问题的解法从大的方面说，可以分为间接解法和直接解法两大类。间接解法又分为转化成一个单目标问题的解法，转化成多个单目标问题的解法和非统一模型的解法。其中，转化成一个单目标问题的解法应用较多。

所谓转化成一个单目标问题的解法是指：首先，设法将原多目标问题转化成一个单目标问题；然后，再利用非线性规划的有关算法求解此单目标问题，并将所求得的解作为该多目标优化问题的最优解。该方法的关键是要保证所构造的单目标问题的最优解是多目标问题的有效解或弱有效解。

下面简要介绍两种典型的多目标优化问题转化成一个目标优化问题求解方法——主要目标法和加权函数法的基本原理。

(1) 主要目标法

主要目标法的基本思想是：在多目标问题中，根据问题的实际情况，确定一个目标为主要目标，而把其余目标作为次要目标，并且根据决策者的经验，选取一定的界限值。这样就可以将次要目标作为约束条件来处理，于是就将原多目标问题转化成一个在新的约束条件下，求主要目标的单目标最优化问题。

主要目标方法简单易行，它保证在次要目标允许取值的条件下，求出主要目标尽可能好的值，因此对实际问题常常很适用。

(2) 加权函数法

在加权函数法中，较常用的是线性加权和法，其次还有指数权函数法和平方加权函数法等，下面简要介绍上述各种方法的函数形式。

1) 线性加权和法　所谓线性加权和法就是按照多个目标的重要程度，分别乘以一组权系数，然后相加作为目标函数，再对此目标函数在多目标问题的约束集合上求最优解。即构造如下的单目标问题：

$$\min_{x\in R}y(x)=\sum_{j=1}^{p}\lambda_j f_j(x)=\lambda^{\mathrm{T}}f(\boldsymbol{x})$$

求此单目标问题的最优解，并将它称为多目标问题在线性加权和意义下的最优解。

线性加权和法是一种常用的算法，而且在理论上有重要意义。在线性加权和法中，权系数的确定是十分重要的。一般说来，目标函数重要的，相应的权系数就要给得大些（当然最大不能超过 1）；而不很重要的目标函数，其相应的权系数就要给得小些（甚至可为 0）。

2) 指数权函数法　指数权函数的具体表达式为

$$\min_{x\in R}y(x)=\exp\left[\sum_{j=1}^{p}\lambda_j f_j(\boldsymbol{x})\right]$$

3) 平方加权函数法　若 $f(x)\geqslant 0$，或是求各目标函数的零点，则可用平方权函数：

$$\min_{x\in R}y(x)=\sum_{j=1}^{p}\lambda_j[f_j(\boldsymbol{x})]^2$$

编写：王广春（山东大学）

第 3 章　基于有限元模拟技术的反向模拟式设计

在实际锻造生产中，大多数锻件的几何形状是相当复杂的，这些锻件当用形状规则的初始毛坯（圆坯或方坯）直接成形（一次成形）时，终锻件往往会出现严重的质量问题，如充不满、折叠、裂纹等缺陷。为保证获得高质量的锻件，在进行终锻前，必须进行一次或多次预锻，即预成形（preforming），预成形的目的是为了合理地分配金属，使金属在终锻时完全充满模腔，且又不会出现缺陷。此外，还可以避免一次成形时变形力太大的问题，从而延长模具和设备寿命。预成形件的形状及预成形工序的数目是获得理想终锻件的重要因素，因此，预成形设计包括确定预成形工步数目、预成形件和预成形模具形状以及合理的毛坯尺寸。在现行生产中多数情况下，人们仍是凭实际经验来进行类比设计。

金属成形中的预成形设计（preform design）指的是设计工件的初始或中间的形状与尺寸，这种形状的毛坯经有关变形后，能够成功地变形为性能合乎要求、无缺陷、料耗少的产品，从而降低生产成本，提高材料利用率和锻件质量。在各种金属成形工艺中，如模锻、缩口、轧制及板料成形都会遇到预成形设计的问题。设计一个最佳预成形件要求同时确定最佳工艺条件，本章介绍的是在给定工艺条件下如何确定最佳预成形件及其预成形模具的形状。

进行锻造工艺设计，掌握金属流动规律是关键，从目前来看，适合于正、反向模拟金属流动规律的数值方法有上限元技术（Upper Bound Elemental Technique，UBET）和有限元法，模拟技术有正向模拟技术（forward simulation）和反向模拟技术（backward or reverse simulation）或反向跟踪技术（back - tracing）。如果模拟是从已知模具参数和初始坯料形状尺寸的条件下开始进行，模拟自初始坯料到终锻件整个成形过程中的金属流动规律，则这种模拟称为正向模拟或加载模拟；若不是从可变的初始条件出发，而是从给定的最终锻件形状和工艺条件出发，沿与成形过程相反的方向模拟变形过程中任意中间时刻的锻件形状和尺寸，以及最终确定成形终锻件所需的初始或中间毛坯形状与尺寸，则这种模拟方式称之为反向模拟或反向跟踪，也称为复原模拟。具体讲，反向模拟就是从终锻件开始，假想地使上、下模彼此逐步分离，材料随之脱离模具，逆向成形为一定形状的中间或初始毛坯的过程。

正向模拟与实际锻造生产过程相似，即由毛坯生产出锻件；而反向模拟过程则与生产中技术人员进行工艺设计的过程相似，即由已知的终锻件设计预成形件、毛坯形状以及预锻模具形状，但实际设计一般以经验为主，而反向模拟式设计却有一定理论可循。金属成形过程的正向模拟对于实际工艺设计具有很大的指导和验证作用，但这种作用是被动的，即当我们根据经验数据或资料设计出工艺方案后，采用正向模拟技术验证设计方案是否可行，若不够理想，则修改或调整某些工艺参数。而反向模拟技术对于成形工艺设计却有积极的指导作用。通过反向模拟，反演出预成形件、预锻模具及初始坯料的形状与尺寸，从而得到令人满意的设计方案。因此，反向模拟式设计对于实际生产具有重要的意义和作用。尽管反向模拟式设计方法仍存在许多尚未开发的难点，但可以预见锻造工艺的反向模拟式设计是提高锻造工艺设计科学水平的重要途径之一。

本章将介绍基于刚（黏）塑性有限元法的反向模拟技术，给出在实施反向模拟过程中所遇到的若干技术问题及其解决途径，叙述反向模拟系统的构造方法，指出其中的几个难点和未来的研究方向，最后介绍几个应用实例。

1　有限元反向模拟技术的原理

基于刚（黏）塑性有限元法的反向模拟技术在国内外仍是一个令人感兴趣、又有待于进一步研究开发的领域。S. Kobayashi 等应用有限元法对缩口、轧制和圆柱体自由锻粗过程进行了预成形设计，提出了基于有限元法的反向模拟技术，并指出有限元法在金属成形预成形设计问题的各个方面都具有重要作用，反向模拟技术在锻造工艺设计中的应用是对广大金属成形研究工作者的未来挑战。国内也开展了这方面的研究工作。

图 4.3-1 表示反向模拟过程示意图，设终锻时刻为 t_0，锻件形状为 X_0，其中 X_0 为有限元节点坐标的集合，取一个反向时间增量（$-\Delta t$）进行逆向计算，即以 t_0 时刻的形状 X_0 为参考状态，求得此时刻的收敛速度场 V_0，由式（4.3-1）得到 t_{0-1}时刻的预成形件形状。然后再以 X_{0-1}为参考状态，正向模拟求出此时刻的收敛速度场 V'_{0-1}，由式（4.3-2）再求出 t_0 时刻的正向解。

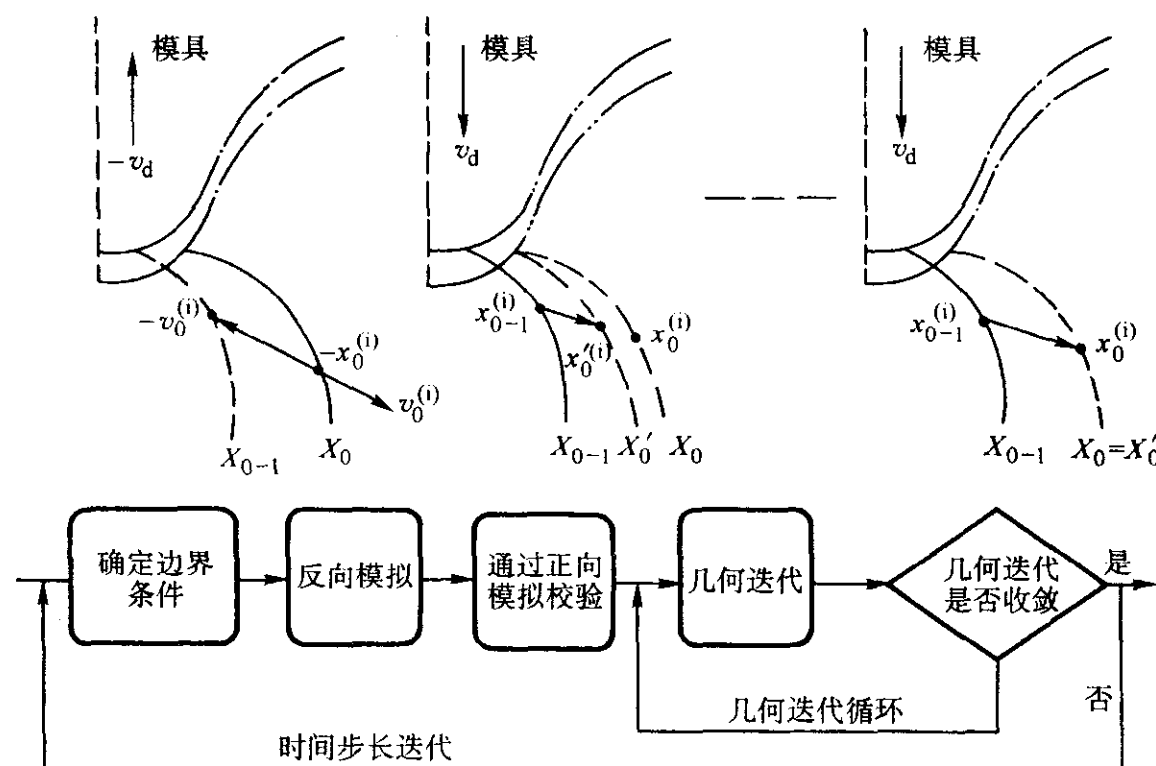


图 4.3-1　反向模拟技术几何迭代过程示意图

$$\left.\begin{aligned} X_{0-1} &= X_0 - V_0\Delta t \\ t_{0-1} &= t_0 - \Delta t \end{aligned}\right\} \tag{4.3-1}$$

$$X'_0 = X_{0-1} + V'_{0-1}\Delta t \tag{4.3-2}$$

一般地，$X'_0 \neq X_0$，即反向模拟与正向模拟计算结果并非互逆，这是因为：

1）在计算时进行了简化，对时间采用了差分代替微分关系；

2）在 t_{0-1}和 t_0 时刻的边界条件和收敛速度场不尽相同。

不难看出，只有 X_0 与 X'_0 重合，才能保证预成形件为 X_{0-1}时，正向模拟所得到的锻件形状满足给定的要求，即终锻件形状为 X_0。欲使二者重合，必须反复调整 t_{0-1}时刻的节点位置（预成形件形状），从而使 t_{0-1}时刻的预成形件正向模拟一步后所达到的位置（或形状）恰好与 t_0 时刻的锻件形状 X_0 重合。这种调整或修正过程需要通过多次迭代对所有节点进行，直至满足 $X_0 = X'_0$ 的条件。在计算中还应注意到边界条件受模具约束，在调整 X_{0-1}时，边界条件可能由于 X_{0-1}的修正而变化，这就出现了在反向模拟过程中如何确定预成形件边界条件的问题，即边界节点与模具相脱离或接触的问题。

一般而论，直接调整预成形件形状 X_{0-1}使 X'_0 与 X_0 重合是很困难的，事实上，调整 X_{0-1}的过程完全可以转变为调整 t_{0-1}时刻、以 X_{0-1}为参考状态的速度场，图 4.3-2 给出

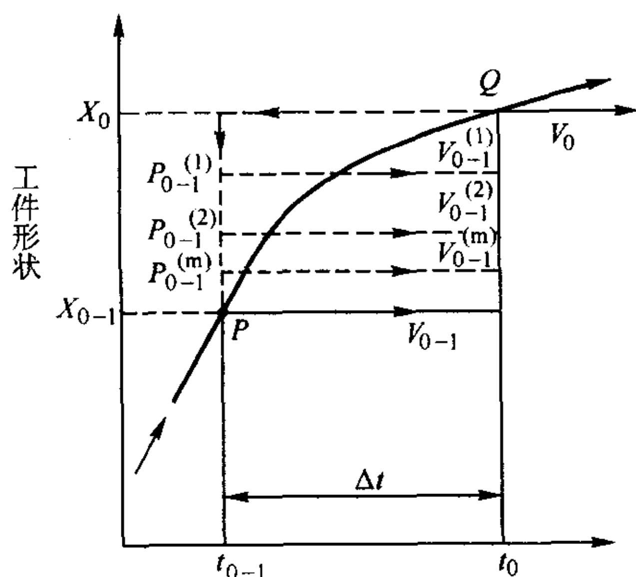


图 4.3-2　速度场调整过程示意图

了这种速度场调整过程的示意图，其具体方法与步骤如下。

1）取 X_0 处（Q 点，Q 为节点坐标集合）的收敛速度场 V_0，确定 X_{0-1}（P 点，P 为节点坐标集合）的第一次估计值：$P^{(1)} = X_0 - V_0\Delta t$。

2）根据 $P^{(1)}$ 点处的形状计算 t_{0-1}时刻的正向加载解 $V_{0-1}^{(1)}$，求出 $Q^{(1)} = P^{(1)} + V_{0-1}^{(1)}\Delta t$，若 Q 与 $Q^{(1)}$ 相当接近，则认为 $Q^{(1)}$ 是合乎要求的反向模拟解，然后以 $Q^{(1)}$ 点为参考状态进行下一个时间步的反向模拟。

3）若 Q 与 $Q^{(1)}$ 不足以接近，则再由 X_0 及 $V_{0-1}^{(1)}$ 计算 P 的第二次估计值 $P^{(2)} = X_0 - V_{0-1}^{(1)}\Delta t$，再由 $P^{(2)}$ 计算 $P^{(2)}$ 处的正向加载解 $V_{0-1}^{(2)}$，即计算 t_0 时刻的第二个正向估计值 $Q^{(2)} = P^{(2)} + V_{0-1}^{(2)}\Delta t$，再比较 Q 与 $Q^{(2)}$。

4）继续迭代和比较，直至 $Q^{(n)}$ 与 Q 足够接近为止，其中，n 为几何迭代次数。

迭代公式为

$$\left.\begin{aligned} P^{(n)} &= X_0 - V_{0-1}^{n-1}\Delta t \\ Q^{(n)} &= P^{(n)} - V_{0-1}^{n}\Delta t \end{aligned}\right\} \tag{4.3-3}$$

下面进一步说明基于有限元法的反向模拟技术。

在有限元离散化后，在 t 时刻与塑性变形相关联的边值问题的解满足一个非线性耦合代数方程组

$$\boldsymbol{K}(\boldsymbol{X}, \bar{\boldsymbol{\varepsilon}}, \dot{\bar{\boldsymbol{\varepsilon}}}, \boldsymbol{T})\boldsymbol{V} - \boldsymbol{F}(\boldsymbol{X}) = 0 \tag{4.3-4}$$

并具有力和速度边界条件。

式中，$\boldsymbol{K}$ 为刚度矩阵，为 $\boldsymbol{X}$，$\bar{\varepsilon}$，$\dot{\bar{\varepsilon}}$，T 的函数；$\boldsymbol{F}$ 为载荷向量矩阵；$\boldsymbol{X}$ 为变形体所有节点的坐标向量；$\boldsymbol{V}$ 为变形体所有节点的速度向量；$\bar{\boldsymbol{\varepsilon}}$为变形体所有单元积分点处的等效应变向量；$\dot{\bar{\boldsymbol{\varepsilon}}}$ 为变形体所有单元积分点处的等效应变率向量；$\boldsymbol{T}$ 为给定的变形体所有节点的温度向量。

由式（4.3-4）中，设 t 时刻的 $\boldsymbol{X}$ 和$\bar{\boldsymbol{\varepsilon}}$已知，并设$\dot{\bar{\boldsymbol{\varepsilon}}}$是 $\boldsymbol{V}$ 的函数，在正向模拟（$t \geq t_0$）中，应用 $t = t_0$ 时刻式（4.3-4）的解 $\boldsymbol{V}_0$ 作为下一步 $t = t_{0+1} = t_0 + \Delta t$ 时刻的解，其显式为

$$\left.\begin{aligned} \boldsymbol{X}_{0+1} &= \boldsymbol{X} + \boldsymbol{V}_0\Delta t \\ \bar{\boldsymbol{\varepsilon}}_{0+1} &= \bar{\boldsymbol{\varepsilon}}_0 + \dot{\bar{\boldsymbol{\varepsilon}}}\Delta t \end{aligned}\right\} \tag{4.3-5}$$

在反向模拟时（$t \leq t_0$），在 $t = t_{0-1} = t_0 - \Delta t$ 时刻，$\boldsymbol{X}_{0-1}$，$\dot{\bar{\boldsymbol{\varepsilon}}}_{0-1}$应当满足下列条件：

$$\left.\begin{aligned} \boldsymbol{X}_0 - \boldsymbol{X}_{0-1} &= \boldsymbol{V}_0\Delta t \\ \bar{\boldsymbol{\varepsilon}}_0 - \bar{\boldsymbol{\varepsilon}}_{0-1} &= \dot{\bar{\boldsymbol{\varepsilon}}}_{0-1}\Delta t \end{aligned}\right\} \tag{4.3-6}$$

对于给定的时间增量 Δt，在 t_{0-1}时刻将式（4.3-6）代入式（4.3-4）得

$$\boldsymbol{K}(\boldsymbol{K}_{0-1}, \dot{\bar{\boldsymbol{\varepsilon}}}_{0-1}, \boldsymbol{X}_0, \bar{\boldsymbol{\varepsilon}}_0, \boldsymbol{T})\boldsymbol{V}_{0-1} - \boldsymbol{F}(\boldsymbol{V}, \boldsymbol{X}_0) = 0 \tag{4.3-7}$$

式中，$\boldsymbol{X}_0$，$\bar{\boldsymbol{\varepsilon}}_0$ 已知，$\dot{\bar{\boldsymbol{\varepsilon}}}_{0-1}$是 $\boldsymbol{V}_{0-1}$的函数，$\dot{\bar{\boldsymbol{\varepsilon}}}_{0-1} = \dot{\bar{\boldsymbol{\varepsilon}}}(\boldsymbol{V}_{0-1})$。

由上式可知，如果 t_{0-1}时刻的边界条件被指定，则式（4.3-7）完全可以求解，重复计算 t_{0-1}，t_{0-2}，…，t_{0-n}时刻的解 $\boldsymbol{V}_{0-1}$，$\boldsymbol{V}_{0-2}$，…，$\boldsymbol{V}_{0-n}$，即可完成整个成形过程的反向模拟，获得各个时刻的预成形件及其有关参量的分布。

2　反向模拟关键应用技术

上节介绍了基于有限元方法的反向模拟技术的基本原理，在将该模拟技术应用于金属塑性成形过程的预成形设计时，还遇到一些关键应用技术，例如加工硬化问题的处理、预锻模具形状的设计、边界条件的确定准则等，这些关键技术直接影响到反向模拟设计方法的实际应用，下面着重介绍几个关键技术。

2.1　反向模拟与预成形设计的步骤

基于刚（黏）塑性有限元方法的反向模拟过程与正向模拟过程一样采用步进方法进行，在当前模拟步内，采用收敛的节点速度刷新有限元网格或工件的形状，节点速度是通过在施加相应的边界条件下求解非线性方程组而获得，典型的求解方法为 Newton - Raphson 迭代方法，即摄动法。反向模拟与预成形设计的基本步骤如下：

1）利用终锻件的几何形状、终锻模具形状和工艺参数等建立反向模拟的初始有限元模型，包括边界节点与模具接触的条件等；

2）对于当前模拟步，确定反向模拟时边界节点与模具的接触或脱离条件，这通常由预先确定的边界接触节点脱离模具的准则来确定；

3）进行速度场迭代，求得收敛的速度场，在前一反向模拟步的基础上，应用收敛的速度场刷新工件形状和模具位置，完成该步的反演。然后，对该反向模拟步进行正向模拟验证，即从反向刷新的工件形状、模具位置、边界条件等开始，进行有限元正向模拟，求得收敛的速度场，并正向刷新工件的几何形状；

4）比较该模拟步正向验证模拟所得到的工件形状和该

模拟步反向模拟开始时的工件几何形状，如果二者几何形状足够接近，则该反向模拟步结束，并进行下一模拟步的反演。若二者几何形状不够接近，则重复步骤3）继续进行几何迭代，直至满足几何迭代的误差条件；

5）对所有模拟步重复步骤1）~4），直至整个反向模拟过程结束。原则上讲，当工件的边界节点全部脱离模具时，反向模拟过程结束。

2.2　加工硬化现象的处理

由式（4.3-7）看出，在反向模拟开始时，必须预知终锻件形状 X_0 和等效应变分布 $\bar{\varepsilon}_0$，其中 X_0 已知，即为终锻件几何形状，而 $\bar{\varepsilon}_0$ 未知，对于理想刚塑性材料 $\bar{\sigma}$ = 常量，与 $\bar{\varepsilon}_0$ 无关，因此，不必预先知道锻件内的等效应变分布。然而，对于加工硬化材料的反向模拟，则必须预先知道终锻件内的等效应变分布。一般地，终锻件内的等效应变分布取决于预成形件的形状，而预成形件的形状事先未知，它是反向模拟所寻求的目标。因此，应通过某一途径来获得终锻件内的应变分布或近似分布，具体方法是：

1）首先应用无加工硬化的材料进行反向模拟，求出对应的预成形件或毛坯的形状与尺寸；

2）应用求出的无加工硬化材料的预成形件或毛坯对加工硬化材料进行正向模拟，求出对应的终锻件形状及其应变分布，该终锻件形状可能不同于实际锻件形状，但却非常接近指定的实际终锻件形状，然后根据实际锻件的形状修正正向模拟得出的终锻件形状，并保持相同的应变分布，这种应变分布可近似地作为加工硬化材料反向模拟开始时的应变分布 $\bar{\varepsilon}_0$，这样，即可反向模拟出加工硬化材料的预成形件形状。

在反向模拟过程中，工件内的等效应变称为残余等效应变（residual effective strain），第 n 步反向模拟结束时的残余应变为

$$\bar{\varepsilon}_{0-n} = \bar{\varepsilon}_{0-(n-1)} - \dot{\bar{\varepsilon}}_{0-n}\Delta t \tag{4.3-8}$$

严格来讲，当反向模拟结束时，反向模拟得出的毛坯内的残余应变应减少为零。

2.3　预锻模模腔形状设计

终锻模（型腔部分）的形状是根据终热锻件的形状及尺寸来设计的，同理，预锻模形状也可以按照预锻件的形状来设计。因此，预锻模模腔形状的设计实际上就转换为根据终锻件形状设计预锻件形状，只要能够求出合理的预锻件形状及尺寸，相应地就能设计出预锻模型腔形状，这也是进行反向模拟的目的之一。因此，可根据预锻件的形状设计出预锻模具的形状。

2.4　预锻件形状的选择及工序数目的确定

有限元反向模拟技术是一种步进式（或增量式）模拟技术，在整个模拟过程中，可以得到成形过程中任意时刻的锻件形状，必须进行合理取舍，留下最少和最必须的几个中间锻件做为预成形件。在正向模拟过程中，有一个确定的模拟终止条件，即上、下模打靠位置，正向模拟的结束是以上、下模之间的距离达到规定的数值为条件，或者以飞边达到一定的高度为标志。而反向模拟则是以上、下模完全脱离锻件，即全部边界接触节点成为自由节点为终止条件。若反向模拟终止时的毛坯形状接近于圆坯或方坯，则简化成圆坯或方坯，并保证体积不变；若反向模拟结束时，锻件形状仍很复杂，则可作为一个预成形坯，并继续以此预成形坯为初始状态进行反向模拟，从而求出一个形状较规则的初始毛坯，中间任意时刻的锻件形状均可作为预成形件，但为了减少成形工序数目，可适当选择一到两个中间锻件作为预成形件。在确定了预成形工序数目及相应的预成形形状后，再进行全过程正向模拟验证，观察是否能够得到给定的终锻件形状，是否出现缺陷以及飞边是否较小等，以确保预成形设计的正确性。若结果不够理想，则再进行适当调整。

2.5　反向模拟过程中动态边界条件的确定

反向模拟技术是有限元法在金属成形问题中的一个独特应用，这种技术已经应用于多种成形工艺过程的预成形设计，同时也揭示出在进一步开发这种技术中的关键问题之一是边界条件的处理。

在正向模拟中，非稳态成形过程中的边界条件是不断变化的，而且变化的方式又取决于预成形件的形状，实际上，金属充填模腔的过程就是其自由边界（边界节点）与模具内表面不断贴合（接触）的过程，随着边界节点的不断接触与脱离，边界条件随之发生变化。这种变化的边界条件在正向模拟中通过自由节点的触模及接触节点的脱模判断（参见文献［9］）是完全可以确定的。而在反向模拟时，却没有初始坯料或中间预成形件的边界条件，因为这正是我们进行反向模拟所寻求的目标。例如，对于锻造问题，反向模拟开始时，假设终锻件完全充满型腔，除飞边处的个别节点为自由节点外，其余边界节点均为与模具边界接触的节点。在反向模拟过程中，这些接触节点应随着反向模拟的不断进行，逐步地脱离模具成为自由节点。与有限元正向模拟不同，这种接触节点的脱离判断是无法单纯地依靠节点运动和模具形状或位置来确定的。因此，在进行反向模拟时，必须预先确定接触边界节点何时、何处、以何种顺序和方式脱离模具，即所谓的边界条件的控制准则（control criteria）。

在反向模拟过程中，沿最终锻件轮廓上的工件与模具接触面上的节点应当随上、下模的不断分离而逐步脱离模具，当反向模拟结束时，边界上的所有节点应成为自由节点，因此，从这种意义上讲，可以想像出三种边界条件的控制准则：

1）控制预成形件的几何形状；

2）控制工件内的变形；

3）边界接触节点的受力特征。

控制预成形件几何形状的准则是人为地规定反向模拟过程中边界接触节点脱模的时机与顺序。终锻件的形状、可以想象出的预成形件形状、正向模拟时边界节点的贴模过程和顺序以及实际锻造生产中所期望的最佳成形特点，均可作为确定边界接触节点脱模时机和顺序的参考。

对于控制工件内变形的准则，在实际锻造生产中，变形体的变形很不均匀，而我们所希望的是尽可能获得变形均匀的锻件。若假设初始终锻件内的应变分布均匀，则在反向模拟过程中，变形体内的应变（残余应变）分布就表征变形体内不同区域变形的大小及其规律，残余应变大的区域，变形较小；残余应变小的区域，变形较大。同时，残余应变分布又是模拟软件的输出结果之一。因此，可根据变形体内的残余应变分布来选择节点脱离模具的时机与顺序，以便使反向模拟出的预成形件或毛坯当正向模拟时所得到终锻件内的应变分布尽量均匀，从而得到性能均匀的终锻件。

根据边界接触节点的受力状态确定脱模规律。从边界节点的受力状态来看，若边界接触节点的节点力（或应力）大于零（拉力），则表明该节点将脱离模具，若节点沿模具表面法线方向的节点力（或应力）小于零（压力），则表明该节点仍与模具相接触。

此外，在反向模拟过程中，已经脱离模具的节点（边界

自由节点）也可能再次与模具接触，这种接触的判断方法与有限元正向模拟过程中节点的接触判断方法相同。

采用不同的边界条件控制准则，可求得不同的预成形件形状，这就为我们提供了灵活设计预成形件形状的可能性，例如，我们可以根据所要求的充填过程和变形规律来确定所期望的工艺方案。但是这种控制准则的不惟一性，给反向模拟设计实际应用带来了较大的困难，也较难实现反向模拟设计的自动化和通用化。因此，对于基于有限元的反向模拟预成形设计方法，其关键问题是如何确定合理的反向模拟过程中边界条件的变化，即动态边界条件的确定准则。目前，国内外还尚未有通用的边界条件控制准则。在本章后面的3~5节中将结合具体实例介绍几种边界条件的控制准则。

2.6 有限元反向模拟系统的构造

如图4.3-3所示为基于有限元方法的反向模拟程序流程框图。在反向模拟过程中，加载模拟（正向模拟）与反向模拟交替进行，无论是正向还是反向模拟都离不开在一定几何和工艺条件下求其收敛的速度场，即速度场的迭代过程。因此，反向模拟系统中的绝大部分模块与正向模拟系统相同，可以借用正向模拟系统中的大多数模块，不同之处大致有以下几点。

1）在反向模拟系统中需增加一个主程序模块，以控制反向模拟程序的运行。

2）确定动态边界条件的模块，该模块的编制主要基于边界条件的控制准则，即控制边界接触节点何时、何地、以何种顺序脱离模具。

3）增加几何迭代收敛判断模块和几何迭代次数循环语句，控制反向模拟的结果精度。几何迭代公式见式（4.3-3），式中 n 为几何迭代次数。收敛条件是使 $Q^{(n)}$ 与 Q 足够接近，具体来讲，可以采用下列几何收敛判断准则：

$$\left.\begin{aligned} G &= \| Q - Q^{(n)} \| / \| Q \| \leqslant G_0 \\ \| Q - Q^{(n)} \| &= \sqrt{\sum_{i=1}^{M} (x_i - x_i^{(n)})^2} \\ \| Q \| &= \sqrt{\sum_{i=1}^{M} x_i^2} \end{aligned}\right\} \tag{4.3-9}$$

式中，M 为节点坐标分量总数，对于二维问题 $M=2N$，N 为节点总数；x_i 为节点坐标分量；G 为几何误差范数；G_0 为几何误差极限，取为 $10^{-3} \sim 10^{-4} \times$ 行程步长。

显然，G_0 取得越小，反向模拟结果越准确，但计算效率会降低，这是因为，G_0 越小，几何迭代次数越多，而在每次迭代中，又要进行多次速度场的 Newton - Raphson 迭代，若几何迭代次数为 n，速度场迭代次数为 m，则进行一步反向模拟所需的速度场迭代次数为 $n \times m$。可见，反向模拟时速度场的迭代次数远远大于单纯正向模拟的速度场迭代次数，呈 n 倍关系。经有关计算表明，n 一般为2~4。

此外，步长越大，差分与微分的近似程度越差，因此，G_0 的选取也与步长大小有关，一般设定为线性关系。

反向模拟系统的可靠性与通用性，一方面取决于反向模拟系统本身的三个模块，即主控模块、边界条件控制模块和几何迭代收敛判断模块；另一方面又取决于正向模拟系统中与速度场 Newton - Raphson 迭代相关模块的可靠性与通用性。因此，反向模拟系统的构造必须建立在可靠的有限元正向模拟系统之上。图4.3-3所示的反向模拟系统结构框图中标有“*”者为正向模拟系统中的有关模块。

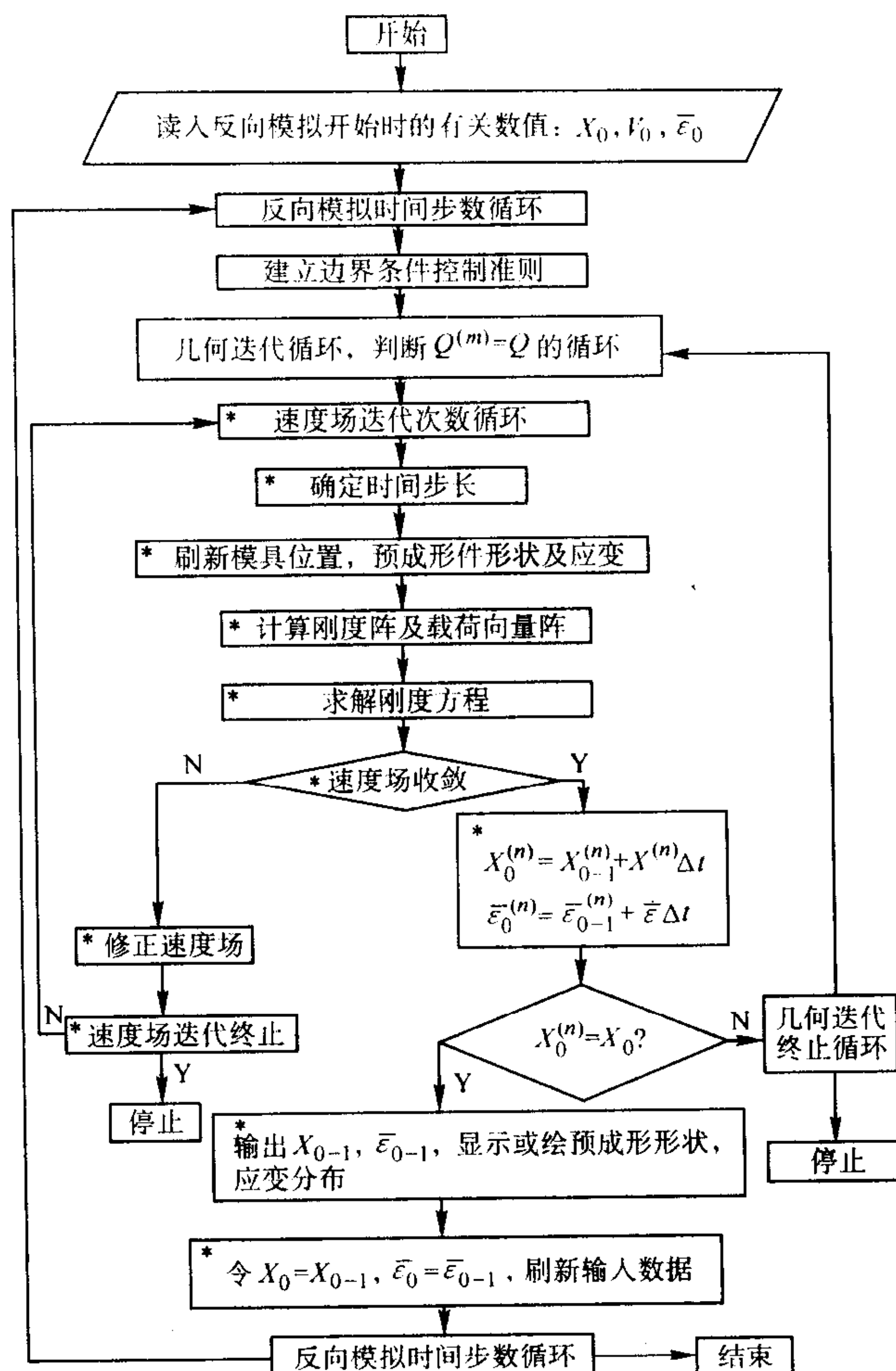


图4.3-3 基于有限元方法的反向模拟程序流程框图

3 缩口过程的反向模拟及预成形设计

3.1 室温缩口过程的预成形设计

缩口工艺是通过施加压力使得筒形件进入某种形状的模具，从而使筒形件开口端部的直径减小的成形工艺。在现行的缩口工艺方法中，缩口件的内部型腔成形精度较高，而机加工往往限于缩口件的外表面。因而，确定使缩口件能够达到最终形状精度并具有所要求的壁厚分布的预成形形状显得很有必要。在缩口成形过程中的金属流动较复杂，对于某些大型缩口件，由于在开口端部要求具有很大的变形，往往要在高温下进行缩口成形。

对于缩口工艺设计，有关定量设计的资料较少，在大多数缩口成形中，为了设计出缩口成形工艺，仍需通过较多的试验或实验来获得满意的成形工艺方案。Nadai 最早对缩口工艺进行研究，给出了进行应变分析的方法。利用 Nadai 的方法，Carlson 提出了在最终形状已知的情况下决定最初管坯形状的方法。在 Carlson 方法的基础上，Lahoti 等编写了预成形工艺设计的计算机程序。Kobayashi 推导出了一种根据模具表面上切向速度分布来确定初始管坯形状的近似方法。然而，上述方法只能给出缩口工艺的近似结果。

Park 等应用反向模拟方法来确定缩口过程的预成形工艺，采用设计出的预成形件，通过缩口后可得到壁厚一致的缩口工件。对于如图4.3-4所示的模具几何形状和缩口件形状，可以指定出两种预成形形状，预成形中所需的壁厚分布可以通过在预成形时改变其内径或外径来调整。这一准则可用来控制反向模拟过程中的边界条件。

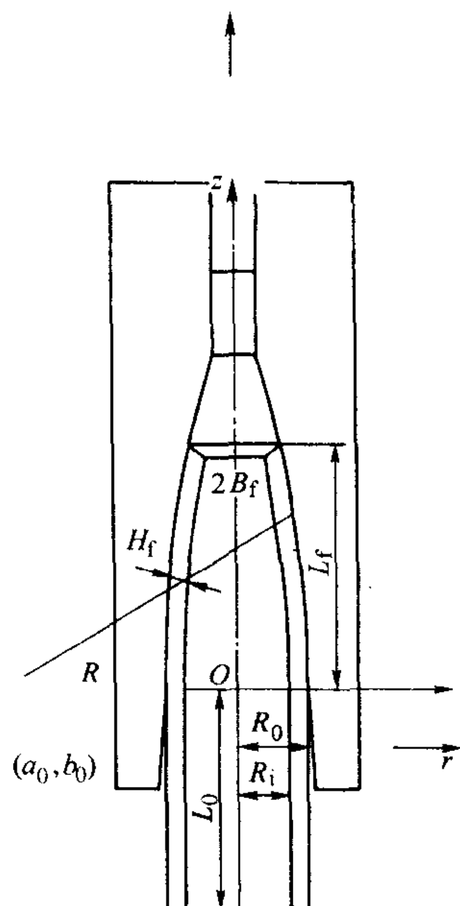


图 4.3-4 模具尺寸与缩口件形状

$L_0 = 47.26$ mm, $L_f = 48.84$ mm, $R = 21.526$ mm,
$R_i = 14.648$ mm, $B_f = 14.47$ mm, $H_f = 6.88$ mm（不变）,
$a_0 = 237$ mm, $b_0 = 18.75$ mm

对于外径不变的预成形形状（O形预成形件），对所有的有限元节点 n 有 $R_0^n = R_0$，其中 R_0^n 为沿缩口件外表面分布的有限元节点的半径，R_0 为模具入口处半径。这种类型的预成形件在反向模拟过程中边界条件的控制准则为：当满足条件 $R_0^n = R_0$ 时，与模具接触的节点（$R_0^n \leqslant R_0$）成为自由节点。

对于内径不变的预成形形状（Ⅰ形预成形件），边界条件控制的原则为：对于 $R_i^p < R_i$ 的节点，一旦满足条件 $R_i^p = R_i$ 时，与模具接触的外表面节点就成为自由节点，其中 R_i^p 为沿着缩口件内表面分布的节点的半径。

假设缩口件的材料为理想刚塑性材料（无加工硬化），针对上述缩口工艺进行反向模拟。每步压下量（增量行程步长）为2 mm。对于速度场的 Newton - Raphson 迭代，一般需要6~7次迭代即可获得收敛的速度场。要完成一步反向模拟计算，对于边界条件不变的情况，需要进行1~2次几何迭代；对于边界条件发生改变的情况，一般进行2~3次几何迭代。如果反向模拟中两个几何形状 Q 与 $Q^{(n)}$ 的误差小于给定的值时，即认为几何迭代收敛。图4.3-5将两种预成形形状与最终所得的缩口件形状进行了对比。其中，ξ_0^0 与 ξ_0^1 分别为O形预成形件与Ⅰ形预成形件的长度，即缩口成形后成为缩口件缩口部分长度 L_f 的预成形件那一部分的长度。由于缩口成形时管坯被拉长，预成形件要比缩口件短。

如果考虑加工硬化对于预成形设计的影响，就需要了解缩口件的应变分布和几何形状。但缩口件的应变分布取决于预成形件的形状，且预先未知（或不确定）。图4.3-6给出了考虑加工硬化影响的反向模拟和预成形设计的步骤。

首先对非加工硬化材料的缩口过程进行反向模拟和预成形设计，得到两种类型的预成形设计。然后利用这两种预成形进行加工硬化材料的缩口成形过程的正向模拟。模拟结束后获得缩口件的等效应变分布以及与指定的加工硬化材料缩口件几何形状十分接近，但又不完全相等的缩口件形状。将这种缩口件形状修正成精确的缩口件形状，并保持同样的等效应变分布。采用该等效应变分布，即可进行反向模拟和确定加工硬化材料的预成形件形状。在反向模拟每一步中，还

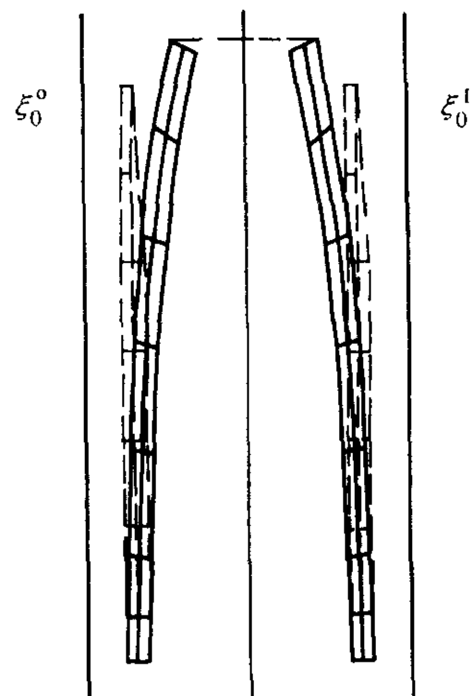


图 4.3-5 缩口件和无加工硬化材料的两种类型的缩口件和预成形件的形状

应根据材料的应力应变曲线对应变和相应的应力进行反向模拟。需要提及的是，如果采用非加工硬化材料的预成形形状的近似解，可消除图4.3-6所示的第一步计算过程。

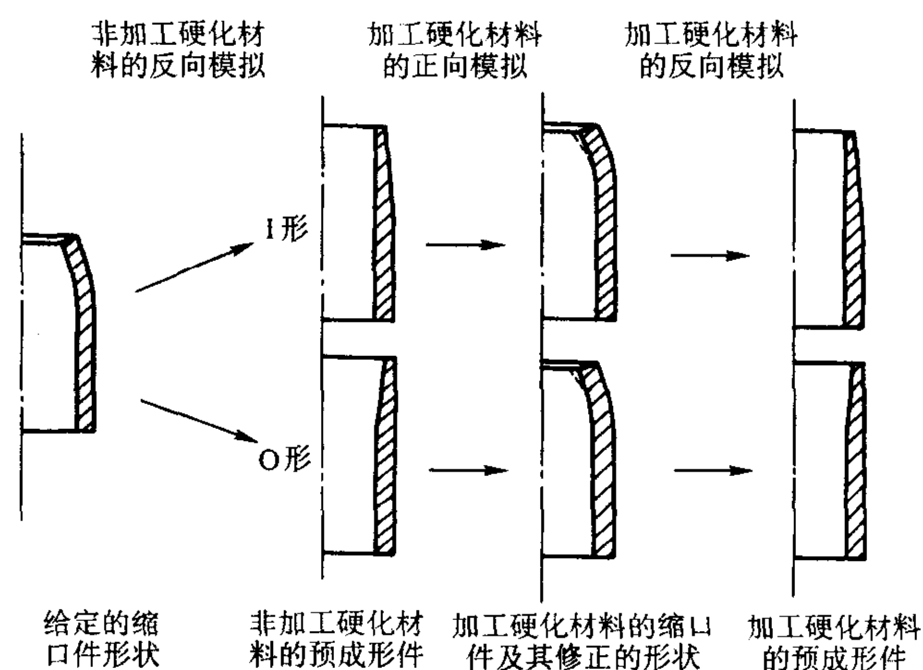


图 4.3-6 考虑到加工硬化效应的预成形设计

3.2 热缩口成形过程的预成形设计

对于热黏塑性变形过程的预成形设计问题，反向模拟过程就应当考虑温度变化的影响。除了满足式（4.3-1）所示的速度方程外，根据边界条件，还必须满足下面的温度方程和相应的边界条件：

$$\text{在 } t \text{ 时刻}\quad \boldsymbol{K}_c\boldsymbol{T} + \boldsymbol{C}(\boldsymbol{X})\dot{\boldsymbol{T}} - \boldsymbol{Q}(\boldsymbol{X}, \bar{\varepsilon}\,\dot{\bar{\varepsilon}}, \boldsymbol{T}) = 0 \tag{4.3-10}$$

式中，$\boldsymbol{K}_c$ 为热传导矩阵；$\boldsymbol{C}$ 为热容矩阵；$\boldsymbol{Q}$ 为热流向量；$\boldsymbol{T}$ 为节点温度向量；$\dot{\boldsymbol{T}}$ 为节点温度速率向量。

对于反向跟踪（$t \leqslant t_0$），在 $t = t_{0-1} = t_0 - \Delta t$，$\boldsymbol{X}_{0-1}$ 与 $\bar{\boldsymbol{\varepsilon}}_{0-1}$ 应满足式（4.3-6）给出的条件，并且 T_{0-1} 应该满足下列条件

$$T_0 - T_{0-1} = [\beta\dot{T}_0 + (1-\beta)\dot{T}_{0-1}]\Delta t \tag{4.3-11}$$

式中，$0 \leqslant \beta \leqslant 1$，在常用的几种差分格式中，$\beta$ 的取值有4种情况：$\beta = 0$，为向前差分（显式 Euler 方法）；$\beta = 1$，为向后差分（隐式 Euler 方法）；$\beta = 1/2$，为 Crank - Nicolson 差分格式；$\beta = 2/3$，为 Galerkin 差分格式。

在 t_{0-1} 时刻，将式（4.3-6）代入式（4.3-10），得

$$\boldsymbol{K}_c(\boldsymbol{X}_0, \boldsymbol{V}_{0-1})\boldsymbol{T}_{0-1} + \boldsymbol{C}(\boldsymbol{X}_0, \boldsymbol{V}_{0-1})\dot{\boldsymbol{T}}_{0-1}$$

$$-Q(X_0, V_{0-1}, \bar{\varepsilon}_0, \dot{\bar{\varepsilon}}_{0-1}, T_{0-1}) = 0 \qquad (4.3\text{-}12)$$

在式（4.3-12）中，X_0，$\bar{\varepsilon}_0$，V_{0-1}，$\dot{\bar{\varepsilon}}_{0-1}$应已知或通过求解速度方程获得。式（4.3-12）中温度 T_{0-1}是由式（4.3-11）求出的 $\dot{T}_{0-1}$的函数。既然 T_0 和 $\dot{T}_0$ 已知，在 T_{0-1}时刻通过假设边界条件就可以由式（4.3-12）解出 $\dot{T}_{0-1}$。一旦 $\dot{T}_{0-1}$已知，就可以算出 T_{0-1}并与假设的边界值相比较。反复迭代就可以完成温度计算的反向模拟。

对于高温缩口过程的反向模拟和预成形设计，已知缩口模具和缩口件的尺寸、工件与模具接触面上的摩擦、材料(AISI 1045 钢)、初始温度分布以及缩口成形速度（270 mm/s)。预成形设计过程由以下三步组成。

1）加载模拟。在已知缩口件和模具内的应变、温度场的条件下进行反向模拟，为了预先获得这些场的分布，利用一个预先设想的预成形件及其初始温度分布进行加载模拟，假想预成形件的设计应遵循预成形件体积与缩口件体积相等的原则，并假设预成形件的内部轮廓为直线，然后根据外部轮廓为直线的要求确定出假想预成形件的几何形状。假想预成形件的形状及其内部的初始温度分布如图 4.3-7 所示。

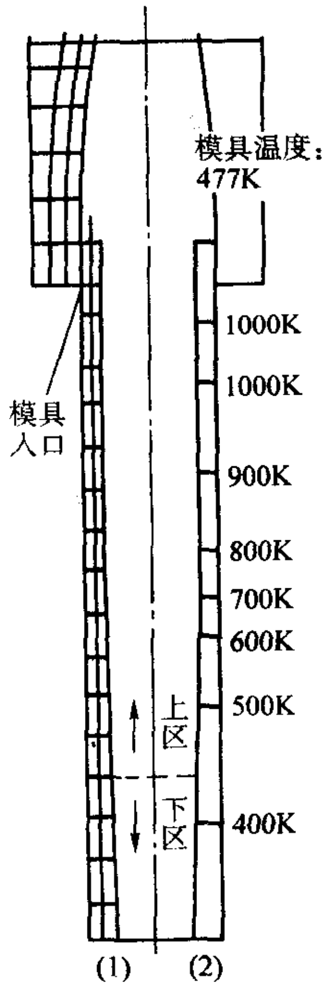


图 4.3-7
（1）缩口过程正向模拟的假想预成形件形状及其有限元模型；
（2）预成形件中初始温度分布

2）反向模拟准备。经步骤 1）模拟得到的最终缩口件形状与预先假想的不完全一致，需要对某些有限元节点位置进行调整，这样就可得到预先假想的缩口件形状和其中的应变、温度分布，如图 4.3-8 所示。

3）反向模拟和边界条件的控制。反向模拟从假想的缩口件形状和相应的应变、温度场开始。为了使预成形件的外表面为直线，通过假设每个与模具接触的节点在模具入口处脱离模具，然后沿着轴向运动，由此来控制反向模拟过程的边界条件。图 4.3-9（1）为最终的预成形件形状，图 4.3-9（2）为通过反向模拟得到的初始温度分布。图 4.3-9（2）中的温度分布与图 4.3-8（2）中所示的初始温度分布十分相近，表明不必再进行迭代。

反向模拟方法是进行预成形设计的一种新技术。对于缩

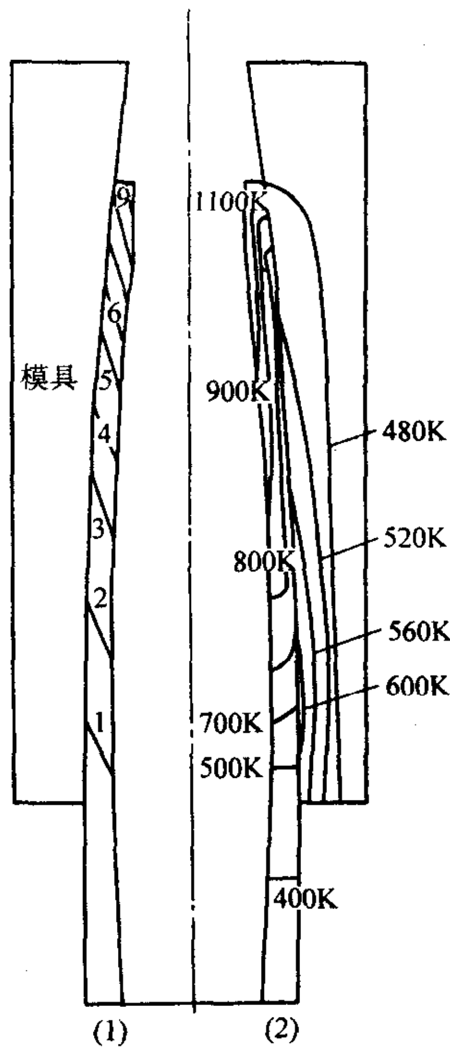


图 4.3-8
（1）假想的缩口件的应变分布；
（2）假想的缩口件和模具的温度分布

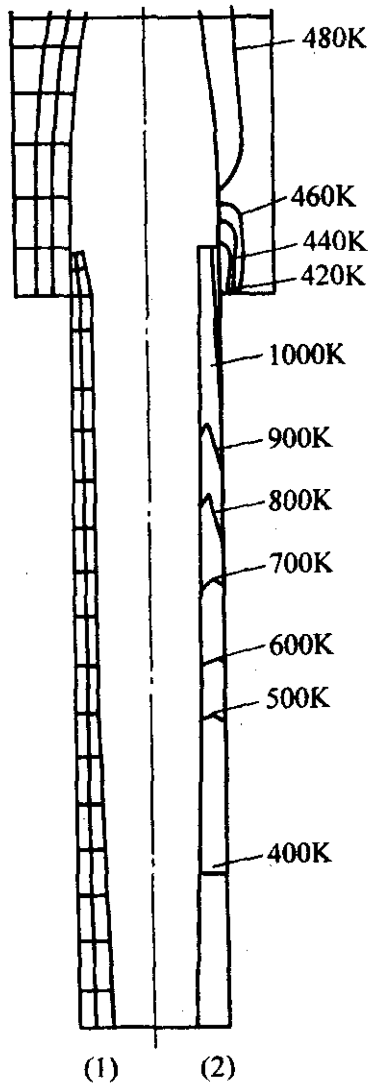


图 4.3-9
（1）预成形件形状；
（2）反向模拟结束后得到的初始温度分布

口成形问题，其边界条件的确定准则为几何控制准则，该准则在变形过程中维持不变。但是对于较复杂的变形过程，反向模拟路径是不断变化的。因而，一般成形过程的变形路径和反向模拟边界条件的确定就显得尤为重要。

4 反向模具接触跟踪方法

4.1 边界条件的确定

反向模具接触跟踪方法是一种确定反向模拟过程中接触节点脱离模具的判断准则。根据终锻件的形状和预测的金属流动情况，假想一个初始的预成形件，进行正向模拟，在正向模拟过程中记录模具边界与工件的接触时间及其顺序。采用一种与有限元网格无关的记录方法记录下模具边界与工件的接触顺序。得到该接触时间顺序后，可根据终锻件的成形情况，适当修正时间顺序，然后用作反向模拟边界条件的确定准则。

首先将终锻模具边界形状划分成一系列的直线段或圆弧段，采用一个假想的预成形件进行有限元正向模拟，在模拟过程中计算并记录每个模具边界线段与变形体的接触时间。这样，就不必保持原始的有限元网格及其边界节点的数目。如果某个模具边界线段重复与变形体接触或脱离，则最后接触的时间被认为是该线段的接触时间。为了较明确地反映边界条件的变化，模具边界线段的长度应小于边界单元的最小边界长度。采用这种模具边界线段记录方法，处于工件边界节点之间的模具边界线段与边界节点不相接触。但事实上，只要两个相邻的边界节点均接触了模具边界线段，则处于这两个相邻边界节点之间的所有模具边界线段都应与工件相接触。因此，这些处于中间的模具边界线段的接触时间顺序通过两个相邻的接触线段的接触时间作为端点数据进行线性插值求出。这样就给出了模具边界线段与工件接触的连续的时间顺序记录曲线。

如果在正向模拟结束后发现预先给定的预成形件不能满足预成形设计的目标或要求，则应对得到的模具边界线段进行修正，修正的目的是确保终锻模具型腔充满和材料具有合理的流动特征。修正的程度取决于初始选择的预成形件的形状与所期望的预成形形状的接近程度。一般地具有以下边界条件的修正准则。

1）如果存在未充满区域，则将未充满区域深腔处的模具边界线段的接触时间定义为上模到达所要求的最终位置的时间，在该线段两边的模具边界线段的接触时间逐步减小，最终与两边的接触线段的时间顺序相匹配。

2）对于具有多深腔的锻造模具，应当调整最早充满的型腔处的模具边界线段的接触时间，以使得所有模具深腔在终锻结束时同时充满。如果某些模具型腔未能充满，则边界条件的修正准则可按照准则 1）进行。这种修正准则可确保预锻成形后获得合理的材料体积分布。

3）如果在正向模拟结束时，终锻件存在着较大的飞边，应调整飞边处模具边界线段的接触时间，以获得所要求的飞边尺寸。在飞边槽内最外部的模具边界线段的接触时间定义为上模运动到所要求的最终位置的时间。

4）由于对所记录的边界条件进行了不同程度的修正，在反向模拟过程中，工件的某些边界节点可能穿透模具。为防止这种情况产生，应对这些节点施加额外的接触约束，并保持这种约束直至这些节点不再穿透模具。对于一个反向模拟步，当应用收敛的速度场刷新工件的几何形状后，通过节点与模具的几何关系检测穿透模具的节点。一般而言，这些节点大都为先前模拟步中脱离模具的边界节点。对于穿透模具的节点应施加边界接触约束，使其只能沿着模具表面的切线方向运动。然后返回到当前模拟步的开始状态，重新确定工件内的收敛速度场，进行模具位置和工件几何形状的刷新。当前模拟步结束后，立即释放对穿透模具的节点的接触约束。连续进行这种检测直至完成整个反向模拟过程。

在反向模拟过程中，根据修正的边界条件记录，如果一个模具边界线段在该模拟步内将脱离工件表面，则位于该模具边界线段内的边界接触节点应脱离模具。在正向模拟中最后接触模具边界的有限元节点，在反向模拟过程中将最早脱离模具而成为自由节点。当反向模拟过程结束后，除了初始接触的模具线段外，其余模具线段将成为自由边界。反向模拟结束后对应的预成形件的形状即为所要求的预成形件。该预成形件是由一系列直线段连接而成的。金属成形过程的有限元步进解法及其增量式的边界条件施加，可能使得预成形形状出现小齿状的表面。这种齿状表面不能直接用作为预成形形状，应当根据该预成形件的形状，通过对表面线段的光滑处理，得到光滑过渡的预成形形状。光滑时应尽量保证与原始齿状的预成形形状接近。一般地，所进行的光滑处理程度较小，这种光滑调整可通过手工或曲线拟合方法进行。

4.2 工序数目的确定

预成形设计的另一个重要方面是工序数目的确定。若从产品的成本考虑，则应以尽可能少的工序完成塑性成形产品的生产。影响工序数目的因素很多，例如：成形材料的可加工性、工件形状的复杂程度、工件可能产生的成形缺陷等等，这些缺陷包括材料开裂、模具受力过大、模具未充满、不良表面质量、不良的纤维分布、折叠、不允许的组织变化、微观裂纹等等。在考虑工序数目时，都应当考虑这些因素的影响，但如何定量地考虑这些因素的影响以及对工序数目的综合效应，目前尚未有方法可言。因此，当前能够从理论上确定合理的工序数目还是相当困难的，这也是十分重要的研究课题。目前生产中采用的方法仍然是经验型的设计方法。

采用有限元技术进行反向模拟和预成形设计，可以采用下列方法进行工序数目的确定。

1）根据要求的终锻件和模具形状，采用一个初始给定的预成形形状进行终锻阶段的正向模拟，通过反向模拟设计预锻模具形状。

2）应用正向模拟技术，模拟从初始坯料到预锻件的预锻过程，通过变换不同的坯料尺寸，选择一种较合理的初试坯料尺寸，如果可得到合理的预锻件形状，则不再需要其他预成形工序，反之，则需要另一个预成形工序。

3）如果需要另一个预成形工序，则应用第 1）步中的预锻模具形状进行反向模拟，以确定另一个预锻工序的预成形件形状及预锻模具形状。

4）采用初始坯料形状进行预锻工序的正向模拟，如果在预锻工序结束时，未获得所要求的预锻件形状，则重复上述步骤，设计下一个预锻工序；如果获得了所要求的预锻件形状，则工序数目设计完毕。

4.3 通用透平圆盘锻造过程的预成形设计

如图 4.3-10 所示为一透平圆盘（Generic Turbine Disk - GTD）锻件，属轴对称锻造问题。该锻件的轴截面形状由一系列的直线和圆弧段组成，为方便计，用 c_l（$l = 1, 2, 3, \cdots, 36$）表示不同的线段。GTD 锻件的形状复杂，其径向尺寸远远大于高度尺寸，若采用一次锻造，难以从初始毛坯锻造出终锻件。为了使材料在终锻阶段尽可能少地产生径向流动，预锻阶段应使初始毛坯具有合理的体积分布。GTD 锻件材料为镍基合金，设材料的应力应变率关系为 $\bar{\sigma} = Y_0 \dot{\bar{\varepsilon}}^{0.3}$，$Y_0 = 68.93$ MPa。假使在工件和模具接触面上的摩擦为常摩擦，摩擦因子为 $m = 0.2$，上模的运动速度为 -0.02 m/s，下模静止不动。

（1）单工序成形坯料优化设计

在预成形设计之前，通过正向有限元模拟可得到有关变

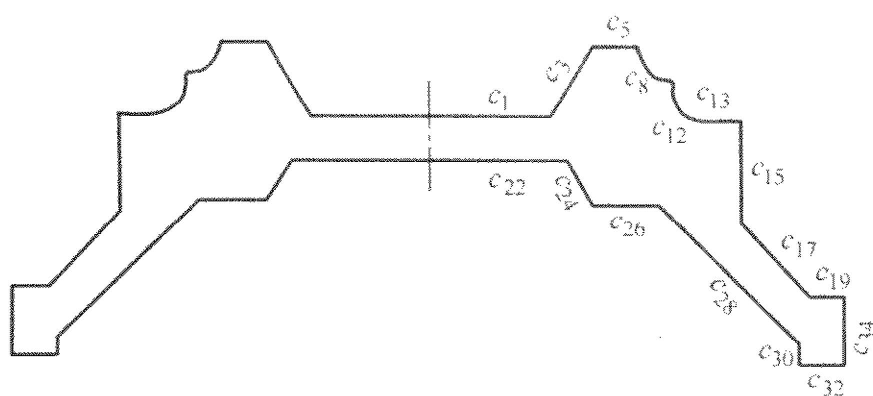


图 4.3-10　通用透平圆盘锻件

形特点、模具载荷、应变分布、材料流动等信息，对于了解 GTD 锻件的成形过程特点具有一定作用。对于采用单工序成形，即从初始坯料开始直接进行终锻成形的情况，模拟了 4 种不同高径比的初始坯料的成形过程，模拟结果如图 4.3-11 所示。分析 GTD 锻件的形状特征，可看出处于模具线段 c_{15} 和 c_{17} 之间的拐角处的半径 R_c 为临界半径，其对应的坯料的临界高径比为 0.286 3。因此，如果坯料半径大于 R_c，则材料的变形特征将与小于 R_c 的坯料的变形特征相差较大。如果坯料半径小于 R_c 或高径比大于 0.286 3，则材料可同时充满内外型腔。对每一种情况的材料流动，采用两个高径比进行正向模拟。$H/D = 0.151\ 3$ 和 $H/D = 0.218\ 1$ 为一种金属流动模式。$H/D = 0.286\ 3$ 和 $H/D = 0.340\ 9$ 为另一种金属流动模式。$H/D = 0.340\ 9$ 时，坯料的半径与 R_c 和 c_{12} 表面拐角处半径的平均值相等。$H/D = 0.218\ 1$ 时，坯料的半径与 R_c 与 c_{17} 和 c_{19} 之间的拐角处的平均值相等。$H/D = 0.151\ 3$ 表示坯料的最大半径（在 c_{17} 和 c_{19} 之间的拐角处半径）。

图 4.3-11a 是当 $H/D = 0.151\ 3$ 时，锻造过程中某一时刻的锻件形状。在开始变形阶段，材料产生弯曲变形，其主要原因是坯料高径比较小，坯料形状呈扁盘型。当材料接触模具底部的倾斜表面 c_{28} 后，就会向上流入上模型腔，材料逐渐地接触上模的 c_3 和 c_5 面。当上模的压下量等于 55.73 mm，材料就会与下模的 c_{24} 表面分离，同时在锻件表面形成凹坑。随着上模压下量的增加，凹坑扩展为折叠缺陷。凹坑的形成可由变形体的速度场分布说明。由于材料与上模的 c_3 和 c_{15} 表面接触，材料就会流入上模的深腔，对称轴附近的材料也会沿轴线向外流动。这一流动方式使材料与下模的 c_{24} 表面分离。在分离表面上，上层材料比下层材料流动快。因此，这一区域的材料会产生回流，从而在锻件表面产生凹坑。凹坑逐渐扩展，直至材料接触上模的 c_{12} 表面。随后由于上模的逐步下压，凹坑就会成为折叠缺陷。当锻造过程完成后，同时也形成了一个薄而宽的飞边。由于飞边冷却较快，飞边槽附近的模具很容易损坏。通过模拟可得到以下有用的信息。

1）与 GTD 锻件的体积分布相比，虽然该尺寸的坯料沿径向具有很好的体积分配，但是，为了消除或减少飞边，应适当增加原始坯料的高径比。

2）为了避免下模 c_{22} 和 c_{24} 表面之间拐角处的折叠缺陷，应使材料与上模的 c_{12} 表面接触，而不能与上模的 c_1、c_3 和 c_{17} 接触。

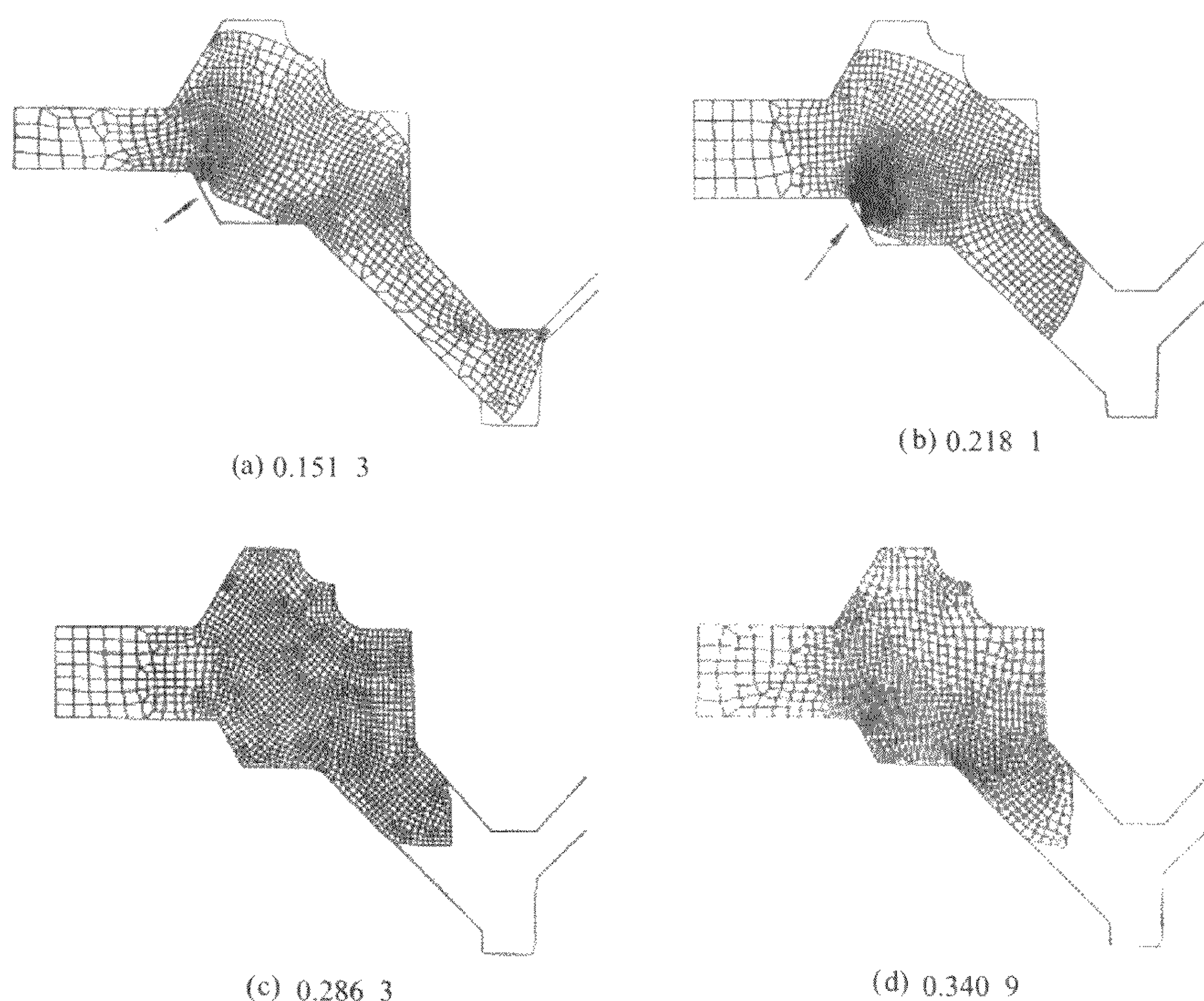


图 4.3-11　不同高径比（H/D）坯料的单工序锻造过程的正向模拟

图 4.3-11b 是当 $H/D = 0.218\ 1$ 时，锻造过程某一时刻的锻件形状。这种情况的材料流动特性与 $H/D = 0.151\ 3$ 时相似。同样在 c_{22} 和 c_{24} 之间的拐角处会发生折叠缺陷，但是在锻造过程结束时可得到一个无飞边的终锻件。上述两种情况（$H/D = 0.151\ 3$ 和 $H/D = 0.218\ 1$）的锻造过程表明，对于单工序锻造过程，当坯料高径比大于临界值 0.286 3 时，总会使终锻件产生锻造折叠缺陷。

对于 $H/D = 0.286\ 3$ 的情况，初始坯料的半径比 c_{13} 和 c_{15} 之间拐角处的半径 R_f 稍小一点。图 4.3-11c 给出了该种坯料锻造过程中某一时刻的锻件形状。由于模具 c_{15} 表面对材料流动的约束，在初始阶段材料主要是向内部型腔流动。当上模向下移动到图 4.3-11c 所示的位置时，上模型腔已经完全充满。上模的继续移动会使得材料通过类似于大飞边槽的斜通道流入外部型腔。这种材料的流动特性类似于挤压时材料的流动特性，使模具载荷增加迅速。斜通道入口（c_{15} 和 c_{17} 之间及 c_2 和 c_{28} 之间的转角）处附近的模具表面会经

历较长的材料与模具表面相对滑动过程，同时，由于材料和模具之间存在着较大的接触应力和较高的滑动速度，会导致模具过早磨损，但在整个锻造过程中，锻件尚未发生折叠缺陷，并且当锻造过程结束时，可得到无飞边的终锻件。因此，采用高径比为 0.286 3 的坯料进行单工序锻造，可实现少无飞边和无缺陷的锻造过程。

图 4.3-11d 表示坯料高径比 $H/D=0.3409$ 时，锻造过程中某一时刻的锻件形状。由图可以看出，材料的流动性与 $H/D=0.2863$ 时相似，只是材料更容易充满模具的内部形腔。窄通道入口（c_{15} 和 c_{17} 之间，c_{26} 和 c_{28} 之间的拐角）处附近的模具表面与材料相对滑动的时间较高径比 $H/D=0.2863$ 时情况更长。因而会使模具磨损更大。

对比四种不同坯料高径比的锻造过程表明，如果忽略模具磨损，第三种高径比（$H/D=0.2863$）的原始坯料最适合于进行单工序锻造。用这种高径比的坯料可以实现无飞边、无缺陷的通用透平圆盘锻造。如果原始坯料的半径比 c_{15} 和 c_{17} 之间的拐角处半径大，则在进行通用透平圆盘锻造时，锻件会在 c_{22} 和 c_{24} 之间的拐角处产生折叠缺陷。但是单工序通用透平圆盘锻造过程，会造成较大的模具磨损，使模具磨损较小的预锻模具设计就成为两工序锻造工艺设计的主要工作。

(2) 两工序锻造的预成形设计

预成形过程的主要作用之一是获得符合终锻件形状要求的体积分布。以上进行的单工序正向模拟给出了有助于预成形设计的有关信息。从上述四种不同高径比坯料的锻造情况可以发现，$H/D=0.1513$ 的情况与其他三种情况相比，可以得到较合理的径向体积分布。但是，在 $H/D=0.1513$ 时，却存在着折叠缺陷和较大飞边。为了避免产生折叠缺陷，所期望的预锻件在进行终锻的开始位置，应使得终锻模上模的 c_{12} 表面能够在终锻开始时就接触到预成形件，而且，在终锻初始位置，上模的 c_1 和 c_3 表面不能接触预成形件。为了消除或减少飞边，应在预成形时就减少飞边槽附近材料的体积。这样，根据 $H/D=0.1513$ 时锻造过程中间一个时刻的锻件形状，通过上述两种修正，可以设计出如图 4.3-12 所示的预锻模具形状。采用该预锻模具和初始的圆柱体坯料，经过预锻可得到图 4.3-12 中所示的预成形件形状，此预成形件用作反向模拟的假想预成形件。

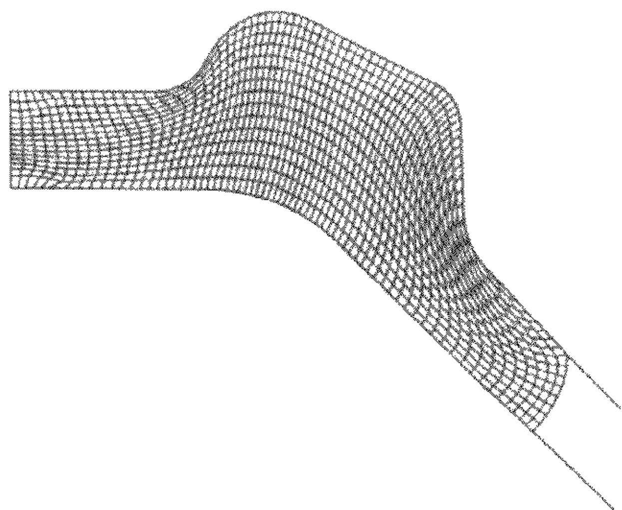


图 4.3-12 用于反向模拟的假想预成形件形状

1) 正向试验模拟 图 4.3-13 所示为终锻阶段（第二工序）正向试验模拟的材料流动过程。图 4.3-13a 表示预成形件和终锻模具的初始位置。显然，终锻模具仅在 c_{12} 和 c_{15} 表面上与工件接触。这种接触可防止在 c_{22} 和 c_{24} 之间的圆角部分产生凹坑，这样就可以避免产生折叠缺陷。在正向试验模拟中，记录了工件与模具上下模的边界接触状况。在模拟过程结束后，可根据材料流动特性和边界条件的修正准则对记录的边界条件进行修正。修正后的边界接触条件可用来控制反向模拟时节点与模具的脱离。在正向试验模拟中，c_{22} 和 c_{24} 之间的拐角处的折叠缺陷得以避免。当上模移动到图 4.3-13b 所示的位置时，内腔完全充满。但是外腔却未充满。这意味着预成形件的形状需要更大的变形才能获得合理的径向体积分布。由于内外模腔充满的程度不同，需要对于记录的边界接触条件进行修正。外腔左侧的所有模具边界线段的接触时间都减去同样的值，该值等于所要求的终锻行程时间和内腔完全充满时所对应的行程时间之间的差。另一方面，外模型腔材料的流动速度不稳定，尤其是外模腔下模左拐角处。为了避免出现这种现象，需要调整外模腔底部的模具边界线段的接触时间，以使在终锻过程的初始阶段预成形件就几乎与模具表面相接触。通过以上边界条件的调整，可获得用于反向模拟的上下模的边界条件。

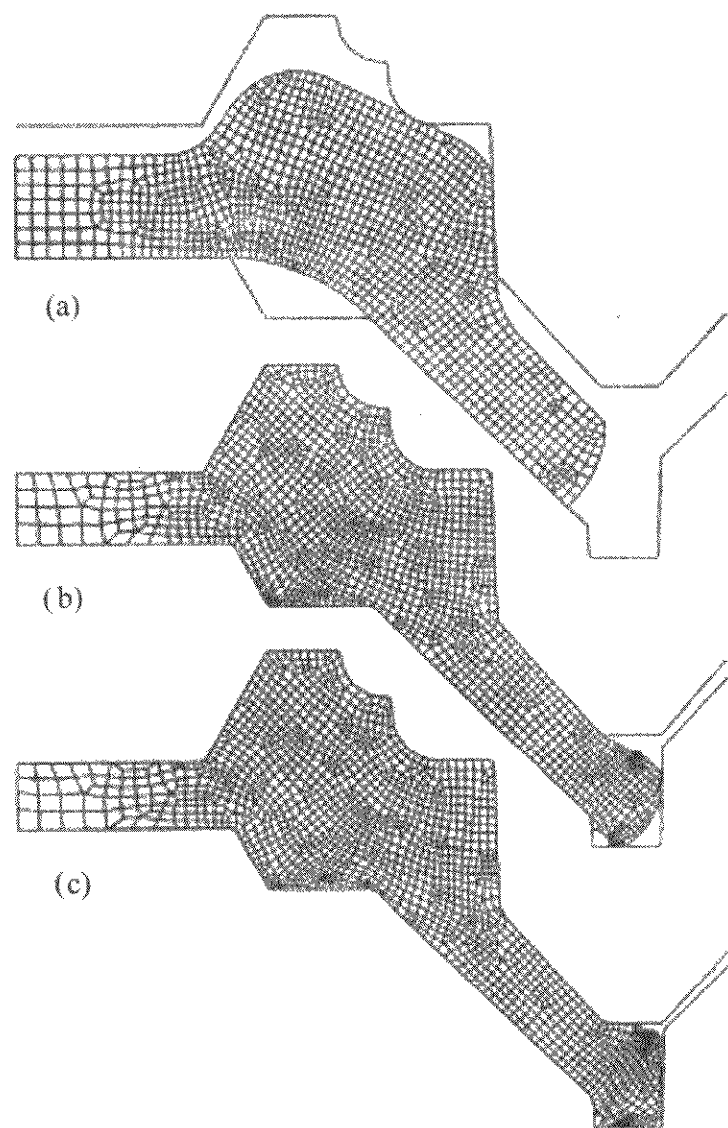


图 4.3-13 采用假想预成形件的正向模拟试验

2) 反向模拟 利用修正的边界条件，即可进行反向模拟和预锻模具形状设计。图 4.3-14 表示所要求的既无飞边又完全充满型腔的终锻件。在反向模拟结束时，图 4.3-14d 所示的中间锻件形状即为所要求的预成形件形状。通过对反向模拟求得的预成形件表面进行少量光滑修整，可得到如图 4.3-15 所示的预锻模具形状。

3) 正向模拟验证 为了验证预成形设计的有效性，需要进行从原始坯料，经预锻过程到终锻过程的整个变形过程的有限元正向模拟。图 4.3-15a 所示为初始圆柱体坯料和预锻模具形状，应注意，初始坯料的尺寸也会影响预成形件的形状和质量，如果不能正确的选择或设计初始坯料的高径比 H/D，也会使预锻件产生缺陷。例如，当 $H/D=0.1513$ 时，会使预锻件在下模拐角处产生凹坑，如图 4.3-16 所示。因此，初始坯料的高径比应大于 0.2，才能避免凹坑产生。如果再考虑到坯料在模具型腔的体积分布要求，$H/D=0.2$ 是选择最优初始坯料高径比的下限值。图 4.3-15c 表示终锻阶段结束位置和对应的预锻件形状，该预锻件可用来进行终锻成形的预锻件形状。

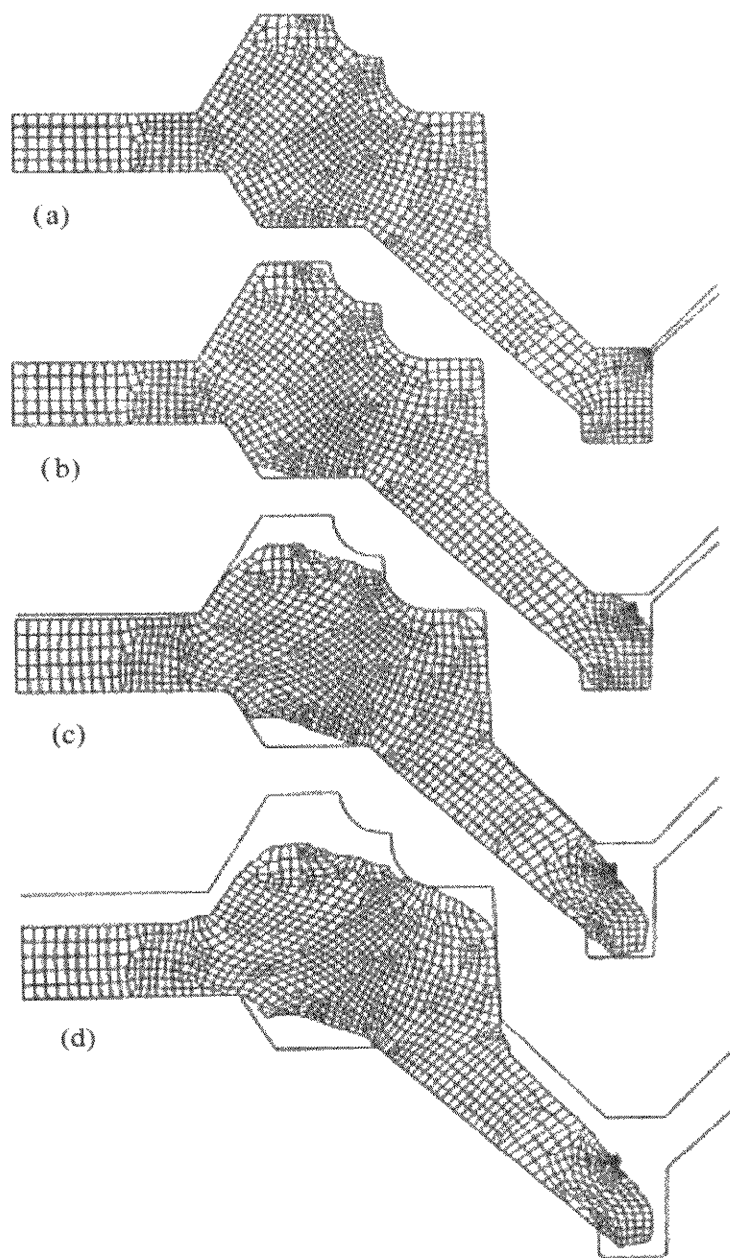


图 4.3-14　终锻过程反向模拟的材料流动模式

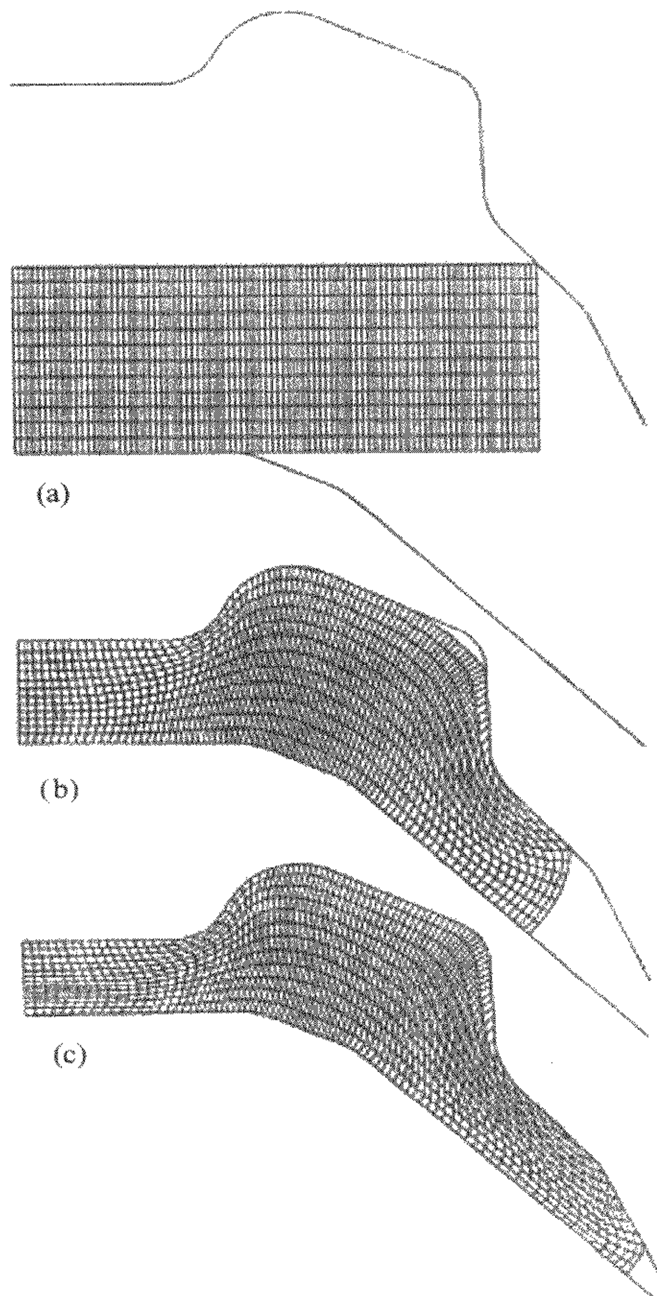


图 4.3-15　用反向模拟设计的预锻模具进行预锻过程的正向模拟验证

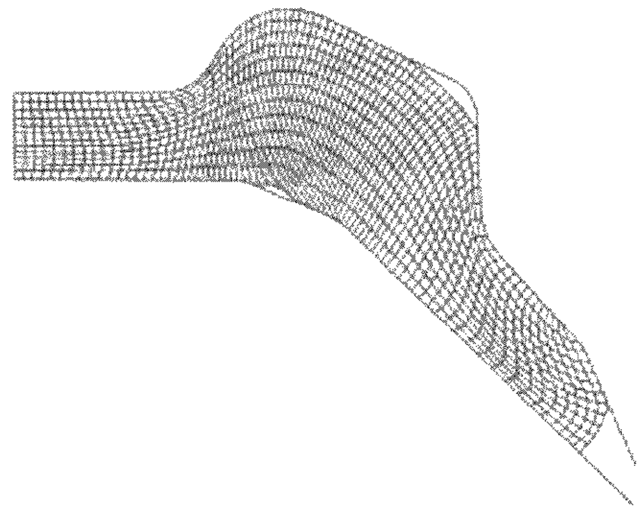


图 4.3-16　不合适初始坯料尺寸导致的预锻件缺陷

图 4.3-17 表示终锻阶段正向模拟过程的材料流动模式，其中图 4.3-17a 为终锻阶段的初始位置。由模拟结果可见，所设计的预锻模具锻造出的预锻件形状为终锻阶段提供了合理的材料体积分布，内外模腔之间几乎没有材料转移。在初始阶段，通过上模向工件的 c_{12} 表面施加压力，防止了在 c_{22} 和 c_{24} 之间的过渡处凹坑的产生。因此，避免了终锻件折叠缺陷的产生。在终锻工序结束阶段，材料先充满型腔，最后才向飞边槽方向流动。因此，飞边是在终锻过程中的最后时刻形成。如果坯料下料尺寸精确，完全可以实现无飞边锻造，而且能够完全充满模具型腔。内外型腔几乎同时充满说明了沿锻件径向预锻件的体积分布与终锻件的体积分布要求相符。通过反向模拟设计，所设计的预锻模具形状和经预锻后所得到的预锻件形状完全可使终锻过程实现无飞边、充满型腔、无折叠缺陷和模具受力状况良好的通用透平圆盘锻件。

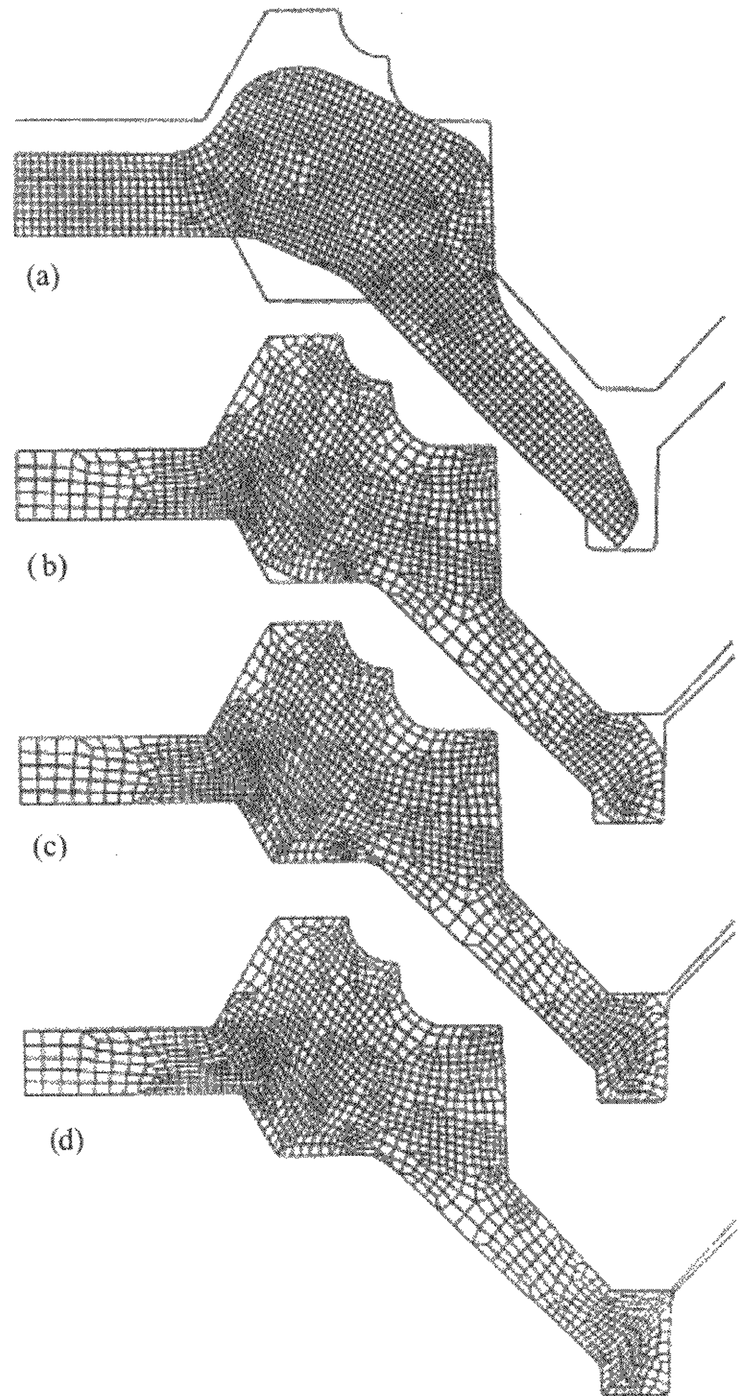


图 4.3-17　预锻模具和预锻件设计结果的正向模拟验证

图4.3-18给出了通用透平圆盘锻造工序顺序、预锻件形状、初始坯料形状和镦粗制坯的镦粗比。整个圆盘锻件的工序为下料、镦粗、预锻和终锻。如果可以获得高径比为 $H/D=0.2$ 的初始毛坯，则可省略镦粗工步，此时锻造工序为下料、预锻和终锻。

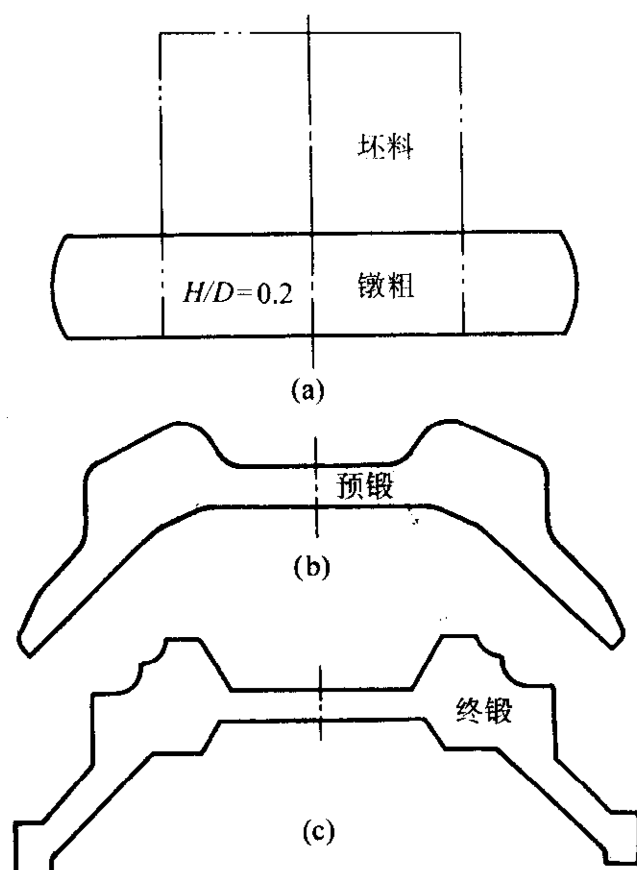


图4.3-18 通用透平圆盘锻造工序设计结果

5 基于锻件形状复杂程度的控制准则

本节介绍另一种有限元反向模拟和预成形设计的边界条件控制准则。该准则利用工件的形状复杂性系数，其目的是为建立一系列与模具相接触的边界节点在反向模拟过程中的脱离顺序。

5.1 工件形状复杂系数及边界条件控制准则

在锻造生产中，具有深腔槽、小圆角半径和薄截面的锻件增加了锻造成形的难度。实际上，锻件形状的复杂程度代表了锻造成形的难度。初始简单形状（例如圆柱体）的坯料经过锻造成形成了具有复杂形状的锻件，一般而言，终锻件形状越复杂，需要的成形工序越多。因此，工序数目和每个工序对应的模具形状与工件形状的复杂系数相关。对于轴对称锻件，不论其多么复杂，工件都是从圆柱形坯料开始进行锻造成形。为此，有必要确定终锻件形状和初始坯料形状的关系。由 Teterin 提出的锻件形状复杂系数建立了棒料与最终形状之间的联系，并被广泛地应用于锻模飞边槽的设计。轴对称锻造中的形状复杂系数 S（也称为几何难度系数）是轴对称锻件截面几何形状复杂程度的表征。其值如下：

$$S=\alpha\beta$$

式中，α 为轴向形状复杂系数；β 为形状复杂系数。

其定义如下：

$$\alpha=\frac{X_f}{X_c}$$

$$X_f=\frac{P^2}{F}$$

$$X_c=\frac{P_c^2}{F_c}$$

$$\beta=\frac{2R_g}{R_c}$$

式中，P、F 分别为轴向横截面的周长和表面积；P_c、F_c 分别为轴向横截面形状外接圆柱体的周长及其表面积；R_g 为对称轴到半个轴截面的重心或质心的距离；R_c 为外接圆柱体的半径。

对于多工序锻造过程来说，终锻工序的形状复杂系数由下式决定：

$$S_f=\frac{S_{ff}}{S_{fp}}$$

式中，S_{ff}为终锻件的形状复杂系数；S_{fp}为预锻件的形状复杂系数。

若锻件形状为一圆柱体，则 Teterin 形状复杂系数等于1。当给定的终锻件的形状复杂系数越来越大时，这就意味着需要进行多工序锻造。因此，形状复杂系数是一个重要的工艺设计参考因素。在锻造过程有限元正向模拟过程中，形状复杂系数随着模具的不断压下而增大。当金属充满锻模型腔时，锻件的形状复杂系数增至最大。在模具行程的结束阶段，使金属尽量能够较迅速地填充深腔，从而减少锻造力较大的模具行程，在充满型腔后再使多余金属溢出形成飞边，是我们想达到的理想结果。因此，在终锻过程的结束阶段，理想的预锻和终锻过程应使变形体的形状复杂系数随着上模压下行程而迅速增大。相应地，如果进行反向模拟，随着模具的分离，形状复杂系数应该以同样的方式减小。这就定义了节点脱离模具的控制准则，即在反向模拟中，最大限度地减小锻件的几何复杂系数。

在反向模拟中，每一步从模具上释放一个接触节点。节点的选择是通过试验搜寻而获得的。在搜寻过程中，轮流释放每个边界接触节点，进行反向模拟，同时计算对应的形状复杂系数并进行比较，产生最小形状复杂系数的节点将在该步内脱离模具。在反向模拟过程中，可根据锻件的几何形状，对锻件边界节点进行分组，每组中的节点相邻，每组的端部节点一般取为模具型腔内凹处的节点，处于每组节点端部的节点是可能较早脱离模具的节点。上述对边界节点进行分组处理的目的是为了节省计算时间。选择节点以后，初始的优先权就给了位于模具型腔内角半径附近的节点，这样可保证预锻模具型腔内角半径大于终锻模具的内角半径，这也符合预锻模具的设计规则和常识。

利用形状复杂系数作为控制准则提供了边界条件的控制方法，它仅仅依靠边界节点的坐标而与节点或模具的速度不直接有关。因此，这个准则可以同时应用于运动或静止的模具，从而对于上下不对称的工件，在其上、下部分提供了一致的形状控制方法。

5.2 预成形设计应用举例

图4.3-19所示为一轴对称等温锻造过程，选用应变率硬化材料，其应力应变之间的关系为$\bar{\sigma}=Y_0\dot{\bar{\varepsilon}}^{0.3}$，其中 $Y_0=68.93$ MPa。假设工件与模具之间的摩擦状态为常摩擦状态，摩擦因子为0.15。下模固定，上模运动速度为 -25.4 mm/s。因锻件关于 Z 轴对称，故取1/2模型进行有限元分析。

（1）单工序正向模拟

图4.3-20所示为单工序锻造过程金属流动情况。其中图4.3-20a所示为初始坯料网格和模具位置，圆柱体坯料的半径为131.3 mm，高径比为0.65。1/2横截面模型的初始和结束时的单元数分别为400和827。金属首先充满中心型腔。当上模到达图4.3-20d所示的位置时，中心处的上模腔和下模腔完全充满，金属开始流向飞边槽，但是上模两侧型腔仍有很大的空间未充满。随着上下模具的闭合，金属将流向上模两侧型腔和飞边槽。中心型腔的金属必须经过很长的流动路径才能到达上模两侧型腔，这种长距离的金属转移需要较

大的变形力，并且增加了模具的磨损。

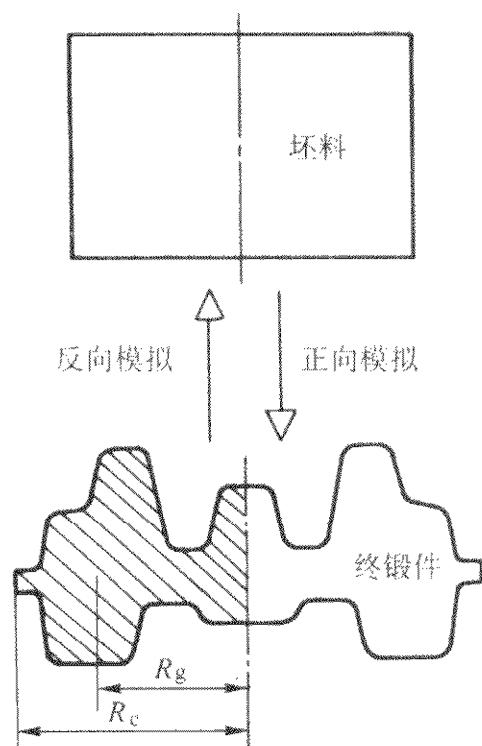


图 4.3-19 坯料和终锻件图

当变形到达如图 4.3-20f 所示的位置时，即模具到达要求的飞边高度 15.24 mm 时，上模两侧型腔仍未充满，并且金属被挤入飞边槽形成飞边。继续闭合模具，行程极限就会超出要求，只会导致更多的材料流入飞边槽，因此无法获得符合高度尺寸要求的锻件。这种伴随着大变形载荷的过量挤压还将引起飞边槽入口处模具的磨损。从以上情况可看出，有必要在终锻之前采用一道预锻工序。

(2) 反向模拟预成形设计

假设采用两工序锻造过程，包括预锻和终锻工序。采用基于形状复杂系数的节点分离准则，并通过反向模拟的方法来设计预锻模具。图 4.3-21 所示为反向模拟过程中的金属流动情况，其中图 4.3-21a 为反向模拟的开始状态，工件完全充满模具，即除了飞边外侧几个节点外，工件表面的所有节点都与模具接触。在反向模拟过程中，每一步只有一个节点从上模或下模脱离。当所有的节点与模具分离时，反向跟踪过程即告结束，所对应的中间锻件形状即为预锻件形状，如图 4.3-21f 所示。通过光滑如图 4.3-21f 所示的中间工件形状即可得到预锻模具的形状。图 4.3-22 所示为预锻模具形状和终锻模具形状的比较。从图 4.3-21 和图 4.3-22 可得以下结论。

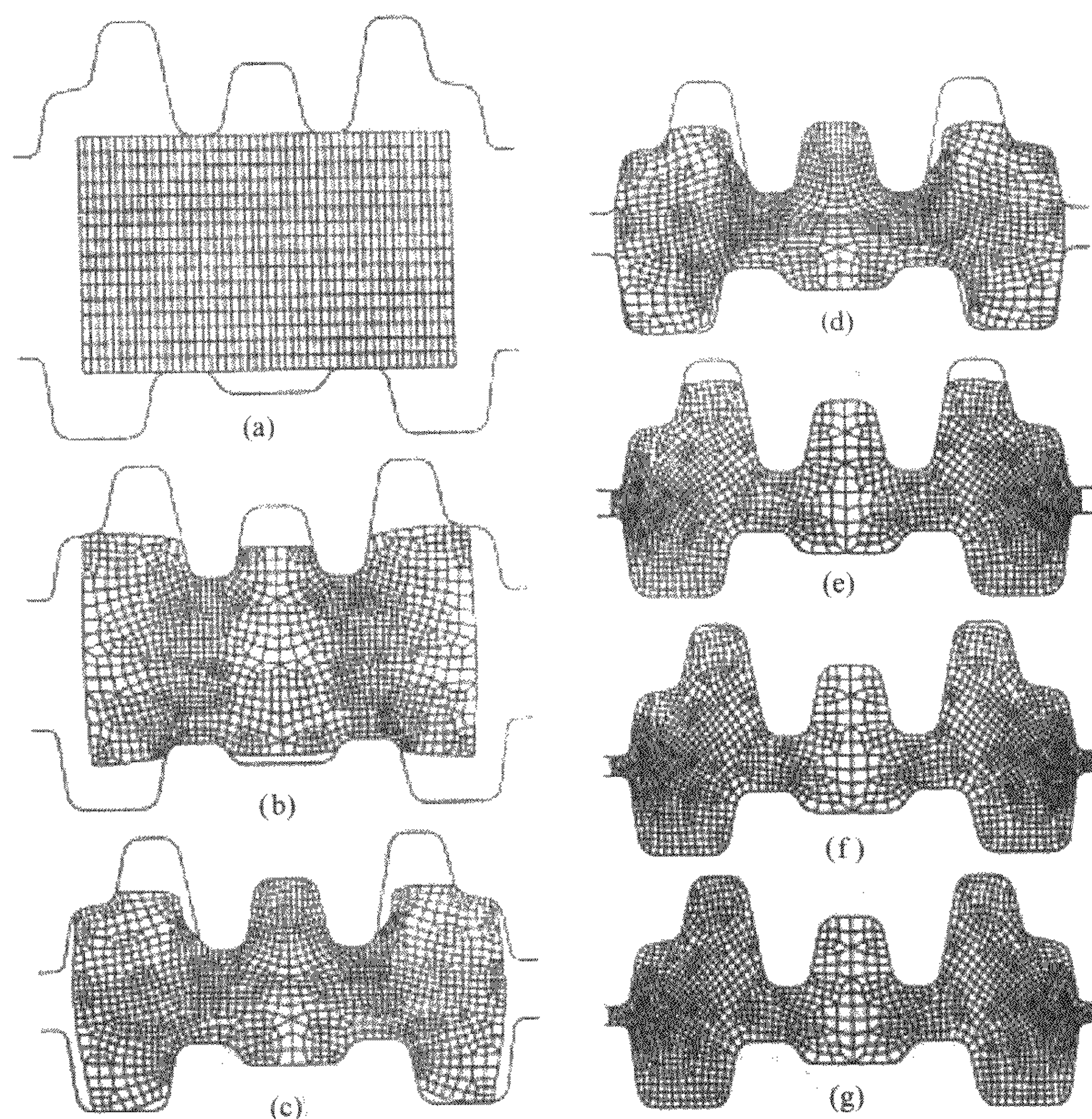


图 4.3-20 单工序锻造过程正向模拟的金属流动情况

1) 根据本方法所获得的预锻件形状较为光滑，而且材料体积分布合理。位于预锻模具中心型腔的材料体积基本上接近于终锻模具型腔所要求的体积，模具两侧型腔的体积分布亦如此。因此在终锻过程中中心型腔内的金属只在中心型腔内局部流动，两侧型腔内的金属也只向两侧型腔填充。这种金属流动模式有效地减少了终锻工序中金属向侧部型腔流动以及向飞边槽流动的趋势，从而降低了模具的磨损和终锻载荷。

2) 两侧型腔的金属在终锻阶段的流动模式为镦挤模式，这就克服了单工序锻造过程流动模式的缺陷。

3) 预锻件或预锻模的过渡圆弧半径大于终锻件或终锻模相应的过渡圆弧半径。因此符合经验设计方法的要求或规则。

(3) 正向模拟验证和比较

为了验证预锻件设计结果的有效性，采用有限元正向模拟方法来模拟两工序锻造过程。图 4.3-23 所示为预锻阶段材料的流动情况，坯料尺寸和模具与工件表面之间的摩擦因数与单工序相同，当上模移动到要求的位置时，材料完全充满模具型腔并且获得了较小飞边的合格预锻件，如图 4.3-23e 所示。图 4.3-24 所示为终锻阶段的金属流动情况，从中可以看出中心和两侧型腔的金属流动都很小，而且两个型腔彼此之间几乎没有金属材料的转移，这是因为大部分的材

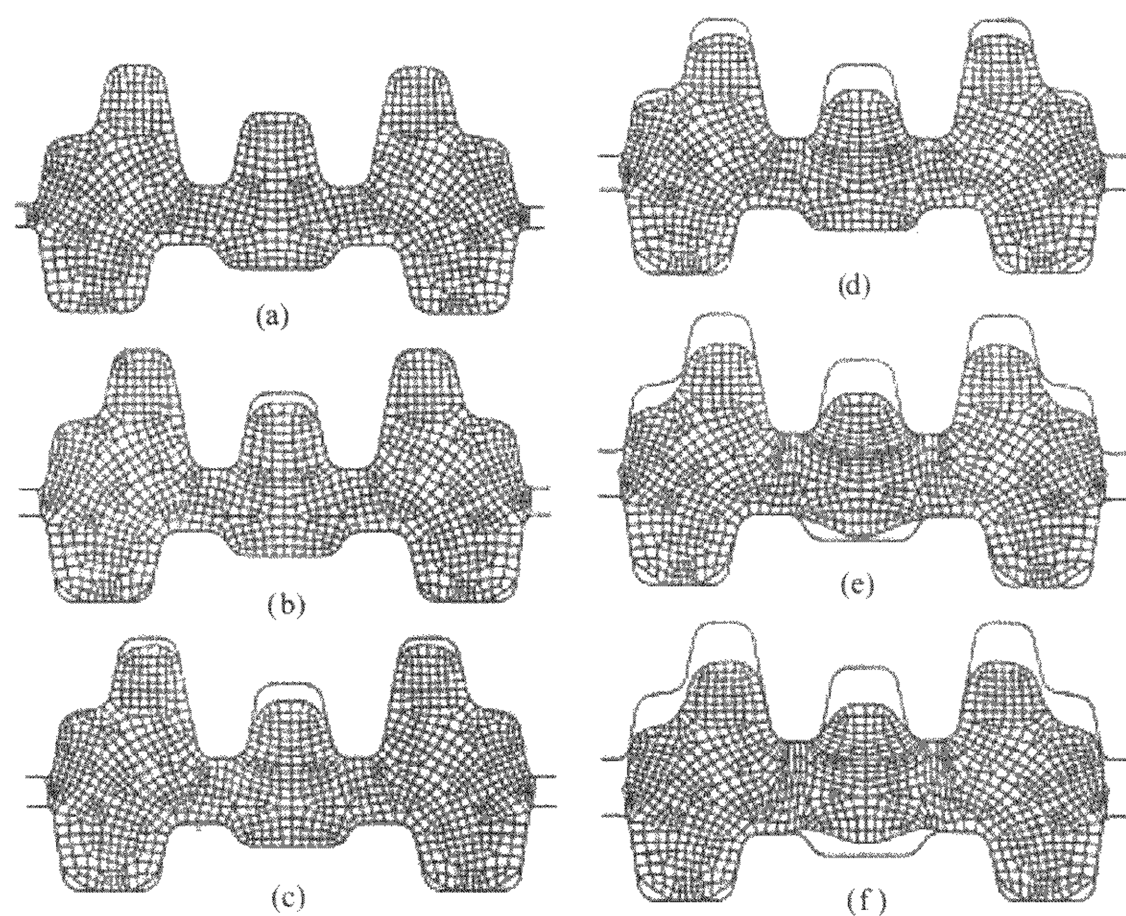


图 4.3-21　反向模拟和预锻件设计

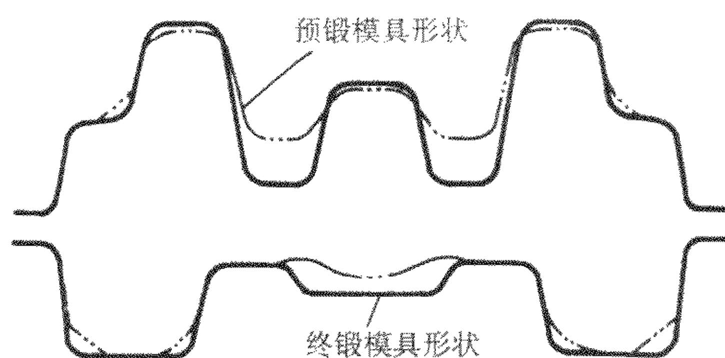


图 4.3-22　预锻件和终锻模具的形状

料分布已经通过预锻过程取得。这就使得两侧型腔的上半部分和中心型腔的上半部分基本上同时充满，从而在行程结束阶段变形载荷具有较大程度的减小。因此，通过对反向模拟方法得到预锻件形状的两工序锻造过程和单工序锻造过程的金属流动模式的比较，不难看出，锻造载荷、模具充满情况、飞边大小和能量要求都得到了较大程度的改善。

以上介绍了基于有限元数值模拟方法的反向模拟技术及其金属塑性成形的预成形设计方法，给出了几种判断节点脱离模具的准则，通过实例证实了有限元反向模拟技术的可行性。在反向模拟技术中，关键问题是边界条件的确定，因此，实用的判别准则是该项技术应用研究的重点。不同的边界条件控制准则，可得到不同的预成形设计结果，这种设计结果的不惟一性给寻求最佳的设计方案带来了一定困难。因此，具有普遍适用性的边界条件的研究开发显得尤为重要。

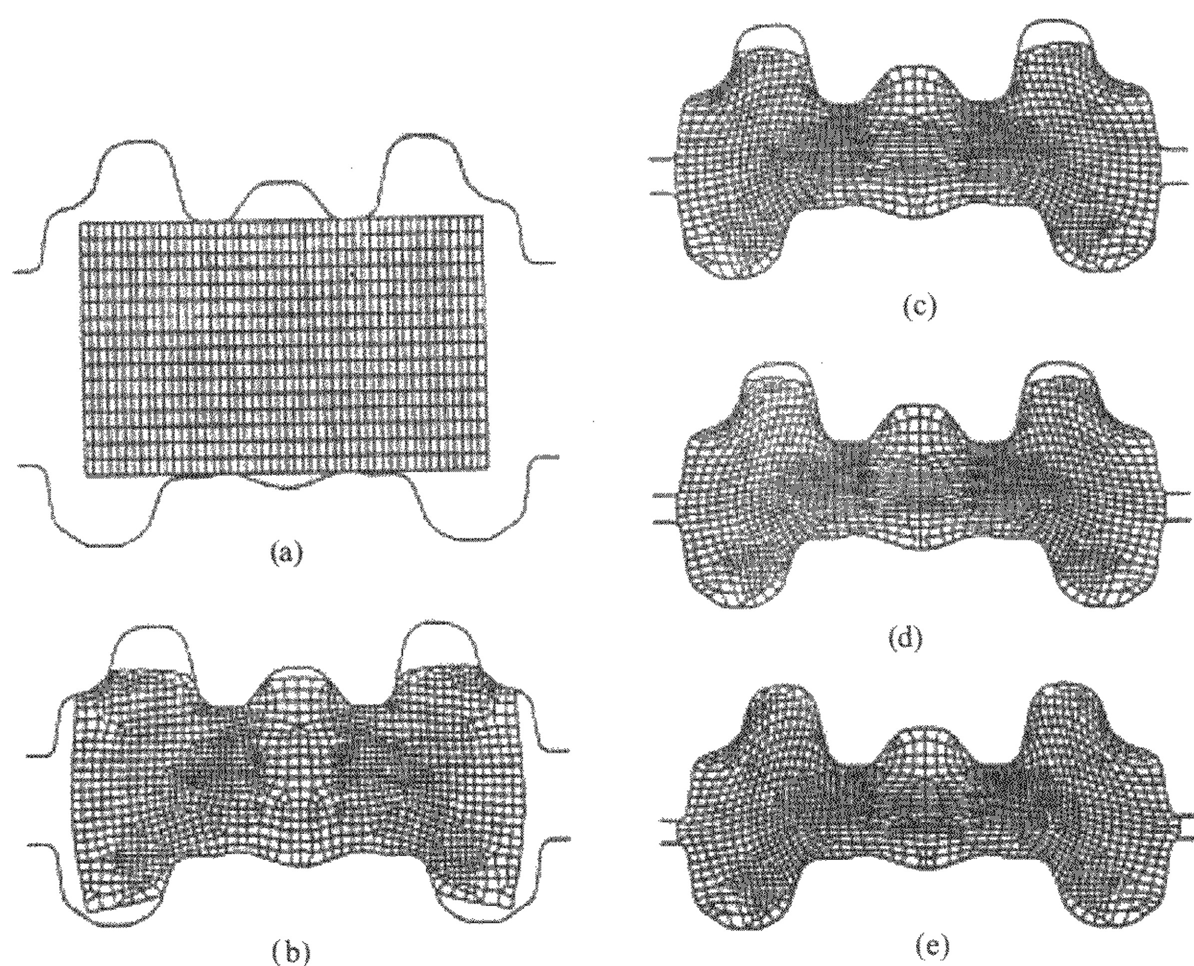


图 4.3-23　预锻过程的正向模拟验证

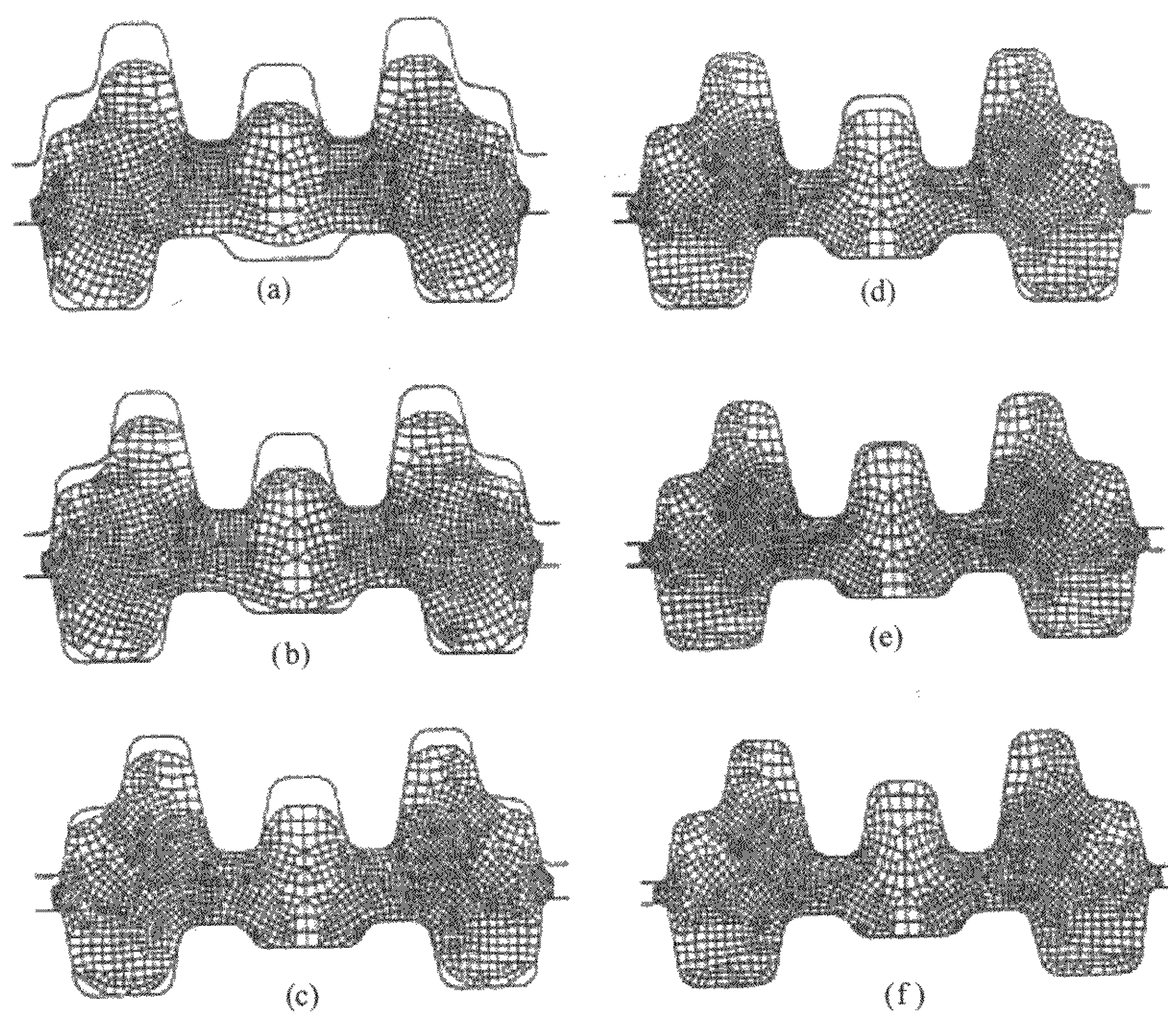


图 4.3-24　终锻过程的正向模拟验证

实施反向模拟与预成形设计的数值模拟技术，除有限元方法外，还有上限元法。基于上限元的反向模拟技术与基于有限元法的反向模拟技术在几种关键问题的处理方法上有相似之处，所不同的是在上限元法反向模拟技术中锻件与模具的边界轮廓被理想化为直边界，并假定在某一时刻某条边界脱模。从这一点讲，上限元法的反向模拟技术较之有限元法的反向模拟技术简单，但预测结果的精度较有限元反向模拟技术粗放。鉴于此，可以建立上限元与有限元的耦合正反向模拟技术，即首先应用上限元法反向模拟出预成形件，然后再用有限元正向模拟技术去验证，并由正向模拟结果来修正反向模拟的结果，直至预测的预成形件用正向有限元模拟时的结果符合设计要求。关于上限元反向模拟技术，有兴趣的读者可参阅相关文献。

金属塑性成形过程是将初始形状简单的坯料经过模具和设备成形为所要求形状的终成形件。从理论上讲，应有一条最佳的成形路径，与该成形路径对应的工艺参数、模具设计应当是我们寻求的设计方案。因此，将工程优化设计与金属塑性成形数值模拟技术相结合，进行塑性成形过程的优化设计至关重要。

编写：赵国群（山东大学）
管延锦（山东大学）

第4章 刚（黏）塑性有限元灵敏度分析与模具优化设计

随着计算机技术与有限元法的发展，形状优化方法已越来越多地应用于复杂结构工程的优化问题。在金属塑性成形生产中，产品的质量主要取决于模具设计的水平。金属塑性成形过程通常为非稳态大变形问题，因其成形件形状的复杂性，绝大多数成形件采用多工序加工，工序数目的确定和各工序模具形状的设计是材料成形工艺与模具设计的主要工作。如图4.4-1所示为一锻造过程，为获得所要求的终锻件，必须经过中间的预锻工步，如何设计预锻件形状和加工此预锻件所用预锻模具的形状是生产合格终锻件的关键，也是模具设计的困难问题之一。在实际生产中，设计者的经验和直觉一直是预成形设计的主要依据，其设计水平、时间和成本都受到一定限制。

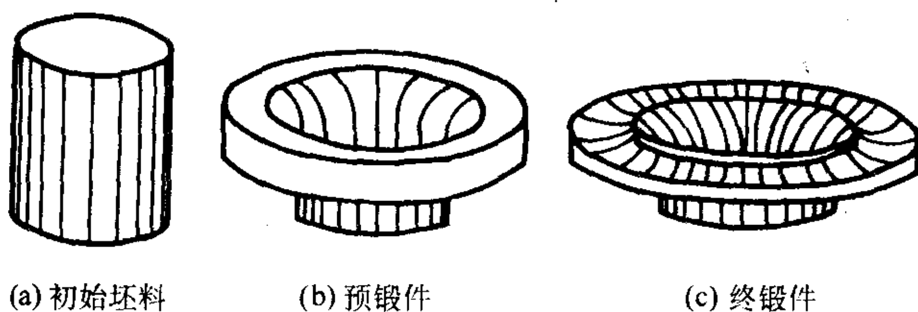

(a) 初始坯料　(b) 预锻件　(c) 终锻件

图4.4-1 典型多工序锻造过程

近年来，以有限元为主的材料塑性成形过程模拟技术为预成形设计提供了良好的数学方法。第3章介绍了有限元反向模拟方法及其关键技术，并给出了实际锻造问题的预成形设计。这种反向模拟方法是从完全充满终锻模腔的终锻件形状出发，以反向变形方式模拟材料变形规律，按照规定的边界条件控制准则，通过解除边界节点的约束条件而得到任意时刻的预成形件形状。然后，根据预成形件形状来设计预锻模具形状。

上述预成形设计方法都采用有限元正、反向模拟方法，都依赖于边界条件的控制准则。由于材料成形路径的多样化，建立通用的或者最佳的边界条件控制准则，仍存在较大困难。变形路径的多样化特征决定了寻求最佳预成形设计和建立具有全局优化特点的优化设计方法的可能性与必要性。

Badrinarayanan等提出了一种拟弹－黏塑性材料的有限元灵敏度分析方法，并应用于大变形过程的预成形优化设计，该方法设计目标是预成形件形状，而不是预成形模具形状，预成形模具形状则根据该预成形件形状设计。但设计出的预成形件形状通常不易成形，例如，圆柱体镦粗过程，为消除终锻件的侧面鼓形，设计出的预成形件侧面呈凹形，这就增加了预成形模具设计和实际生产的困难性。

赵国群等提出了一种基于刚（黏）塑性有限元数值模拟技术的灵敏度分析方法，并建立了相应的预成形模具形状优化设计方法与步骤。该方法以直接设计预成形模具形状为目标，应用有限元、优化和单纯正向模拟技术相结合的方法建立灵敏度分析理论和模具形状优化设计方法。本章主要介绍该方法的理论及其具体应用技术。

1 刚（黏）塑性有限元基本方程

为便于推导刚塑性、刚（黏）塑性有限元灵敏度方程，本节简单列出刚（黏）塑性有限元的基本方程，对于详细的刚塑性、刚（黏）塑性有限元理论与技术，可参考文献［9］。

对于金属塑性成形问题，经有限元离散化后，单元应变率可用下列矩阵形式表示：

$$\dot{\boldsymbol{\varepsilon}} = \boldsymbol{BV} \tag{4.4-1}$$

式中，$\boldsymbol{B}$为单元应变率矩阵；$\boldsymbol{V}$为单元的节点速度向量。

对于四边形四节点等参单元，单元应变率矩阵$\boldsymbol{B}$可表示为

$$\boldsymbol{B} = \begin{pmatrix} E_1 & 0 & E_2 & 0 & E_3 & 0 & E_4 & 0 \\ 0 & H_1 & 0 & H_2 & 0 & H_3 & 0 & H_4 \\ P_1 & 0 & P_2 & 0 & P_3 & 0 & P_4 & 0 \\ H_1 & E_1 & H_2 & E_2 & H_3 & E_3 & H_4 & E_4 \end{pmatrix} \tag{4.4-2}$$

式中，$E_\alpha = \partial N_\alpha / \partial x, H_\alpha = \partial N_\alpha / \partial y (\alpha = 1,2,3,4)$；$N_\alpha$为单元形状函数；$x, y$为单元内任意点的坐标。

对于平面变形问题，式（4.4-2）中的P_α（$\alpha = 1, 2, 3, 4$）等于零，对于轴对称变形问题，$P_\alpha = N_\alpha / r$，r为径向坐标，并表示为

$$r = \sum_{\alpha=1}^{4} N_\alpha x_\alpha \tag{4.4-3}$$

式中，$x_\alpha (\alpha = 1,2,3,4)$为单元四个节点的径向（横向）坐标。

刚（黏）塑性材料成形问题的有限元平衡方程式为

$$\boldsymbol{K}(\boldsymbol{V},\boldsymbol{X})\boldsymbol{V} + \boldsymbol{F}(\boldsymbol{V},\boldsymbol{X}) = 0 \tag{4.4-4}$$

其中

$$K_{ij}(\boldsymbol{V},\boldsymbol{X}) = \int_V \frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}} P_{ij} \mathrm{d}V + Q \int_V C_i C_j \mathrm{d}V \tag{4.4-5}$$

$$F_i(\boldsymbol{V},\boldsymbol{X}) = \int_{S_C} mk \frac{2}{\pi} N_i \arctan\left(\frac{N_j u_{s_j}}{u_o}\right) \mathrm{d}S \tag{4.4-6}$$

式中，K为与材料和过程有关的非线性刚度矩阵；F为施加的节点力向量；X为单元的节点坐标向量；Q为引入塑性变形体积不变条件的惩罚因子；$\bar{\sigma}$为等效应力；$\dot{\bar{\varepsilon}}$为等效应变率；m为常摩擦因子；k为剪切屈服应力。

在工件与模具接触面上的摩擦力用相对滑动速度的反正切函数表示，设v_{s_j}为模具与工件接触面上节点j的相对滑动速度，则滑动速度v_s可采用线性插值求得：

$$v_s = \sum_{j=1}^{2} N_j v_{s_j} \tag{4.4-7}$$

其中u_o是一个与v_s相比很小的正常数。矩阵$\boldsymbol{P}$与$\boldsymbol{C}$由应变率矩阵$\boldsymbol{B}$构造，其元素为

$$\boldsymbol{P}_{ij} = \boldsymbol{B}_{ik}\boldsymbol{D}_{kn}\boldsymbol{B}_{nj} \tag{4.4-8}$$

$$\boldsymbol{C}_i = \boldsymbol{B}_{1i} + \boldsymbol{B}_{2i} + \boldsymbol{B}_{3i} \tag{4.4-9}$$

$$\boldsymbol{C}_j = \boldsymbol{B}_{1j} + \boldsymbol{B}_{2j} + \boldsymbol{B}_{3j} \tag{4.4-10}$$

式中，$\boldsymbol{D}$为一对角线矩阵。

对于材料成形问题，平衡方程式（4.4-4）为一非线性方程组，其求解过程为：首先计算各单元的刚度阵和节点力向量；然后在整个工件上进行组装，形成刚度方程组；最后进行线性化，并采用Newton－Raphson迭代方法进行求解。

2 目标函数与设计变量

对于二维塑性成形问题，假设G_0和G分别表示最终成形件的理想形状和对应于任意一个预成形模具形状实际取得

的最终成形件形状。优化设计的目标是通过设计预成形模具形状使实际工件形状接近或等于理想工件形状。因此，G_0 与 G 之间的差别是预成形模具形状的函数，其大小反映了预成形模具形状设计的准确性。目标函数定义为两个工件形状不重合区域的面积，如图 4.4-2 所示。因此，优化问题陈述为通过设计预成形模具形状，使形状 G 接近于 G_0。

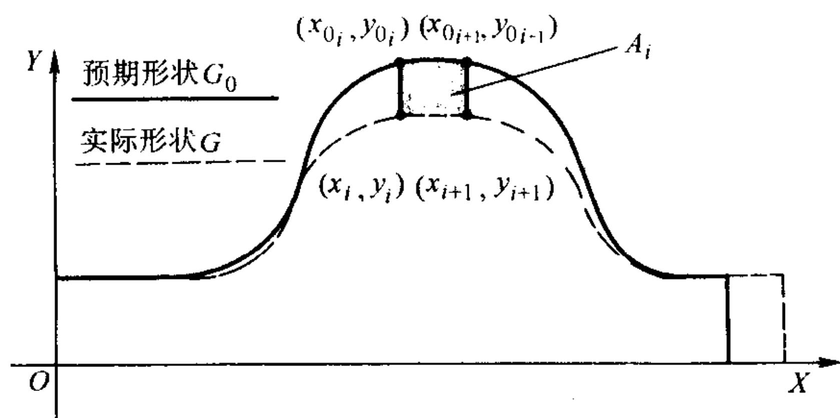


图 4.4-2　实际取得的终成形件与理想终成形件形状

对于一个经有限元离散化后的工件边界，假设所取得的实际工件 G 有 N 个边界节点，其坐标分别为 (x_i, y_i)，$i=1, 2, 3, \cdots, N$，通过这些边界节点 i 并沿其法向向理想工件边界引线，其交点构成了另一组在理想工件上的边界节点 (x_{0i}, y_{0i})，$i=1, 2, 3, \cdots, N$，连接这两组边界节点，在两个工件形状 G_0 与 G 不重合区域内构成了一组四边形单元，这些单元的节点按逆时针方向编号，如图 4.4-3a 示，单元面积为

$$A_j = \frac{1}{2}(x_{32}y_{12} - x_{12}y_{32}) + \frac{1}{2}(x_{14}y_{34} - x_{34}y_{14})$$

其中，$x_{ij} = x_i - x_j$，$y_{ij} = y_i - y_j$，$i, j = 1, 2, 3, 4$。

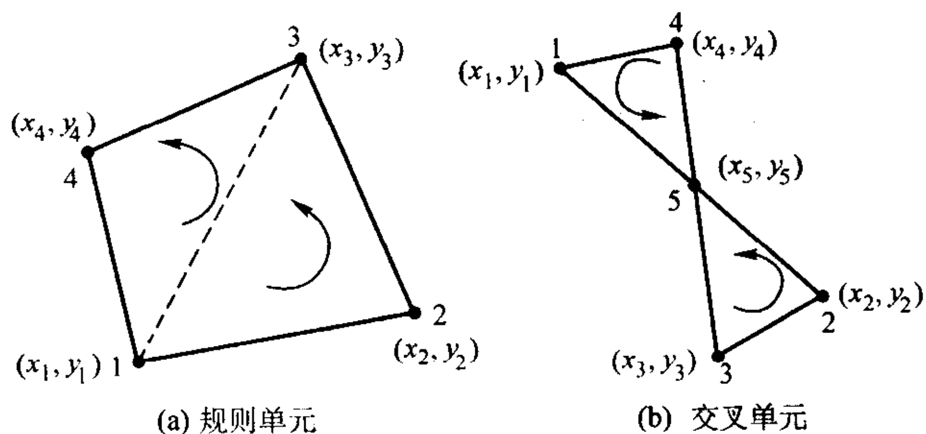


图 4.4-3　计算目标函数的单元

在特殊情况下，实际工件形状与理想工件形状相交，在边界上形成相交单元，如图 4.4-3b 所示。这种特殊单元的面积为

$$A_j = \frac{1}{2}(x_{35}y_{25} - x_{25}y_{35}) + \frac{1}{2}(x_{14}y_{54} - x_{54}y_{14})$$

其中 (x_5, y_5) 为交点坐标。

应用各单元的面积 A_j 的平方和作为目标函数 ψ，即

$$\psi = \sum_{j=1}^{N} A_j^2 \tag{4.4-11}$$

当 ψ 接近零时，实际获得的最终成形件形状 G 将与理想终成形件形状 G_0 趋于一致。优化问题定义为通过使目标函数 ψ 取最小值的 B 样条曲线控制点的坐标来定义最佳预成形模具形状。

采用三次 B 样条函数表示二维成形问题的预成形模具形状，变化 B 样条曲线的控制点坐标，可得到不同的 B 样条曲线形状。对于每个控制点，有两个自由度 (p_{x_i}, p_{y_i})，$i=1, 2, \cdots, K$，K 为定义预成形模具形状的 B 样条曲线的控制点总数。因此，共有 $2K$ 个设计变量，对这种无约束的优化设计问题，本书采用 BFGS 优化方法进行优化。

3　灵敏度分析

根据式（4.4-11）给出的目标函数表达式，推导目标函数 ψ 对优化设计变量 p_l 梯度如下：

$$\frac{\partial \psi}{\partial p_l} = \sum_{i=1}^{N} \frac{\partial \psi}{\partial x_i}\frac{\partial x_i}{\partial p_l} + \sum_{i=1}^{N} \frac{\partial \psi}{\partial y_i}\frac{\partial y_i}{\partial p_l}, \quad l = 1, 2, \cdots, 2k \tag{4.4-12}$$

式中，$\frac{\partial \psi}{\partial x_i} = 2\sum_{j=1}^{N} A_j \frac{\partial A_j}{\partial x_i}$，$\frac{\partial \psi}{\partial y_i} = 2\sum_{j=1}^{N} A_j \frac{\partial A_j}{\partial y_i}$。

对于两个最终成形件不重合区域内的一般性单元 j 有

$$\frac{\partial A_j}{\partial \chi_1} = \frac{1}{2}(y_{34} - y_{32})$$

$$\frac{\partial A_j}{\partial x_2} = \frac{1}{2}(y_{32} - y_{12})$$

$$\frac{\partial A_j}{\partial y_1} = \frac{1}{2}(x_{32} - x_{34})$$

$$\frac{\partial A_j}{\partial y_2} = \frac{1}{2}(x_{12} - x_{32})$$

对于特殊的交叉单元 j 有

$$\frac{\partial A_j}{\partial x_1} = \frac{1}{2}y_{54}, \frac{\partial A_j}{\partial y_1} = -\frac{1}{2}x_{54}, \frac{\partial A_j}{\partial x_2} = -\frac{1}{2}y_{35}, \frac{\partial A_j}{\partial y_2} = \frac{1}{2}x_{35}$$

在有限元速度场迭代收敛后，应用收敛的速度场刷新该时间增量内的节点坐标，其刷新公式为

$$\boldsymbol{X}^{(t+\Delta t)} = \boldsymbol{X}^{(t)} + \boldsymbol{V}^{(t)}\Delta t \tag{4.4-13}$$

式中，$\boldsymbol{X}^{(t+\Delta t)}$ 是 $(t+\Delta t)$ 时刻节点的坐标向量，$\boldsymbol{X}^{(t)}$ 表示 t 时刻节点的坐标向量，$\boldsymbol{V}^{(t)}$ 表示 t 时刻节点的速度向量，将式（4.4-13）对优化设计变量 p_1 求偏导，得

$$\frac{\partial \boldsymbol{X}^{(t+\Delta t)}}{\partial p_1} = \frac{\partial \boldsymbol{X}^{(t)}}{\partial p_1} + \frac{\partial \boldsymbol{V}^{(t)}}{\partial p_1}\Delta t \tag{4.4-14}$$

上式即为有限元节点坐标对优化设计变量的灵敏度刷新公式，当 $t=0$ 时，有 $\partial \boldsymbol{X}^{(0)}/\partial p_1 = 0$。

可见，一旦确定节点速度对优化设计变量的灵敏度 $\partial \boldsymbol{V}^{(t)}/\partial p_1$，即可根据式（4.4-14）确定出节点坐标对优化设计变量的灵敏度，进而由式（4.4-12）计算出目标函数对优化设计变量的灵敏度。

3.1　节点速度灵敏度

对于任意一个单元，t 时刻的节点速度灵敏度向量 $\frac{\partial \boldsymbol{V}^{(t)}}{\partial p_l}$，简记为 $\boldsymbol{V}_{,p_l}$，可通过将单元平衡方程式（4.4-4）对优化设计变量求偏导得：

$$\left(\frac{\partial \boldsymbol{K}}{\partial \boldsymbol{V}}\boldsymbol{V} + \boldsymbol{K} + \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{V}}\right)\frac{\partial \boldsymbol{V}}{\partial p_l} = -\left(\frac{\partial \boldsymbol{K}}{\partial \boldsymbol{X}}\frac{\partial \boldsymbol{X}}{\partial p_l}\boldsymbol{V} + \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{X}}\frac{\partial \boldsymbol{X}}{\partial p_l}\right) \tag{4.4-15a}$$

上式简写为

$$\boldsymbol{R}\boldsymbol{V}_{,p_l} = \boldsymbol{F}_{,p_l} \tag{4.4-15b}$$

式中，$\boldsymbol{R}$ 为单元的刚度灵敏度矩阵，$\boldsymbol{F}_{,p_l}$ 即 $\partial \boldsymbol{F}/\partial p_l$，是单元的节点力灵敏度向量，根据式（4.4-15a）、式（4.4-5）和式（4.4-6），经推导得出矩阵 $\boldsymbol{R}$ 与向量 $\boldsymbol{F}_{,p_l}$ 的计算公式如下：

$$R_{ij} = \sum_{n=1}^{8}\sum_{m=1}^{8}\int_V \left(\frac{1}{\dot{\bar{\varepsilon}}}\frac{\partial \bar{\sigma}}{\partial \dot{\bar{\varepsilon}}} - \frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}^2}\right)\frac{1}{\dot{\bar{\varepsilon}}}P_{in}v_n v_m P_{mj}\,\mathrm{d}V + K_{ij} + \frac{\partial F_i}{\partial v_j}$$

$$(i, j = 1, 2, \cdots, 8) \tag{4.4-16a}$$

$$F_{i,p_l} = -\left(\sum_{n=1}^{8}\frac{\partial F_i}{\partial x_n}\frac{\partial x_n}{\partial p_l} + \sum_{j=1}^{8}\sum_{n=1}^{8}\frac{\partial K_{ij}}{\partial x_n}\frac{\partial x_n}{\partial p_l}v_j\right)$$
$$(i=1,2,\cdots,8) \qquad (4.4\text{-}16b)$$

当用式（4.4-16a）和式（4.4-16b）计算出所有单元的刚度灵敏度矩阵和节点力灵敏度向量后，在整个工件上，对所有单元进行组装，得到一组关于节点速度灵敏度的线性方程：

$$\bar{\boldsymbol{R}}\,\bar{\boldsymbol{V}}_{,p_l} = \bar{\boldsymbol{F}}_{,p_l} \qquad (4.4\text{-}17)$$

式中，$\bar{\boldsymbol{R}}$ 为整体刚度灵敏度矩阵；$\bar{\boldsymbol{V}}_{,p_l}$ 为整体节点速度灵敏度向量；$\bar{\boldsymbol{F}}_{,p_l}$ 为整体节点力灵敏度向量。

对于给定的一个初始预成形模具形状，首先进行预成形工序的有限元模拟，在每一时间增量模拟步内，进行速度场迭代，当速度场收敛后，按照式（4.4-17）求解节点速度灵敏度，然后，利用式（4.4-14）刷新节点坐标灵敏度。当整个变形过程（包括预成形与终成形工步）结束时，根据求得的最终节点坐标灵敏度、实际求得的终成形件形状与理想终成形件形状，利用式（4.4-11）与式（4.4-12）分别求出目标函数值及其梯度值。最后，调用优化工具进行优化迭代，判断或检查优化迭代是否已达到所规定的优化条件，若还未满足优化条件，则利用得到的新控制点坐标刷新预成形模具形状，进行下一次优化迭代，直至达到所要求的优化条件。

3.2 $\dfrac{\partial K_{ij}}{\partial x_n}$，$\dfrac{\partial F_i}{\partial x_n}$，$\dfrac{\partial F_i}{\partial v_j}$的确定

根据式（4.4-5）与式（4.4-6），经推导可得式（4.4-15a）中的偏导数项$\frac{\partial K_{ij}}{\partial x_n}$，$\frac{\partial F_i}{\partial x_n}$和$\frac{\partial F_i}{\partial v_j}$的计算公式如下：

$$\frac{\partial K_{ij}}{\partial x_n} = \int_V\left(\frac{1}{\dot{\bar{\varepsilon}}}\frac{\partial\bar{\sigma}}{\partial\dot{\bar{\varepsilon}}} - \frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}^2}\right)\frac{1}{2\dot{\bar{\varepsilon}}}\frac{\partial(\boldsymbol{V}^{\mathrm T}\boldsymbol{P}\boldsymbol{V})}{\partial x_n}P_{ij}\mathrm dV + \int_V\frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}}\frac{\partial P_{ij}}{\partial x_n}\mathrm dV + Q\int_V\frac{\partial C_i}{\partial x_n}C_j\mathrm dV + Q\int_V C_i\frac{\partial C_j}{\partial x_n}\mathrm dV + \int_V\frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}}P_{ij}\frac{\partial(\mathrm dV)}{\partial x_n} + Q\int_V C_iC_j\frac{\partial(\mathrm dV)}{\partial x_n} \qquad (4.4\text{-}18)$$

$$\frac{\partial F_i}{\partial x_n} = \sum_{a=1}^{8}\sum_{k=1}^{8}\int_S\frac{m}{\sqrt{3}\pi}\frac{\partial\bar{\sigma}}{\partial\dot{\bar{\varepsilon}}}N_i\frac{1}{\dot{\bar{\varepsilon}}}v_a\frac{\partial P_{ak}}{\partial x_n}v_k\arctan\left(\frac{N_jv_{s_j}}{u_o}\right)\mathrm dS + \int_{S_C}mk\frac{2}{\pi}N_i\arctan\left(\frac{N_jv_{s_j}}{u_o}\right)\frac{\partial(\mathrm dS)}{\partial x_n} \qquad (4.4\text{-}19)$$

$$\frac{\partial F_i}{\partial v_j} = \int_S mk\frac{2}{\pi}N_iN_j\frac{u_o}{u_o^2+(N_kv_{s_k})^2}\mathrm dS \qquad (4.4\text{-}20)$$

其中，式（4.4-18）和式（4.4-19）中的$\frac{\partial C_i}{\partial x_n}$，$\frac{\partial C_j}{\partial x_n}$，$\frac{\partial P_{ij}}{\partial x_n}$与$\frac{\partial(\boldsymbol{V}^{\mathrm T}\boldsymbol{P}\boldsymbol{V})}{\partial x_n}$的计算公式可由式（4.4-8）、式(4.4-9）和式（4.4-10）推导得出如下：

$$\frac{\partial C_i}{\partial x_n} = \frac{\partial B_{1i}}{\partial x_n} + \frac{\partial B_{2i}}{\partial x_n} + \frac{\partial B_{3i}}{\partial x_n} \qquad (4.4\text{-}21a)$$

$$\frac{\partial C_j}{\partial x_n} = \frac{\partial B_{1j}}{\partial x_n} + \frac{\partial B_{2j}}{\partial x_n} + \frac{\partial B_{3j}}{\partial x_n} \qquad (4.4\text{-}21b)$$

$$\frac{\partial P_{ij}}{\partial x_n} = \sum_{k=1}^{4}\sum_{m=1}^{4}\left(\frac{\partial B_{ki}}{\partial x_n}D_{km}B_{mj} + B_{ki}D_{km}\frac{\partial B_{mj}}{\partial x_n}\right) \qquad (4.4\text{-}21c)$$

$$\frac{\partial(\boldsymbol{V}^{\mathrm T}\boldsymbol{P}\boldsymbol{V})}{\partial x_n} = \sum_{m=1}^{8}\sum_{k=1}^{8}v_m\frac{\partial P_{mk}}{\partial x_n}v_k \qquad (4.4\text{-}21d)$$

其中，导数矩阵$\frac{\partial\boldsymbol{B}}{\partial\boldsymbol{X}}$可由 $\boldsymbol{B}$ 矩阵式（4.4-2），通过计算$\frac{\partial E_i}{\partial x_j}$，$\frac{\partial H_i}{\partial y_j}$，$\frac{\partial E_i}{\partial y_j}$，$\frac{\partial H_i}{\partial x_j}$而确定，$\frac{\partial\boldsymbol{B}}{\partial\boldsymbol{X}}$矩阵中的诸元素计算公式为

$$\frac{\partial E_i}{\partial x_j} = -E_iE_j$$
$$\frac{\partial H_i}{\partial y_j} = -H_iH_j$$
$$\frac{\partial E_i}{\partial y_j} = -A_{ij}/(8|J|) - E_iH_j$$
$$\frac{\partial H_i}{\partial x_j} = -A_{ij}/(8|J|) - H_iE_j$$
$$\frac{\partial P_i}{\partial x_j} = -N_iN_j/r^2$$
$$\frac{\partial P_i}{\partial y_j} = 0$$

式中，i，$j=1$，2，3，4，$|\boldsymbol{J}|$为坐标转换矩阵的雅可比行列式的值，其表达式为

$$|\boldsymbol{J}| = \frac{1}{8}[(x_{13}y_{24}-x_{24}y_{13}) + (x_{34}y_{12}-x_{12}y_{34})\xi + (x_{23}y_{14}-x_{14}y_{23})\eta] \qquad (4.4\text{-}22)$$

对于平面变形问题，$\frac{\partial P_i}{\partial x_j} = \frac{\partial P_i}{\partial y_j} = 0$。

经推导，A_{ij}的计算公式由下列矩阵 $\boldsymbol{A}$ 给出：

$$\boldsymbol{A} = \begin{pmatrix} 0 & \eta-1 & \xi-\eta & 1-\xi \\ 1-\eta & 0 & -(1+\xi) & \xi+\eta \\ \eta-\xi & 1+\xi & 0 & -(1+\eta) \\ \xi-1 & -(\xi+\eta) & 1+\eta & 0 \end{pmatrix}$$

可见，$\boldsymbol{A}^{\mathrm T} = -\boldsymbol{A}$，故矩阵 $\boldsymbol{A}$ 为一反对称矩阵，其中（ξ，η）为自然坐标。

应注意，式（4.4-18）、式（4.4-19）和式（4.4-20）中的体积积分 $\mathrm dV$ 与面积积分 $\mathrm dS$ 也取决于节点坐标向量 $\boldsymbol{X}$。

对于平面变形问题有：

$$\frac{\partial(\mathrm dS)}{\partial x_n} = \frac{\partial|\boldsymbol{J}_{\mathrm S}|}{\partial x_n}\mathrm d\xi$$
$$\frac{\partial(\mathrm dV)}{\partial x_n} = \frac{\partial|\boldsymbol{J}|}{\partial x_n}\mathrm d\xi\mathrm d\eta$$

对于轴对称问题有：

$$\frac{\partial(\mathrm dS)}{\partial x_n} = \left(|\boldsymbol{J}_{\mathrm S}|\frac{\partial r}{\partial x_n} + r\frac{\partial|\boldsymbol{J}_{\mathrm S}|}{\partial x_n}\right)\mathrm d\xi$$
$$\frac{\partial(\mathrm dV)}{\partial x_n} = \left(|\boldsymbol{J}|\frac{\partial r}{\partial x_n} + r\frac{\partial|\boldsymbol{J}|}{\partial x_n}\right)\mathrm d\xi\mathrm d\eta$$

式中，$|\boldsymbol{J}_{\mathrm S}|$为在工件与模具接触面上坐标转换矩阵的雅可比行列式的值，r 是由式（4.4-3）定义的插值点的径向坐标。

4 速度灵敏度边界条件

在工件与模具接触面上，接触节点 i 的速度边界条件（参见图 4.4-4）为

$$v_{i_n} = \boldsymbol{V}_{\mathrm{die}}^{\mathrm T}\cdot\boldsymbol{n} = (v_{\mathrm{die}_x}, v_{\mathrm{die}_y})\begin{pmatrix}-\sin\beta \\ \cos\beta\end{pmatrix} \qquad (4.4\text{-}23)$$

式中，v_{i_n} 为节点 i 在接触面法线方向的速度分量；$\boldsymbol{V}_{\mathrm{die}}$为模具速度向量；$\boldsymbol{n}$ 为节点 i 处的模具表面的单位法矢量；β 为全球坐标系（xoy）的 x 轴与模具表面切向与法向组成的局部坐标系（$x'o'y'$）的 x' 轴之间的夹角。

对于一个由 $y=y(x)=y(p_l)$，($l=1$，2，3，$\cdots 2K$）定义的 B 样条函数，其斜率 $\mathrm dy/\mathrm dx = y_x = \tan\beta$，故由此可导出下列关系：

$$\sin\beta = y_x/\sqrt{1+{y_x}^2} \qquad (4.4\text{-}24)$$
$$\cos\beta = 1/\sqrt{1+{y_x}^2}$$

将式（4.4-23）两边对优化设计变量 p_1 求导可得速度灵敏度边界条件：

$$\frac{\partial v_{i_n}}{\partial p_1} = \boldsymbol{V}_{\mathrm{die}}^{\mathrm{T}}\frac{\partial \boldsymbol{n}}{\partial p_1} = -v_{\mathrm{die}_x}\frac{\partial \sin\beta}{\partial p_1} + v_{\mathrm{die}_y}\frac{\partial \cos\beta}{\partial p_1} \qquad (4.4\text{-}25)$$

式中，$\frac{\partial \boldsymbol{n}}{\partial p_1}$为模具表面法线方向单位矢量对设计变量 p_1 的灵敏度。

将式（4.4-24）代入式（4.4-25）得：

$$\frac{\partial \boldsymbol{n}}{\partial p_1} = \begin{pmatrix} -\dfrac{\sin\beta}{\partial p_1} \\ \dfrac{\partial\cos\beta}{\partial p_1} \end{pmatrix} = \begin{pmatrix} -\dfrac{y_{xp_1}}{(1+{y_x}^2)^{3/2}} \\ -\dfrac{y_{xp_1}y_x}{(1+{y_x}^2)^{3/2}} \end{pmatrix} \qquad (4.4\text{-}26)$$

式中，$y_{xp_1} = \frac{\partial y_x}{\partial p_1}$为预成形模具表面斜率对设计变量的灵敏度。

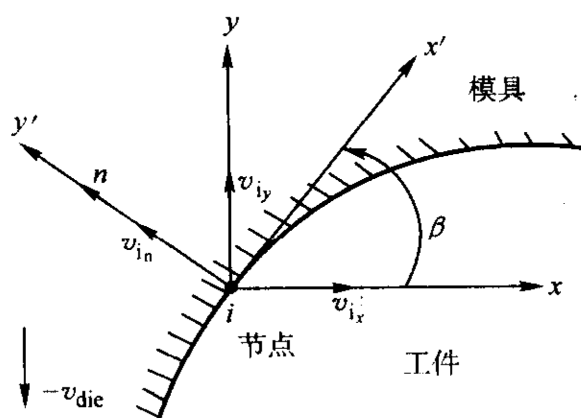


图 4.4-4　接触节点的速度灵敏度边界条件

由式（4.4-25）和式（4.4-26）可知，只要确定 y_x 和 y_{xp_1}，即可确定节点速度灵敏度的边界条件。其中，y_x 和 y_{xp_1} 由 B 样条函数确定。

给定三次 B 样条函数的控制点坐标$(p_{x_i}, p_{y_i})(i=1, 2, \cdots, K)$，则设计变量为 $p=(p_1, p_2, \cdots, p_{2K})=(p_{x_1}, p_{y_1}, \cdots, p_{x_K}, p_{y_K})$，第 i 个三次 B 样条曲线上的(x, y)坐标可用参数方程表示为：

$$x = \frac{1}{6}[(1-s)^3 p_{x_i} + (4-6s^2+3s^3)p_{x_{i+1}} + (1+3s+3s^2-3s^3)p_{x_{i+2}} + s^3 p_{x_{i+3}}] \qquad (4.4\text{-}27a)$$

$$y = \frac{1}{6}[(1-s)^2 p_{y_i} + (4-6s^2+3s^3)p_{y_{i+1}} + (1+3s+3s^2-3s^3)p_{y_{i+2}} + s^3 p_{y_{i+3}}] \qquad (4.4\text{-}27b)$$

式中，$0 \leqslant s \leqslant 1$。

对 B 样条函数求导得：

$$y_x = \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial y}{\partial s}\Big/\frac{\partial x}{\partial s} \qquad (4.4\text{-}28)$$

式中$\partial y/\partial s$，$\partial x/\partial s$ 分别由式(4.4-27a)和式(4.4-27b)导出如下：

$$\frac{\partial y}{\partial s} = \frac{1}{2}[-(1-s)^2 p_{y_i} + (-4s+3s^2)p_{y_{i+1}} + (1+2s-3s^2)p_{y_{i+2}} + s^2 p_{y_{i+3}}] \qquad (4.4\text{-}29a)$$

$$\frac{\partial x}{\partial s} = \frac{1}{2}[-(1-s)^2 p_{x_i} + (-4s+3s^2)p_{x_{i+1}} + (1+2s-3s^2)p_{x_{i+2}} + s^2 p_{x_{i+3}}] \qquad (4.4\text{-}29b)$$

由式（4.4-28）和式（4.4-29），即可很容易地推导出 B 样条曲线斜率对设计变量的偏导数 y_{xpl}，即$\frac{\partial y_x}{\partial p_l}$。

$$\frac{\partial^2 y}{\partial x \partial p_{y_i}} = -\frac{(1-s)^2}{2\partial x/\partial s} \qquad (4.4\text{-}30a)$$

$$\frac{\partial^2 y}{\partial x \partial p_{y_{i+1}}} = \frac{-4s+3s^2}{2\partial x/\partial s} \qquad (4.4\text{-}30b)$$

$$\frac{\partial^2 y}{\partial x \partial p_{y_{i+2}}} = \frac{1+2s-3s^2}{2\partial x/\partial s} \qquad (4.4\text{-}31a)$$

$$\frac{\partial^2 y}{\partial x \partial p_{y_{i+3}}} = \frac{s^2}{2\partial x/\partial s} \qquad (4.4\text{-}31b)$$

$$\frac{\partial^2 y}{\partial x \partial p_{x_i}} = \frac{(1-s)^2}{2(\partial x/\partial s)^2}\frac{\partial y}{\partial s} \qquad (4.4\text{-}32a)$$

$$\frac{\partial^2 y}{\partial x \partial p_{x_{i+1}}} = -\frac{-4s+3s^2}{2(\partial x/\partial s)^2}\frac{\partial y}{\partial s} \qquad (4.4\text{-}32b)$$

$$\frac{\partial^2 y}{\partial x \partial p_{x_{i+2}}} = -\frac{1+2s-3s^2}{2(\partial x/\partial s)^2}\frac{\partial y}{\partial s} \qquad (4.4\text{-}33a)$$

$$\frac{\partial^2 y}{\partial x \partial p_{x_{i+3}}} = -\frac{s^2}{2(\partial x/\partial s)^2}\frac{\partial y}{\partial s} \qquad (4.4\text{-}33b)$$

将由式（4.4-28）和式（4.4-29）给出的 y_x 和由式（4.4-28）和式（4.4-29）推导出的 y_{xpl} [即式（4.4-30）~式（4.4-33）] 代入方程式（4.4-26），即可求得模具表面法向矢量对设计变量 p_l 的灵敏度$\frac{\partial \boldsymbol{n}}{\partial p_l}$。然后将$\frac{\partial \boldsymbol{n}}{\partial p_l}$再代入式（4.4-25），即可确定节点速度灵敏度的边界条件。

对于上下模具对称的情况，只需取上模进行分析即可，由式（4.4-25）计算的速度灵敏度边界条件不为零。单对于上下模具不对称的情况，需要同时对上下模具进行优化，由于下模速度为零，由式（4.4-25）计算的边界节点对所有的下模优化设计变量的速度灵敏度边界条件为零。而 $t=0$ 时，$\partial \boldsymbol{X}^{(0)}/\partial p_1 = 0$，故 $t=0$ 时，式（4.4-15b）式中每个单元节点力灵敏度向量 $\boldsymbol{F}_{,p_1}$ 的元素均为零。组装成的整体节点力灵敏度向量 $\bar{\boldsymbol{F}}_{,p_1}$ 全为零。因此，整体灵敏度刚度方程式（4.4-17）成为：

$$\bar{\boldsymbol{R}}\begin{Bmatrix} \partial v_1/\partial p_l \\ \vdots \\ \partial v_{i_n}/\partial p_l \\ \vdots \\ \partial v_n/\partial p_l \end{Bmatrix} = \begin{Bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{Bmatrix}$$

其中，上述方程右侧向量的元素全为零，而速度灵敏度边界条件也为零。显然求解上述灵敏度刚度方程，得到的节点速度灵敏度为零。根据式（4.4-14）的节点坐标灵敏度刷新公式，刷新后的节点坐标灵敏度仍为零，故模拟结束后，目标函数对下模 B 样条控制点的梯度为零，因而下模形状得不到优化。解决该问题得方法之一是给下模施加一个速度，使其速度灵敏度边界条件不为零。因为在有限元模拟中，只要上下模具的相对速度一定，模拟结果就相同，所以给下模一个向上的速度，不会影响模拟的结果，却实现了对下模形状的优化。下模的速度一般取和上模的速度大小相等，方向相反。

5　温度场灵敏度分析

刚（黏）塑性有限元中的热平衡方程可用下式表示

$$\boldsymbol{C}\dot{\boldsymbol{T}} + \boldsymbol{K}_C \boldsymbol{T} = \boldsymbol{Q} \qquad (4.4\text{-}34)$$

其中，热容矩阵 $\boldsymbol{C} = \int_V \rho c N N^{\mathrm{T}} \mathrm{d}V \qquad (4.4\text{-}35)$

传热矩阵 $K_C = \int_V \lambda MM^T dV$ (4.4-36)

热流矢量 $Q = \int_V \kappa(\overline{\sigma}\dot{\overline{\varepsilon}}) N dV + \int_{S_r} \sigma\varepsilon(T_e^4 - T_s^4) N dS + \int_{S_t} h(T_e - T_s) N dS + \int_{S_C} h_{lub}(T_d - T_w) N dS + \int_{S_C} q_f N dS$ (4.4-37)

节点温度对设计变量的灵敏度,可通过式(4.4-34)两边对设计变量 p_l 求偏导得到。热平衡方程左边各式对设计变量求偏导,可得到以下方程:

$$\frac{\partial(C\dot{T})}{\partial p_l} = \frac{\partial \dot{T}}{\partial p_l}\int \rho c NN^T dV + \dot{T}\int \rho c NN^T \frac{\partial(dV)}{\partial p_l} \quad (4.4\text{-}38)$$

$$\frac{\partial(K_C T)}{\partial p_l} = \frac{\partial T}{\partial p_l}\int k_1 MM^T dV + T\int k_1 \frac{\partial M}{\partial p_l} M^T dV + T\int k_1 M\frac{\partial M^T}{\partial p_l} dV + T\int k_1 MM^T \frac{\partial(dV)}{\partial p_l} \quad (4.4\text{-}39)$$

热流矢量 Q 各项对设计变量求偏导得:

$$\frac{\partial\left[\int_V \kappa(\overline{\sigma}\dot{\overline{\varepsilon}}) N dV\right]}{\partial p_l} = \int_V \kappa\frac{\partial\overline{\sigma}}{\partial p_l}\dot{\overline{\varepsilon}} N dV + \int_V \kappa\overline{\sigma}\frac{\partial\dot{\overline{\varepsilon}}}{\partial p_l} N dV + \int_V \kappa(\overline{\sigma}\dot{\overline{\varepsilon}}) N\frac{\partial(dV)}{\partial p_l} \quad (4.4\text{-}40a)$$

$$\frac{\partial\left[\int_{S_r} \sigma\varepsilon(T_e^4 - T_s^4) N dS\right]}{\partial p_l} = -\int_{S_r} 4\sigma\varepsilon T_s^3\frac{\partial T_s}{\partial p_l} N dS + \int_{S_r} \sigma\varepsilon(T_e^4 - T_s^4) N\frac{\partial(dS)}{\partial p_l} \quad (4.4\text{-}40b)$$

$$\frac{\partial\left[\int_{S_t} h(T_e - T_s) N dS\right]}{\partial p_l} = -\int_{S_t} h\frac{\partial T_s}{\partial p_l} N dS + \int_{S_t} h(T_e - T_s) N\frac{\partial(dS)}{\partial p_l} \quad (4.4\text{-}40c)$$

$$\frac{\partial\left[\int_{S_C} h_{lub}(T_d - T_w) N dS\right]}{\partial p_l} = \int_{S_C} h_{lub}\frac{\partial T_d}{\partial p_l} N dS - \int_{S_C} h_{lub}\frac{\partial T_w}{\partial p_l} N dS + \int_{S_C} h_{lub}(T_d - T_w) N\frac{\partial(dS)}{\partial p_l} \quad (4.4\text{-}40d)$$

$$\frac{\partial\left(\int_{S_C} q_f N dS\right)}{\partial p_l} = \frac{\partial\left[\int_{S_C}\frac{2m}{\pi}\frac{\overline{\sigma}}{\sqrt{3}} q_i \arctan\left(\frac{q_j v_{sj}}{u_0}\right) N dS\right]}{\partial p_l} = \int_{S_C}\frac{2m}{\pi\sqrt{3}}\frac{\partial\overline{\sigma}}{\partial p_l} q_i \arctan\left(\frac{q_j v_{sj}}{u_0}\right) N dS + \int_{S_C}\frac{2m}{\pi}\frac{\overline{\sigma}}{\sqrt{3}} q_i \arctan\left(\frac{q_j v_{sj}}{u_0}\right) N\frac{\partial(dS)}{\partial p_l} \quad (4.4\text{-}40e)$$

以上等式每项可分成三部分。包含 $\partial\dot{T}/\partial p_l$ 部分加到新的热容矩阵 $\overline{C}$,包含 $\partial T/\partial p_l$ 部分加到新的热传导矩阵 $\overline{K}_C$,不包含 $\partial\dot{T}/\partial p_l$ 和 $\partial T/\partial p_l$ 项的部分加到新的热流矢量 $\overline{Q}$ 中。这样,新的平衡方程可写成下面的形式:

$$\left(\overline{K}_C + \frac{\overline{C}}{\beta\Delta t}\right)\frac{\partial T_{\Delta t}}{\partial p_l} = \overline{Q}_{\Delta t} - \overline{C}\frac{\partial\hat{T}}{\partial p_l} \quad (4.4\text{-}41)$$

其中

$$\frac{\partial\hat{T}}{\partial p_l} = -\frac{1}{\beta\Delta t}\frac{\partial T_0}{\partial p_l} - \left(\frac{1-\beta}{\beta}\right)\frac{\partial\dot{T}_0}{\partial p_l} \quad (4.4\text{-}42)$$

节点温度对设计变量灵敏度与节点温度速率对设计变量的灵敏度关系如下:

$$\frac{\partial T_{t+\Delta t}}{\partial p_l} = \frac{\partial T_t}{\partial p_l} + \Delta t\left[(1-\beta)\frac{\partial\dot{T}_t}{\partial p_l} + \beta\frac{\partial\dot{T}_{t+\Delta t}}{\partial p_l}\right] \quad (4.4\text{-}43)$$

对式(4.4-41)在整个工件上进行组装,将得到一组线性方程组,求解该方程组就可得工件节点温度对设计变量的灵敏度值。但从式(4.4-40d)可以看出,在计算热流矢量对设计变量的偏导时,有一项 $\partial T_d/\partial p_l$,即模具节点温度对设计变量的灵敏度需要计算,因此,不仅要对工件进行温度场灵敏度分析,而且需要对模具进行温度场灵敏度分析。与计算工件和模具温度场相类似,在计算工件和模具温度灵敏度时,工件和模具应分别对待,即计算工件温度灵敏度时,认为模具温度灵敏度值不变,同样,在计算模具温度灵敏度时,认为工件温度灵敏度值不变。计算模具和工件节点温度灵敏度的方法是相同的,只不过模具没有塑性变形功,其热流矢量表达式中没有该项罢了。计算工件和模具节点温度灵敏度需要迭代的方式,直到两者都收敛为止。

6　模具优化设计方法的应用技术

非稳态材料成形过程的形状优化设计研究对于提高模具设计水平和有效控制产品质量至关重要,前面给出了基于有限元灵敏度分析的模具优化设计方法的基本理论和方程。形状优化设计的目标函数对设计变量的灵敏度依赖于节点坐标、节点速度对设计变量的灵敏度,其中节点坐标灵敏度是一个与变形历史有关的量,它是整个变形过程的累加值。该方法对整个变形过程进行优化,因此上述方法是一种全局优化设计。所寻求的最佳模具形状以三次B样条函数表示,故设计出的模具形状光滑,易于加工,适合于材料塑性成形生产。所有的解析灵敏度方程是以有限元基本方程为出发点建立的,当目标函数发生改变时,灵敏度方程仍然适用,故对于材料成形过程具有一定的通用性。下面主要介绍该方法的优化设计步骤、具体应用技术及其在模具形状优化设计中的应用。

6.1　模具优化设计步骤

基于有限元灵敏度分析的模具形状优化设计方法的形状目标函数定义为对应于任意一个预成形模具形状所得到的终成形件形状与理想终成形件形状之间的差别,这种差别采用这两个形状不重合区域的面积来度量。采用三次B样条函数描述预成形模具形状,B样条函数的控制点坐标作为优化设计变量。使用令形状目标函数取最小值的设计变量定义最佳预成形模具形状。目标函数是有限元节点坐标、节点速度和优化设计变量的复合函数,目标函数对优化设计变量的梯度是分别通过节点坐标和速度对设计变量的灵敏度来建立的。灵敏度分析的边界条件为速度灵敏度边界条件。选用一种合适的优化算法,例如拟牛顿方法的BFGS优化算法等确定优化搜索方向,BFGS算法的理论性较强,计算效率与可靠性良好。

下面以两工序锻造过程为例,叙述模具形状优化设计的具体步骤。

对于一个多工序锻造问题,模具优化设计的目的是通过优化各预锻工序的模具形状来获得无缺陷、少无飞边的净形状(Net-Shape)终锻件。这种基于有限元方法的成形过程模拟优化设计的计算机程序框图如图4.4-5所示。

其详细的优化设计步骤叙述如下。

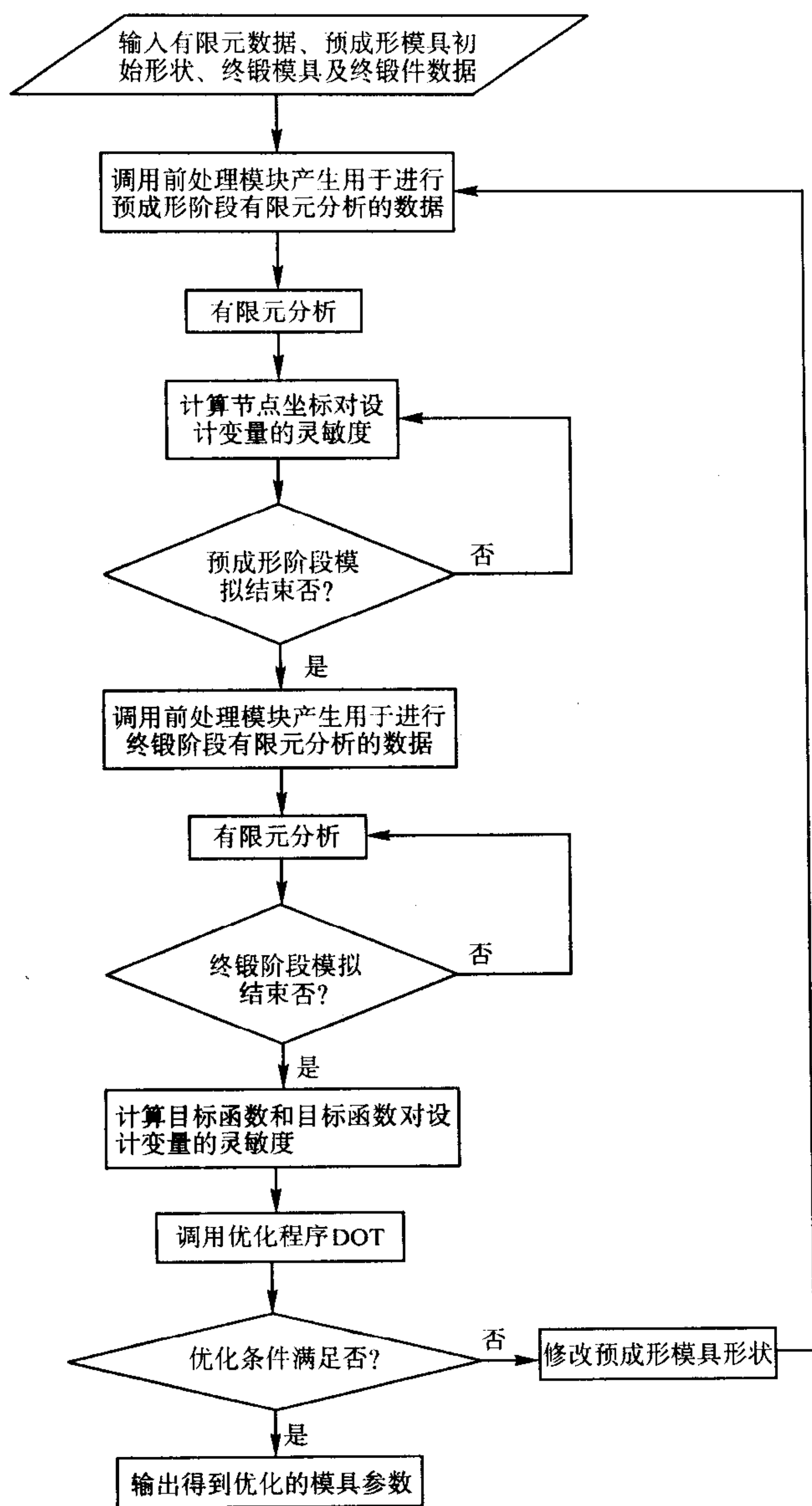


图4.4-5 模具优化设计计算机程序框图

1）假定一个用于优化迭代的初始（猜测）预锻模具形状（即假设B样条函数控制点的初始坐标值）。

2）进行预锻工序的有限元正向模拟，对于每一个时间增量模拟步，首先进行有限元正向模拟的速度场迭代，当速度场收敛后，求解有限元节点速度灵敏度方程，利用求出的节点速度灵敏度刷新节点坐标灵敏度，直至预锻工序的变形过程模拟完毕。

3）在预锻工序模拟结束后，进行终锻工序的有限元正向模拟，由于终锻模具形状预先给定，且固定不变，故在终锻阶段，有限元节点速度灵敏度为零，不必进行终锻阶段的灵敏度分析，只进行有限元仿真。当终锻变形过程模拟结束后，应用所得到的实际终锻件形状、预先给定的理想终锻件形状及求得的累加节点坐标灵敏度，计算目标函数值及其对设计变量的梯度值。

4）调用优化工具进行优化，获得一组新的设计变量值，这组新的设计变量值定义了一个新的预锻模具形状。同时检查优化迭代的收敛性，如果连续两次迭代的目标函数值的绝对变化小于 max $(0.0001|F_0|, 10^{-20})$，式中 F_0 为初次优化迭代的目标函数值，或者连续两次优化迭代的目标函数值的相对变化小于0.001，就认为优化迭代过程收敛。当前的设计变量值就定义最佳预锻模具形状。若不满足优化迭代的收敛性条件，则用新得到的控制点坐标刷新预锻模具形状，重复2）~4）步骤，进行下一次优化迭代，直至满足优化迭代的收敛性条件。

6.2 体积损失的调整

对于刚塑性、刚（黏）塑性材料的有限元方法，通常采用惩罚函数法或拉格朗日乘子法将体积不变条件引入泛函，并进行变分。但在模拟过程中，由于工件几何形状的不断刷新和多次网格再划分等原因，会引起工件体积损失。体积损失量的大小主要取决于工件的变形量、增量步长、网格再划分次数及网格划分的稀密程度等因素。一般情况下，体积损失为工件体积的1%~6%之间。在常规的有限元模拟中，对于形状不甚复杂的工件，体积损失对模拟结果的影响可以忽略不计，但对于较复杂形状的成形问题，特别是形状较复杂的三维成形问题，其体积损失相当可观。例如，对于推土机链轨节第二预锻工步的三维有限元模拟，其体积损失为6%，这就影响了模拟结果的可信度，更困难的是不清楚工件内各部位体积损失的定量关系及分布情况。

如果减少增量步长，则由几何刷新引起的体积损失将减少，但导致计算时间增加。若同时考虑减小时间增量步长以降低几何刷新造成的体积损失和由此导致的计算效率降低，一般取每个时间增量内的相对压下量为0.01。但无论怎样减少时间步长（但不能为零），由几何刷新造成的体积损失总是存在。

网格再划分是影响体积损失的又一主要因素。对于四节点四边形单元（最常用的一种二维单元），其边界为直线。当边界单元与曲线模具边界接触时，对于内凹模具边界，单元的边界总是位于模具轮廓以外，此处单元边界经有限元网格再划分后，一般不会引起单元体积的减少；但对于外凸模具边界，当单元与模具边界接触后，其边界总是位于模具轮廓内部（尽管该单元边界的两个节点必须位于模具边界上），在进行网格再划分时，此处单元一般都被细划，因细划而在单元边界新添加的新网格节点需被调整到该处的模具边界上，因而使得旧单元边界与模具轮廓所包围的区域的材料体积损失掉。这样，网格再划分过程总是使得材料体积受到损失。若对于不同的成形问题采用同一网格畸变判断准则，则形状复杂和变形量较大的问题的网格再划分次数多于形状简单和变形量较小的问题，且体积损失较大。若对于同一成形问题采用不同的网格畸变判别准则，当网格畸变程度达到较严重时才进行网格再划分，即网格再划分次数较少，则体积损失较大；反之，当网格畸变程度达到较小时就进行网格再划分，网格再划分次数较多，则体积损失较少。通过细划初始有限元网格、再划分的网格，会减少材料的体积损失，但同时带来计算时间的显著增加。

若实际用于几何刷新的时间步长大于本步内的最小接触步长，则边界自由节点接触模具是引起体积损失的第三个主要因素。如果从变形材料的流动和模具的运动关系来看，其实质就是变形材料逐步与模具贴合的过程（包括少数节点与模具接触－脱离－再接触的情况）。在一个增量变形步内，规定多少自由边界节点与模具接触对体积损失影响较大，接触节点的数目取决于时间步长的大小。规定接触节点的数目有两种方法：一是规定一个模拟步内最多有一个边界节点接触模具边界，实施这种规定的方法是取实际增量步长为指定的时间步长、最小接触时间步长、该步内最大许可的等效应变增量和最大许可的等效应变率增量的最小值。如果最小接触时间步长为其中的最小值，则该步内只有与该最小接触时间步长对应的节点接触模具边界，否则，该步内无接触节点，以这种接触方法接触模具的节点几乎不影响体积损失，

但这种处理方法使得计算效率降低。二是实际步长大于最小接触时间步长的情况，该步内可能有多于一个的节点接触模具边界，对于这些节点，经几何刷新后节点进入模具内部，应调整这些节点至最近的模具位置，这种调整势必引起工件的体积损失，而且损失量较几何刷新大。

工件体积损失是有限元方法及其过程模拟中出现的，不是实际成形生产中的现象。欲使模拟结果能接近于实际成形情况。在基于有限元灵敏度分析的模具形状优化设计方法中，初始坯料体积等于理想终成形件体积，理想终成形件形状和体积固定不变，目标函数取决于理想终成形件与实际终成形件形状不重合区域的面积。由于体积损失，实际终成形件体积总是小于理想终成形件体积，更重要的是在每次优化迭代过程中体积损失量不同，这就使得较为复杂成形过程的模具形状优化迭代过程收敛较困难。因此，就必须消除或减少模拟时产生的体积损失。

此处介绍一种修正体积损失的方法，该方法可在有限元分析软件中得以实现。其修正原理为：在每个时间增量模拟步结束后，调整与模具相接触的节点在竖直方向（y 或 z 向）的坐标位置，使工件体积等于初始坯料体积，同时，沿竖直方向调整模具同样的距离。一般地，对于一个材料成形过程，需要几百个时间增量步的模拟，在一个增量步内只修正该步的体积损失，因时间增量和在该时间增量内的模具行程都很小，接触节点的调整距离则更小，故这种调整对于材料流动的模拟结果影响很小。另一方面，体积损失主要集中在工件与模具的接触面上，即接触节点处。因此，这种体积损失的调整方法具有可行性。

图4.4-6表示一个与上模接触的单元，其中节点4为接触节点，其坐标为（x_4，y_4），设其在 y 方向上的坐标调整量为 Δy，则节点4的新坐标为（x_4，$y_4+\Delta y$）。

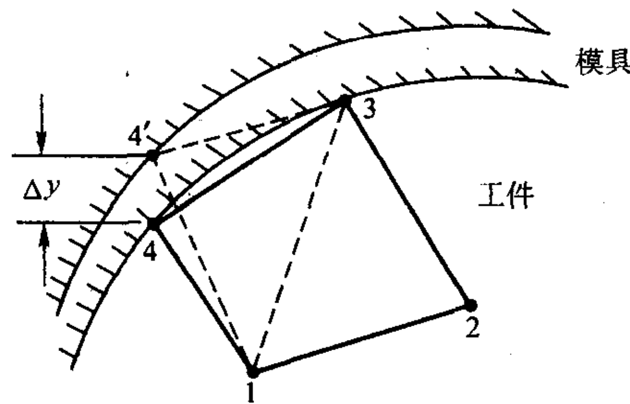


图4.4-6　体积损失的调整

对于轴对称问题，单元$\overline{1234}$的体积为

$$V_e=\frac{1}{2}C_1(x_{32}y_{12}-x_{12}y_{32})+\frac{1}{2}C_2(x_{14}y_{34}-x_{34}y_{14})$$

其中，$C_1=\frac{2}{3}\pi(x_1+x_2+x_3)$，$C_2=\frac{2}{3}\pi(x_1+x_3+x_4)$。

对于平面变形问题，$C_1=C_2=1$。单元$\overline{1234'}$的体积为

$$V_{e'}=\frac{1}{2}C_1(x_{32}y_{12}-x_{12}y_{32})+\frac{1}{2}C_2(x_{14}y_{34'}-x_{34}y_{14'})$$

由于调整节点4引起的体积增加量为

$$\Delta V_4=V_{e'}-V_e=\frac{1}{2}C_2(x_{34}-x_{14})\Delta y=S_4\Delta y$$

其中，$S_4=\frac{1}{2}C_2(x_{34}-x_{14})$。

同理，在每个模拟步内，对所有与模具接触的节点位置进行调整，并对各节点坐标调整引起的体积增加求和，可得到总体积的增加量 ΔV_T

$$\Delta V_T=\Delta y\sum S_i$$

令 ΔV_T 等于该模拟步内的体积损失（V_0-V），其中 V 为该模拟步结束时的工件体积，可得到所有接触节点在竖直方向（y 向）的调整量

$$\Delta y=\frac{\Delta V_T}{\sum S_i}=\frac{V_0-V}{\sum S_i}$$

大量数值计算表明，在一个模拟步内，调整量 Δy 远远小于该模拟步内模具的增量行程，二者的比值仅为0.001～0.01，如此小的调整量不会影响工件的变形流动模式，但在每个模拟步内实现了工件体积不变，从而保证了每次优化迭代后实际终成形件体积等于初始坯料体积，消除了体积损失对优化收敛性的影响。

图4.4-7为一平面变形过程结束状态的有限元模拟结果。在模拟中采用了体积损失补偿方法。初始变形体体积为2 152.4 mm³，步长恒定，模拟总步数为92步，每个时间增量内的压下行程为0.2 mm，假设上模运动速度为－1 mm/s，则时间步长为0.2 s。在整个模拟过程中进行了3次网格再划分。整个变形过程的体积损失为15.7 mm³，占初始体积的0.73%。图4.4-8为每个时间增量内补偿体积损失的上模调整量与加载（模拟）步数之间的关系曲线。由图中曲线可以看出，绝大多数增量模拟步内的调整量为0.001 2mm，与增量行程的比值为0.006。其中有三个增量步内的调节量在0.01～0.015之间，与增量行程的比值在0.05～0.075之间，这三个模拟步为网格再划分模拟步。其他体积损失较大的模拟步为边界接触条件发生变化的模拟步。对于该例，采用的步长固定，因而存在一步内有多于一个节点接触模具的现象，这种边界条件的变化也引起较大的体积损失。图4.4-9为加载步数与体积的变化曲线，当边界接触节点不变时，体积减少大致呈线性，但当接触节点增加时，体积减少较大。



图4.4-7　平面变形问题有限元模拟结果

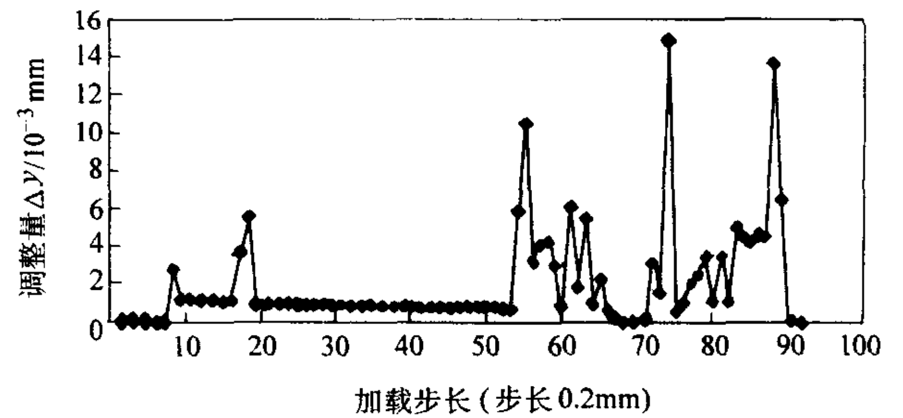


图4.4-8　加载步内的调节量变化曲线

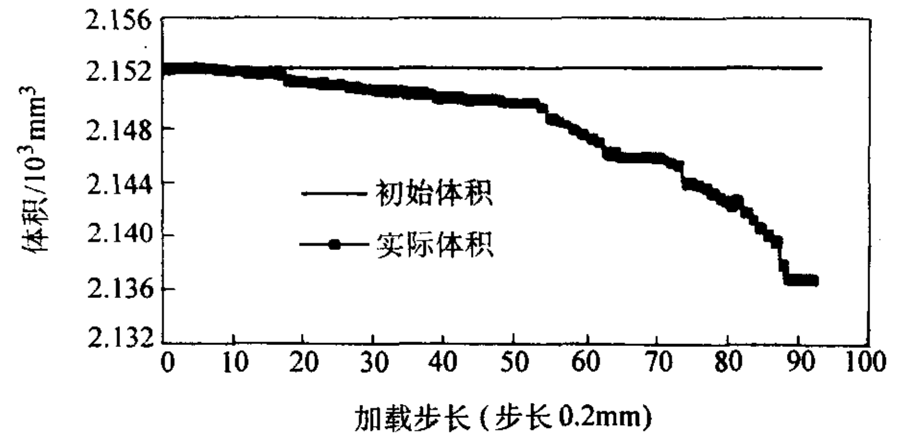


图4.4-9　计算体积随加载步数的关系曲线

图4.4-10为一轴对称锻造过程的刚塑性有限元模拟结果。初始变形体体积为627 028 mm³，步长恒定，增量模拟总步数为88步，每个时间增量内的压下行程、上模运动速度及时间步长与上例相同，在整个模拟过程中无网格再划分。整个变形过程的体积损失为1 845 mm³，占初始体积的0.294%。图4.4-11为每个时间增量内补偿体积损失的上模调整量与加载（模拟）步数之间的关系曲线。由图中曲线可以看出，绝大多数增量模拟步内的调整量为0.001 2 mm，与

增量行程的比值为 0.006。只有少数增量步内的调节量在 0.01 mm 附近，与增量行程的比值为 0.05 左右，这些少数模拟步为网格再划分步或有节点接触模具的模拟步。图 4.4-12 为加载步数与体积的变化曲线，当边界接触节点不变时，体积减少大致呈线性，但当接触节点增加时，体积减少较大。从图中也可以看出，在整个变形过程中，存在两个接触节点增加的阶段，即边界条件变化阶段，一是在 60～70 模拟步内，二是在 80～85 步内。

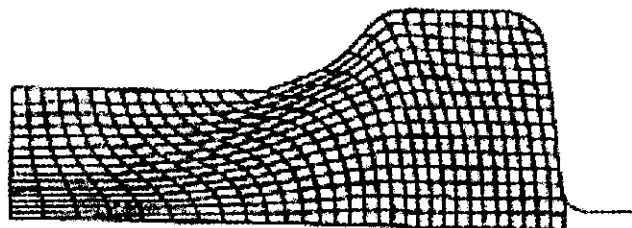


图 4.4-10　轴对称问题的有限元模拟结果

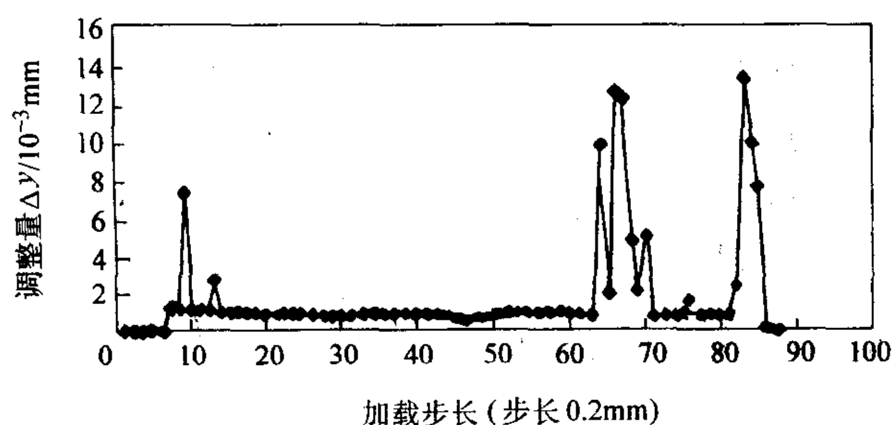


图 4.4-11　加载步内的调节量变化曲线

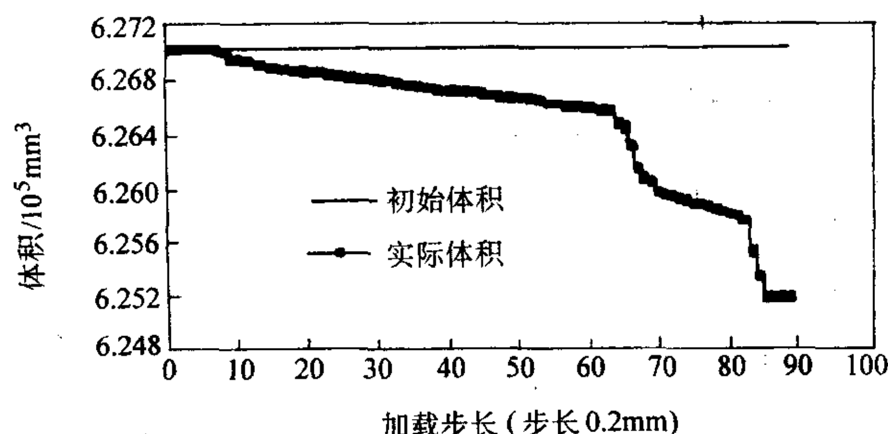


图 4.4-12　计算体积随加载步数的关系曲线

上述两例均为简单成形问题，体积损失较小，可以不计体积损失。但对于较复杂成形问题或变形量较大问题，体积损失较大，就不能不考虑体积损失问题了。

6.3　有限元网格再划分

材料塑性成形过程的变形量及模具表面与工件之间的相对滑动位移均很大，这种大变形与位移使有限元网格出现严重畸变，当外凸的四边形单元变成三角形单元或内凹单元时，等参变换的雅可比行列式的值出现负值或零，致使等参变换失真。为此，当单元畸变时，必须进行网格再划分。对于有限元模拟方法，在进行网格再划分时，应在旧网格区域上生成一套新的有限元网格系统，并将与变形历史有关的信息，如单元等效应变、节点温度等从旧网格上插值转换到新网格系统上去，从而得到新网格系统的用于有限元模拟过程继续进行的数据文件。对于基于有限元灵敏度分析的模具形状优化设计方法，除了常规有限元模拟时网格再划分所进行的数据传递外，还应当包括节点坐标灵敏度的插值转换。因为它也是与变形历史有关的量。节点坐标灵敏度的插值转换步骤为：首先采用单元形状函数插值确定出单元内的坐标灵敏度分布；其次通过新节点与旧单元的包含测试确定出新节点所落入的旧单元；最后利用旧单元上的坐标灵敏度分布，插值求出新网格节点上的坐标灵敏度值。对于节点速度对设计变量的灵敏度，由于其是瞬时量，不必进行插值转换。关于详细的插值转换方法，请参考相关文献。

6.4　提高程序运算速度的方法

有限元灵敏度分析方法在预锻模拟阶段的每一步后，需要计算所有节点坐标对每个优化设计变量的灵敏度，程序对每个优化设计变量求解灵敏度单元刚度矩阵，组装整体刚度方程，然后求解方程。有多少设计变量，就需要求解多少次灵敏度刚度方程。程序大部分时间消耗在求解灵敏度刚度方程上，当节点数目和设计变量数目很大时，程序运行时间较长，因此必须寻求提高程序运行速度的方法。从上述几节介绍的灵敏度方程可以看出，计算单元的灵敏度矩阵较复杂，需要较长的计算时间。然而，从式（4.4-15a）单元刚度灵敏度方程的形式看出，每个单元的灵敏度刚度矩阵和右边的 $\partial \boldsymbol{K}/\partial \boldsymbol{X}$、$\partial \boldsymbol{F}/\partial \boldsymbol{X}$ 项在对每个设计变量求解速度灵敏度过程中保持不变，如果在求解第一个优化设计变量的速度灵敏度之后，将这些值保存起来，用于其余的设计变量，将大大提高程序的运行速度。一般情况下，速度可提高 5～10 倍（取决于设计变量的数目）。此外，在大量实例分析和优化中发现，B 样条曲线的 x 坐标在各次优化设计中变化甚微，因此，可以固定 B 样条曲线上的 x 坐标，只对 y 坐标进行优化，这样就可以减少一半的设计变量，程序运行速度又可提高一倍左右。

7　少无鼓形圆柱体镦粗过程的优化设计

圆柱体镦粗过程是金属塑性成形中常用的制坯工序之一。当模具与工件接触面上存在摩擦时，在一个工序上，圆柱体自由镦粗变形过程不可能产生一个侧面无鼓形的终锻件，而且变形不均匀。应用基于灵敏度分析的模具优化设计方法对圆柱体镦粗过程进行优化设计，其目标是寻求最佳预锻模具形状，当用其锻造后的预锻件在一副平模之间进行终锻后能够获得侧面鼓形最小、满足高度尺寸要求的圆盘终锻件。

因圆柱体镦粗过程为轴对称问题，故只取 1/4 模型进行有限元分析和优化设计。此处介绍高度压下量 50% 的圆柱体镦粗变形过程的模具形状优化设计。假设初始圆柱体坯料尺寸为 ϕ100 mm×100 mm，变形材料为刚（黏）塑性体，变形为等温变形过程，材料的应力和应变率关系为 $\bar{\sigma} = Y_0\,\dot{\bar{\varepsilon}}^{0.145}$，其中 $Y_0 = 33.99$ MN/mm²。假设坯料与模具接触面上的摩擦为常摩擦状态，其摩擦因数为 0.3，下模静止不动，上模的运动速度为 -1.0 mm/s。一般来讲，表示预锻模具形状的 B 样条曲线控制点的数目多少取决于终锻（件）模具形状的复杂程度，如果终锻模具形状复杂，应取较多的控制点数目，但是较多的控制点数目会带来较多的运算时间。对于圆柱体镦粗过程，选取 8 个 B 样条控制点（p_{x_i}，p_{y_i}）（$p_{x_1} > p_{x_2} > \cdots > p_{x_n}$）描述预锻模具形状，这样就有 5 条 B 样条曲线和 16 个优化设计变量 p_l（$l = 1, 2, \cdots, 16$）。为了保证预锻模具形状的起始点的横坐标为零，根据式（4.4-27a），令 $p_{x_1} = -p_{x_3} = b$，$p_{x_2} = 0$，其中 b 为一正常数。因此在优化过程中，对 p_{x_1}，p_{x_2} 和 p_{x_3} 施加上下限，p_{x_1}，p_{x_2} 和 p_{x_3} 上下限分别为 $-b$、0 和 b。其他设计变量无上、下限限制。

假设镦粗压下量为 50%。图 4.4-13 表示目标函数与优化迭代次数之间的关系曲线，当经过 9 次优化迭代后，目标函数值趋于稳定并几乎接近于零。每次迭代过程中的预锻件形状、预锻模具形状及终锻后的终锻件形状如图 4.4-14 所示。经过 9 次优化迭代后所得到的预锻模具和预锻件，经终锻后可得到一个侧面几乎无鼓形的圆盘终锻件，其最小

半径为69.52 mm，最大半径为70.44 mm，侧面鼓形的单边宽度仅为0.92 mm。理想终锻件半径为70.71 mm，本例未考虑体积损失的调整，其有限元分析过程中的体积损失为0.77%。因而使得优化得到的终锻件最大半径比理想终锻件半径小0.27 mm。可见，采用优化方法所得到的最佳预锻模具形状，可大大降低终锻件鼓形现象，其材料利用率提高了8.81%。图4.4-15为三种终锻件形状的比较，显然采用优化方法得出的预锻模具锻造出的终锻件形状几乎与理想终锻件形状相同。图4.4-16给出了预锻模具形状的优化迭代历史。经过9次优化迭代后，预锻模具形状几乎保持不变。

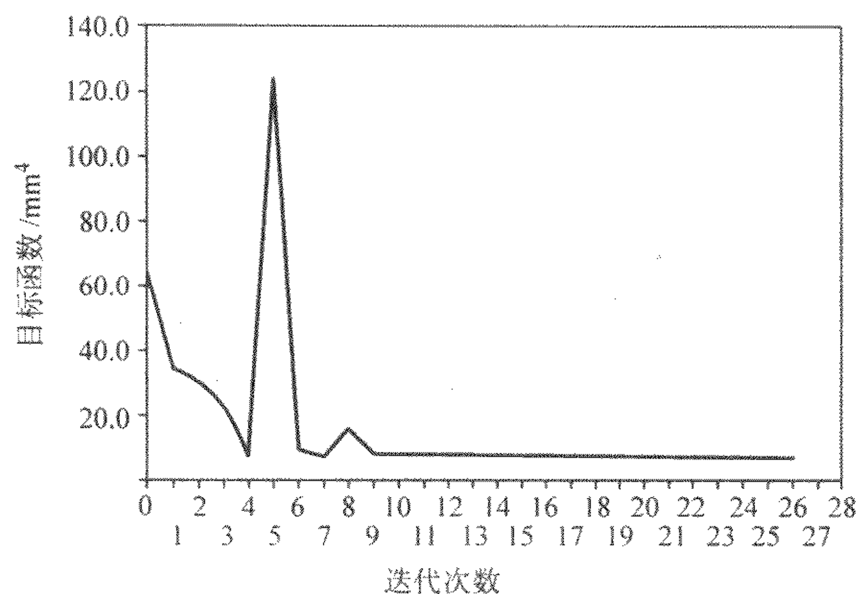


图4.4-13　目标函数值随优化迭代次数的变化曲线

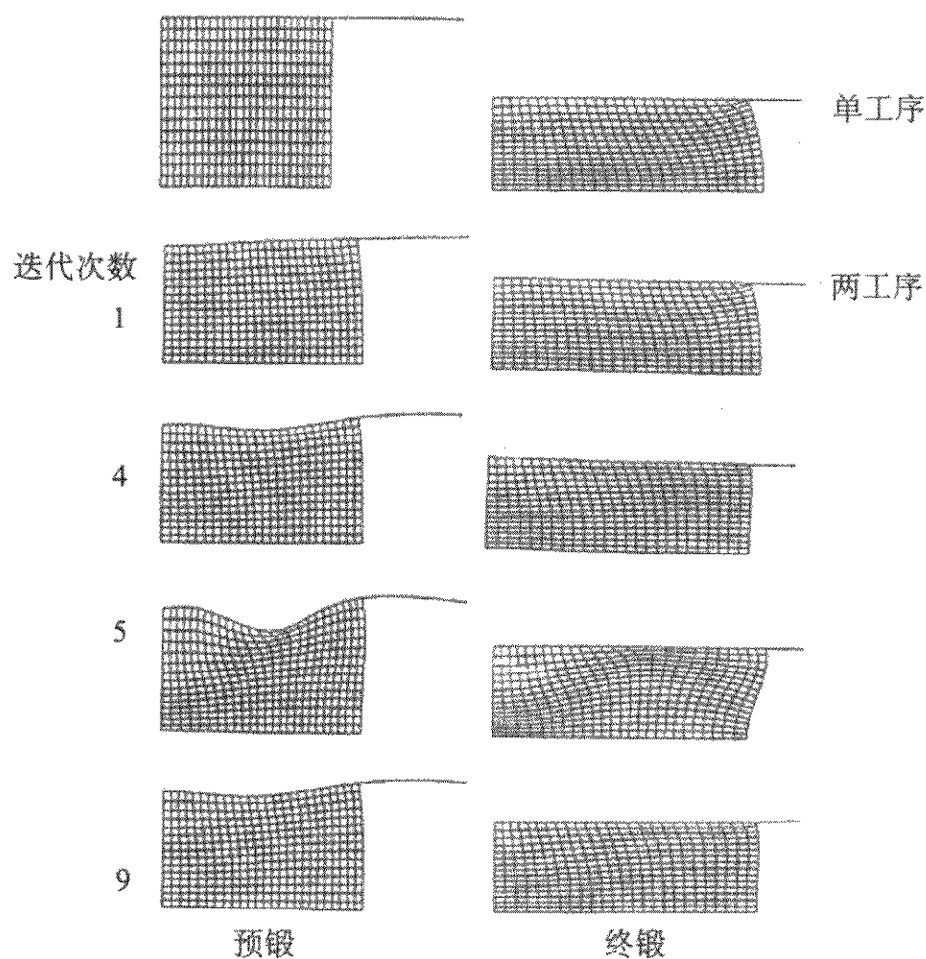


图4.4-14　优化迭代过程中预锻件形状、预锻模具形状与终锻件形状变化

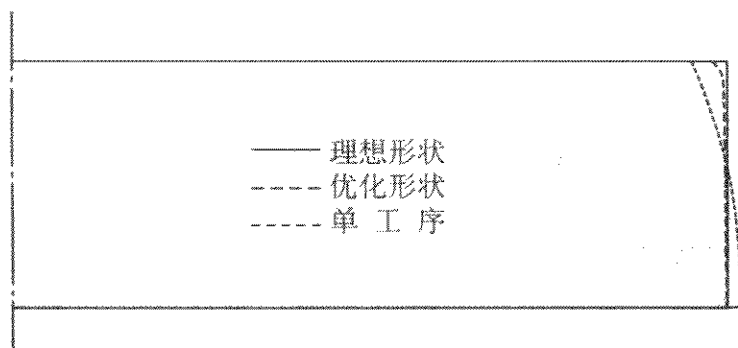


图4.4-15　终锻件形状的比较

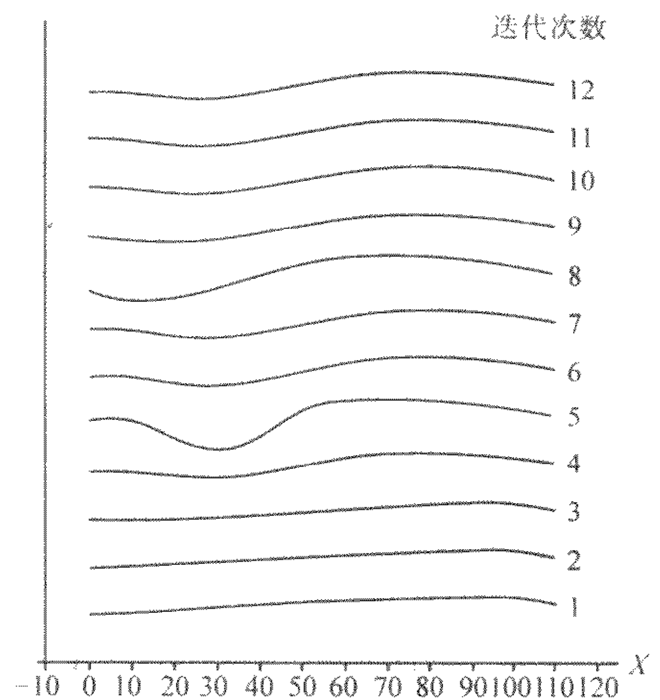


图4.4-16　优化迭代过程中预锻模具形状的变化历史

8　模具有型腔的锻造过程优化设计

上节介绍了简单镦粗过程的预锻模具优化设计过程，本节将介绍模具带有型腔的锻造过程预锻模具形状的优化设计。优化设计的目的是通过优化设计预锻模具形状，使得到的预锻件形状经过终锻过程后得到的终锻件完全充满型腔且少无飞边。

8.1　H形截面轴对称锻件锻造过程（型腔高径比$h/b=1$）

H形截面锻件锻造过程是基本的锻造过程之一。其设计目的是通过优化设计预锻模形状，使其锻造出的预锻件，经终锻后能够获得无飞边、完全充满模腔的终锻件，即实现净形状锻造。本节介绍H形截面轴对称锻件锻造过程的优化设计过程。

假设锻造过程为等温成形过程，材料为刚（黏）塑性体，模具与工件之间的摩擦状态为常摩擦状态，摩擦因数$m=0.2$，坯料为圆柱体，高度与直径分别为100 mm，变形材料为刚（黏）塑性体，变形为等温变形过程，材料的应力和应变率关系为$\bar{\sigma}=Y_0\,\bar{\dot{\varepsilon}}^{0.145}$，其中$Y_0=33.99$ MN/mm²。下模静止不动，上模的运动速度为−1.0 mm/s。由于对称，只取1/4模型进行分析与优化。采用10个B样条控制点表示预锻模具形状，故有20个优化设计变量p_l（$l=1,2,\cdots,20$）。用于优化迭代的预锻模具的初始（猜测）形状取为平模。

在优化迭代过程中，保持预锻件在竖直轴方向上的高度不变，这个高度决定了预锻过程的压下行程。图4.4-17表示得到的不同优化迭代次数对应的预锻模具形状、预锻件形状和终锻件形状，当预锻模具为平模时，终锻件存在飞边，并且在深腔左上角存在未充满区，因此，该终锻件不符合尺寸要求。随着优化迭代的进行，未充满区域的面积逐渐减少，当经过5次优化迭代后，预锻模具形状的优化迭代过程收敛，可得到无飞边且完全充满型腔的终锻件。因此，对应于第5次优化迭代的预锻模具形状即为最优预锻模具形状。图4.4-18所示的目标函数与优化迭代次数的关系曲线也形象地表明了优化迭代的收敛过程，经过5次优化迭代后，目标函数几乎减少至零，即从初始的9.309 mm⁴减少至0.210 mm⁴，此时优化迭代过程收敛。对于该例，存在网格再划分过程，并采用了体积损失调整方法。图4.4-19表示三种终锻件形状的比较，可见优化求得的终锻件形状几乎与理想终锻件形状完全相同。图4.4-20表示预锻模具形状随优化迭代次数的变化关系。

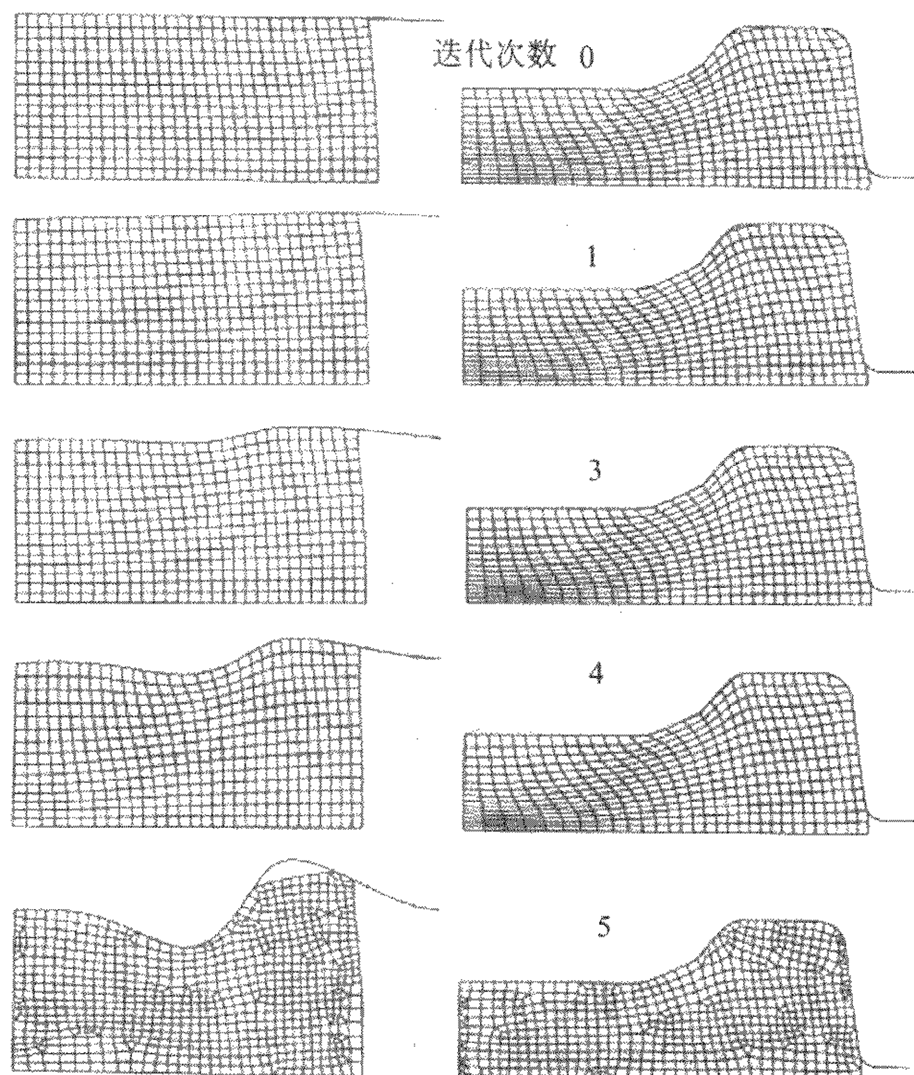


图 4.4-17 不同优化迭代次数对应的预锻件、预锻模具和终锻件形状

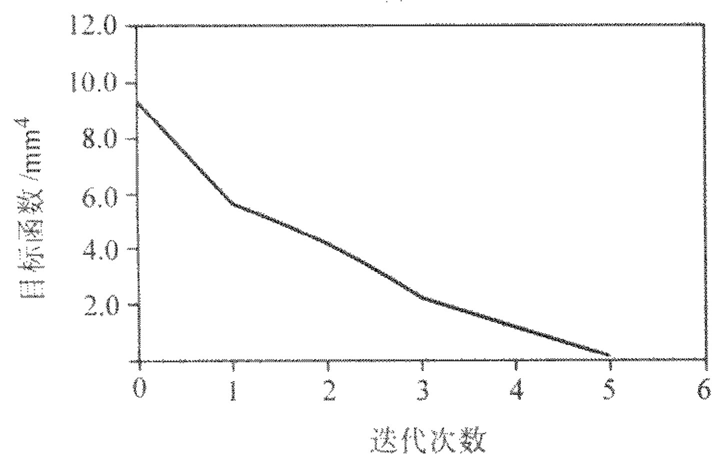


图 4.4-18 目标函数与优化迭代次数的关系曲线

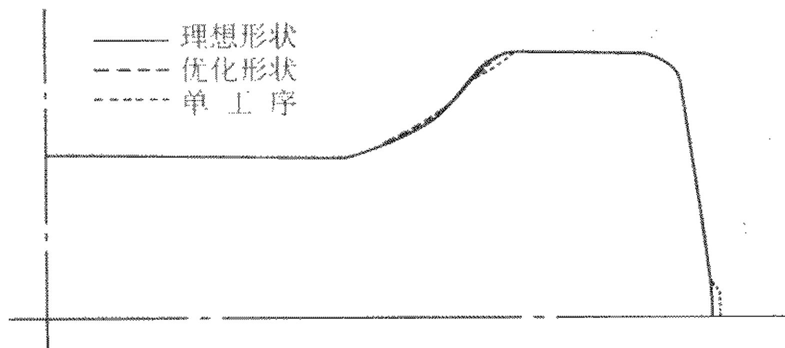


图 4.4-19 三种终锻件形状的比较

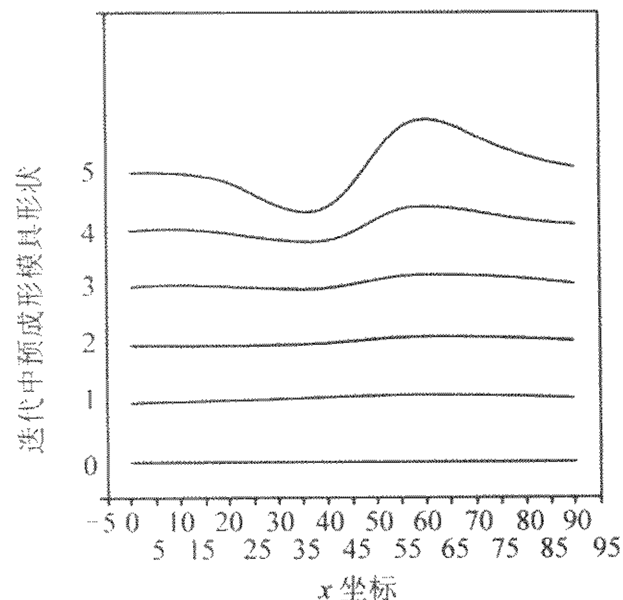


图 4.4-20 预锻模具形状随优化迭代次数的变化历史

8.2 H形截面锻件锻造过程（型腔高径比 $h/b=2$）

型腔高径比 $h/b=2$ 的H形截面锻件锻造过程是典型的两工序锻造过程，也是最基本的锻造过程之一。在实际锻造生产中，往往出现终锻件具有较大飞边的现象，而且在模具的深腔处往往存在未充满区。因此，设计目的是通过优化设计预锻模形状，使其锻造出的预锻件，经终锻后能够获得无飞边、完全充满模腔的终锻件，即实现净形状锻造。将以轴对称变形过程为例介绍其优化设计过程和结果。

初始坯料为一圆柱体，其直径和高度均为100 mm，通过优化设计获得了最佳预锻模具形状，采用该预锻模具所得到的预锻件形状经过终锻后可得到少无飞边且完全充满型腔的终锻件。图4.4-21为对应于每次优化迭代在预锻和终锻阶段结束时获得的预锻模具形状、预锻件形状和终锻件形状。对于初始一次的优化迭代，所获得的终锻件形状存在较大飞边和深腔未充满区。在其后的三次迭代中，模具型腔未充满区逐渐减小，对应的飞边尺寸逐渐减小，这表明优化设计的搜索方向正确。因此，在第5次优化迭代时，优化步长增大，预锻模具型腔具有较大的斜度，即呈陡峭趋势。这种预锻模具型腔的形状变化趋势可增大终锻过程中材料充满型腔和减少飞边尺寸的趋势。但是所获得的预锻件在进行终锻时出现了折叠现象，这意味着预锻模具形状出现过调整现象。出现折叠缺陷后，应当返回上次迭代，减少优化搜索的步长。经过7次优化迭代后，终锻件几乎完全充满型腔，飞边尺寸几乎为零。经过9次优化迭代后，迭代过程收敛，对应的预锻模具形状为优化模具形状，获得了完全充满终锻型腔和几乎无飞边的设计效果。目标函数由初始的73.41 mm^{4}减少到1.24 mm^4。在每次优化迭代过程中，锻造过程均进行了网格再划分。图4.4-22为预锻模具形状随优化迭代次数的变化情况，可见为了充满型腔深处较小的未充满区，预锻模具形状经历了较大的变化。

应指出，上述锻造过程优化设计实例中的理想锻件形状总是要求无任何飞边，这是一个非常严格的要求，对于较复

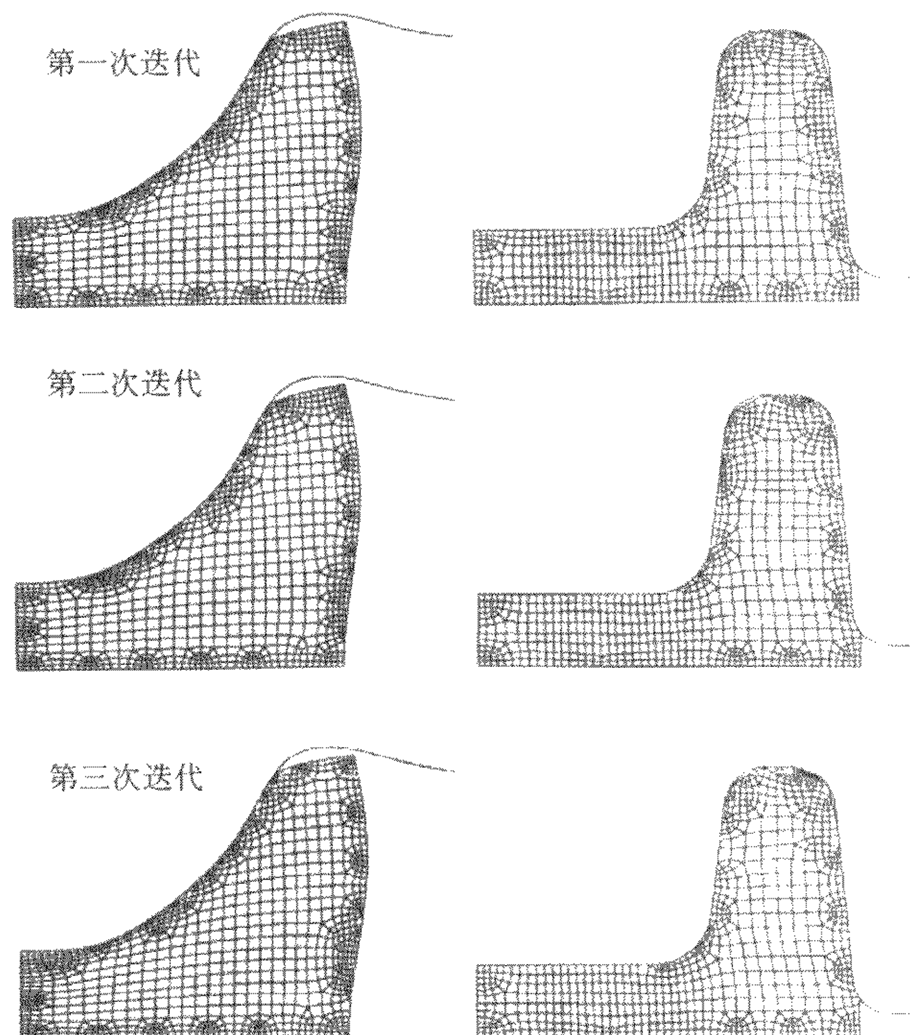
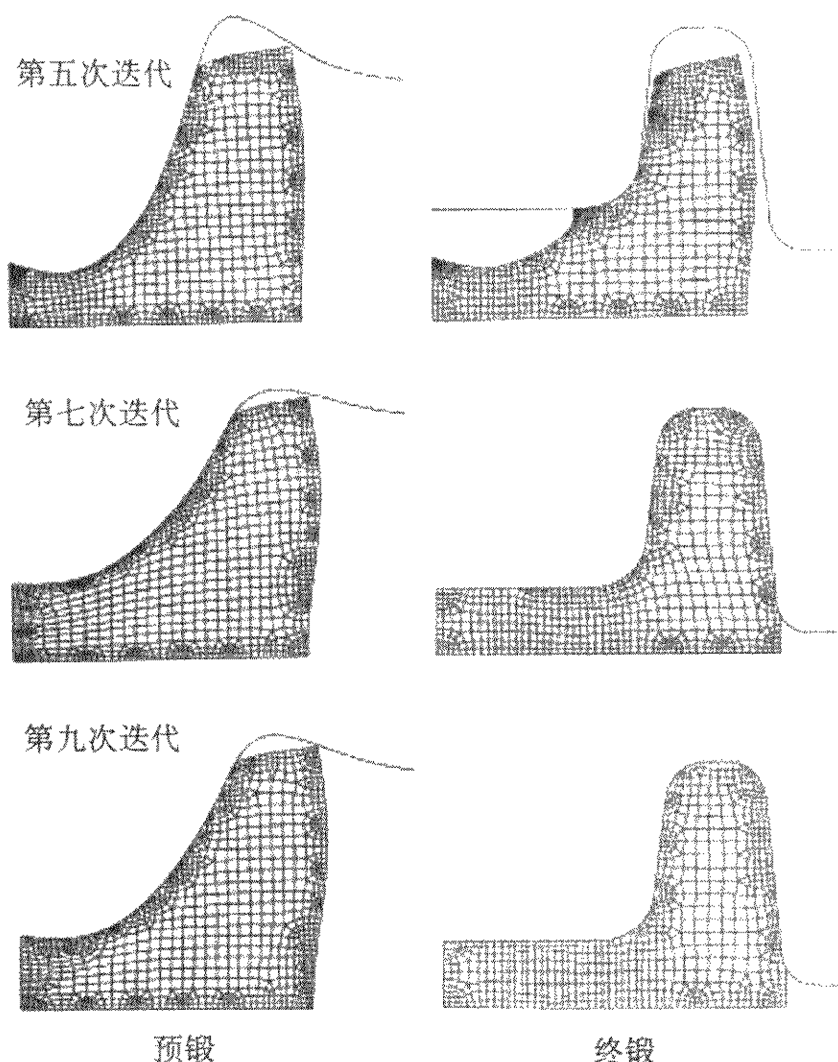


图 4.4-21 不同优化迭代次数对应的预锻件、预锻模具和终锻件形状



续图 4.4-21

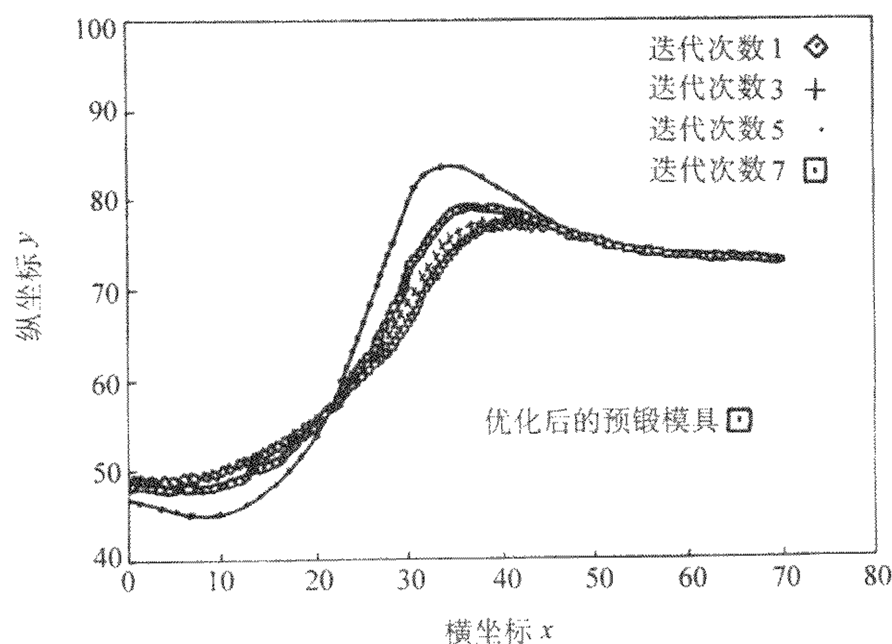


图 4.4-22　预锻模具形状随优化迭代次数的变化关系

杂的锻造问题，设定理想锻件具有一个很小或适量的飞边可有助于优化设计过程的较快收敛。实际上，这在实际锻造生产设计中也是很有必要的，因为坯料的下料一般采用剪切工艺，坯料的长度或体积具有一定误差，因此，为确保型腔完全充满，在理想锻件上适当设置较小的飞边将有助于优化过程的快速收敛，符合实际锻压工艺的需要。

8.3　齿轮坯锻造过程的优化设计

图 4.4-23 所示的锻件为一典型的肋-毂齿轮坯。该锻件由于型腔比较复杂，一次成形时很难充满肋部模腔。本例的设计目的是通过优化预锻模具形状，使金属在终锻时能够充满模腔，得到少无飞边的终锻件。

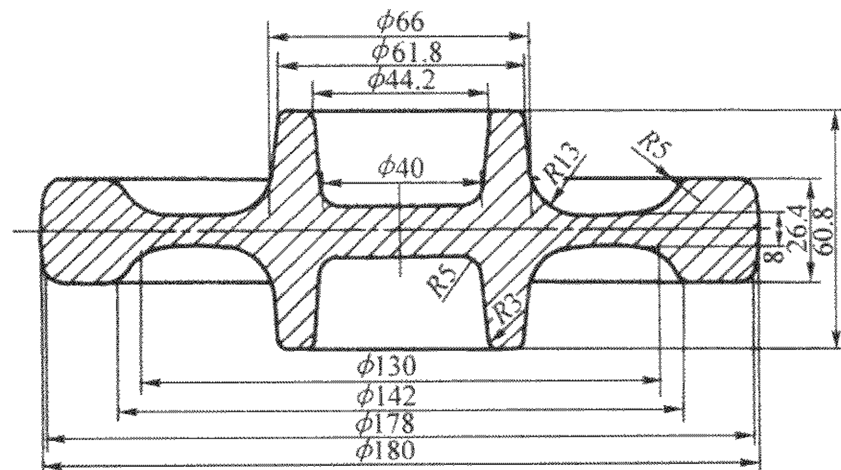


图 4.4-23　齿轮坯锻件

由于本锻件为轴对称变形问题，且上下形状对称，故只取 1/4 模型进行分析和优化，采用 13 个 B 样条控制点表示预锻模具形状。预锻和终锻行程依靠模具到中心线的垂直距离来控制。当预锻模具第一个节点到中心的垂直距离达到 8.4 mm 时，预锻阶段结束，当终锻模具第一个节点到中心线的垂直距离达到 6.4 mm 时终锻阶段结束。选择坯料的体积略大于锻件的体积，坯料形状为 ϕ116.8 mm × 96 mm。假设锻造过程为等温成形过程，材料为刚（黏）塑性体，材料应力应变关系为 $33.99\,\bar{\dot{\varepsilon}}^{n}$ MN/mm^2，其中，n 为 0.145。模具与工件之间的摩擦状态为常摩擦状态，摩擦系数 $m=0.2$。上述工艺参数在各次优化迭代过程中假定不变，图 4.4-24 表示迭代过程中求得的预锻件模具形状、预锻件形状及其对应的终锻件形状。可以看出，用初始的预锻模具所得到的终锻件出现明显的飞边和肋部充不满的现象，经过 9 次优化迭代后，预锻模具形状的优化迭代过程收敛，可得到完全充满模腔，且飞边非常小的终锻件。因此，对应于第 9 次优化迭代的预锻模具形状即为最优预锻模具形状。图 4.4-25 表示第 1、4、9 次优化迭代时的预锻模具形状比较。图 4.4-26 为目标函数值随迭代次数的变化关系，可见，经过 9 次优化迭代后目标函数几乎减少至零。

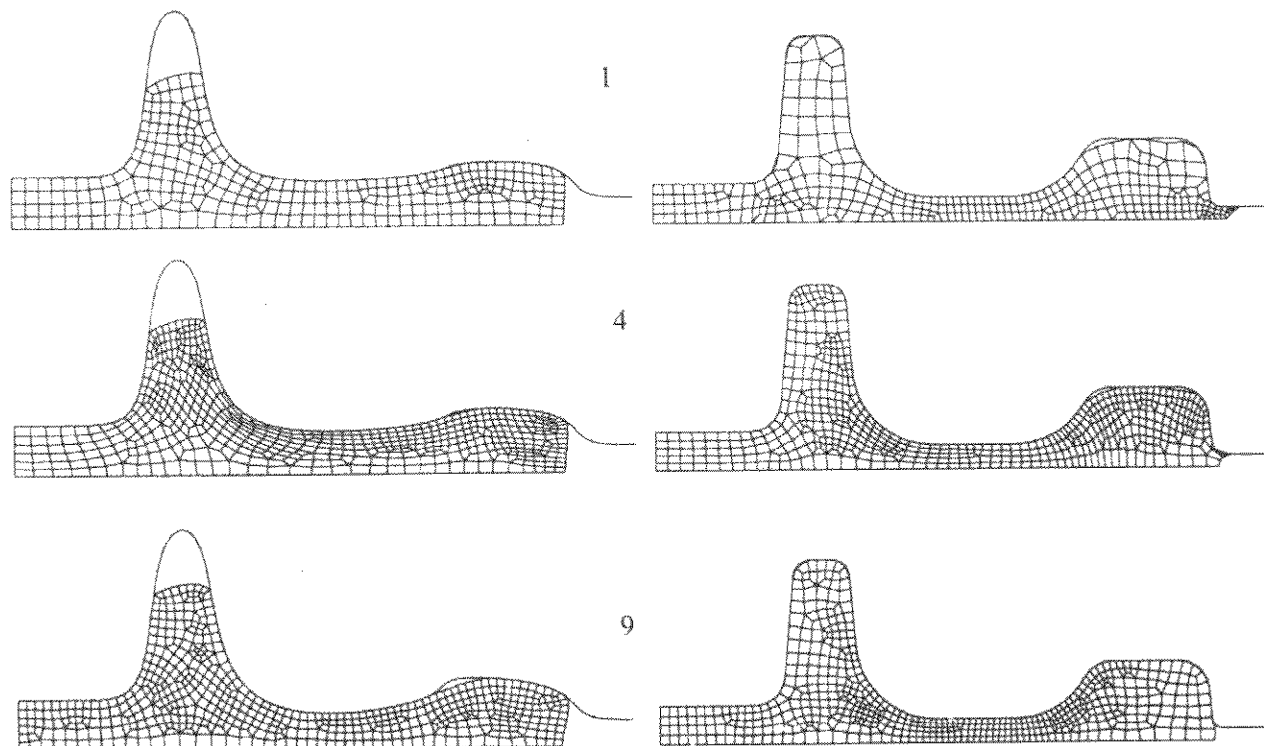


图 4.4-24　优化迭代过程中预锻模具、预锻件和终锻件形状的变化

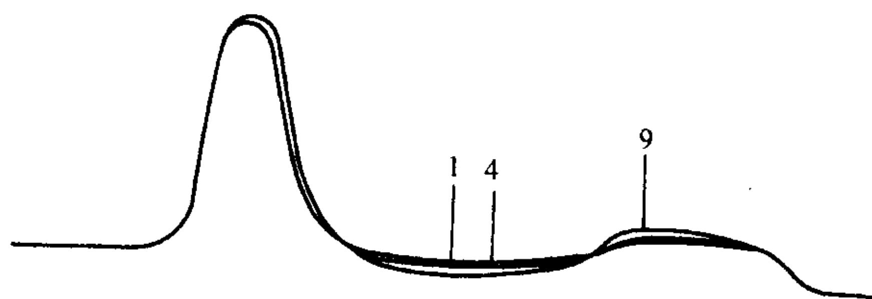


图 4.4-25　第 1、4、9 次优化迭代时的预锻模具形状比较

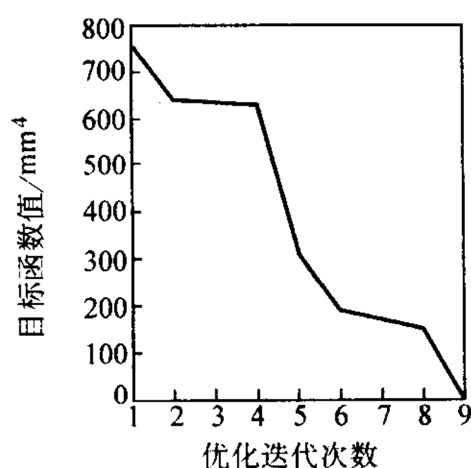


图 4.4-26　目标函数值与迭代次数的关系

从上述实例分析与优化结果来看，基于有限元灵敏度分析的模具形状优化设计方法在塑性成形过程优化中可获得使终锻件少无飞边和完全充满模具型腔的预锻模具形状。对于一般的塑性成形问题，只需要 5～9 次优化迭代即可得到优化解，每次优化迭代只需进行一次变形全过程的有限元模拟，因此，计算效率较高。该方法对于净形状成形过程的模具形状优化设计较有效，为预成形设计提供了一种良好的设计方法。

9　锻件毛坯形状优化设计方法

锻件根据其所需成形工序数量的不同，可分为单工序锻件、两工序锻件及多工序锻件。对于单工序锻件，其形状一般都相对简单，成形仅由一个工序就可以完成。但同时，由于其成形仅由一个工序来完成，终锻模具的形状成为一个定值，所以其毛坯形状就成为影响最终锻件质量的一个重要因素。如何选择合适的毛坯形状就成为单工序锻件形状优化的一个重要内容。

毛坯形状的优化可以分为两种情况。一种是纯意义上的毛坯形状优化。这种情况下，按照体积相等原则，毛坯的体积应与锻件的体积相等（不考虑加工余量和热损耗），这就意味着对于某一个具体问题，毛坯体积是一个常量，另一方面，由于二维锻件的毛坯形状多为圆柱体毛坯或矩形截面毛坯，而圆柱体毛坯或矩形截面毛坯的形状在体积一定的情况下可由其高径比或高宽比唯一确定，所以就转化为以毛坯高径比或高宽比为单一优化变量的一维搜索问题。另一种是实际意义上的毛坯形状优化。这种情况是由现实的生产实际决定的。由于实际生产尤其是对于轴对称锻件来说，其所用的坯料多是圆柱体棒料，而圆柱体棒料的尺寸是由国家标准确定的，与理想锻件所要求的毛坯尺寸会有一定的差距，有时不能满足锻造工艺的要求。常用的解决办法是将标准尺寸的棒料进行镦粗以达到工艺要求的尺寸。而对于镦粗工艺，其压下量的大小是决定其锻件形状及尺寸的主要影响因素，因而，这种情况可转化为以镦粗压下量为单一优化设计变量的一维搜索问题。

由上述分析可知，无论是哪种情况，毛坯形状的优化都可以转化为单一优化设计变量的一维搜索问题。

一维搜索问题是非线性最优化的最基本问题之一，其常用的方法有直接法、曲线拟合法和解析法。其中直接法中的黄金分割法不要求函数的连续性，只要能计算出函数值就可以进行优化求解，简单且优化效果较好，是一维搜索中较常用的一种方法。黄金分割法是一种试探合围法，其基本思路是：首先确定一个包含最优解的计算区间，并在区间内选定两个试算点，计算这两个试算点的函数值，经比较就可以把包含区间缩小一次，然后在缩小的区间内再找一个新的试算点，便可以将包含的区间再度缩小。如此反复进行，直到最后的区间充分小，就得到最优解的近似值。

9.1　目标函数

对于二维成形问题，假设 G 表示对应于任意一个实际坯料尺寸取得的最终成形件的形状，G' 表示模具的形状，L 表示锻件形状的理想边界直线。由于实际锻件未充满整个模具型腔，所以在锻件理想形状边界直线 L 外形成超出部分即为飞边 A。优化设计目标就是通过设计毛坯的高径比（高宽比）或镦粗的压下量使实际工件形状超出理想边界直线部分的体积（如图 4.4-27 中阴影部分 A 所示）最小，从而由体积不变原理可知模具型腔的充填情况最好。所以，以超出

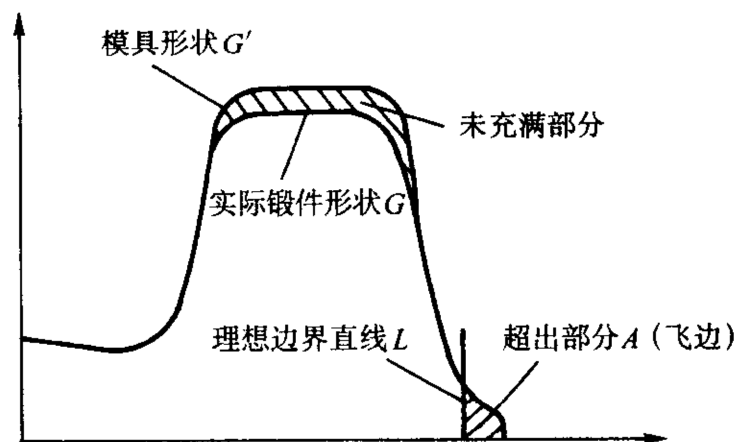


图 4.4-27　目标函数简图

部分的体积作为优化毛坯镦粗压下量问题的目标函数。即以实际锻件超出预定理想直线部分的面积，对于二维问题，目标函数有以下两种情况。

（1）平面应变问题

在平面应变条件下，实际仅考虑其截面。目标函数即为图 4.4-27 中超出部分 A 的面积。欲计算超出部分 A 的面积，首先需要判断预定理想直线与各个单元之间的关系，然后才能准确的算出超出部分的面积。按照理想直线和单元之间的位置，其关系可分为以下三类。

1）整个单元位于直线之内　如图 4.4-28 所示。虽然这种情况下单元和理想直线仍可能会有一个或两个交点，但交点都位于单元边上，因而不会对超出部分造成影响，所以可以直接略过这些单元而不必计算其面积。

2）整个单元位于直线之外　如图 4.4-29 所示。同第一种情况类似，单元和直线之间仍会有一个或两个交点，但交点同样都是位于单元的边上，所以在计算超出部分的面积时可以直接计算整个单元的面积。

设任意单元的节点坐标为 $(x_i,\ y_i)$，$i=1,\ 2,\ 3,\ 4$，则其面积可计算如下：

$$S=\frac{1}{2}(x_{32}y_{12}-x_{12}y_{32})+\frac{1}{2}(x_{14}y_{34}-x_{34}y_{14}) \qquad (4.4\text{-}44)$$

式中，$x_{ij}=x_i-x_j$，$y_{ij}=y_i-y_j$。

3）单元和直线相交　如图 4.4-30 所示。此时，按照单元位于直线之外部分（图中阴影部分）的形状，又可以细分成四边形、三角形和五边形三种情况。无论是哪种情况，在计算面积之前，都需要求出单元与直线交点的坐标，然后再进行计算。

① 四边形　如图 4.4-30（a）所示，首先按逆时针顺序把两个交点和直线之外的两个单元节点组合成一个新的四边

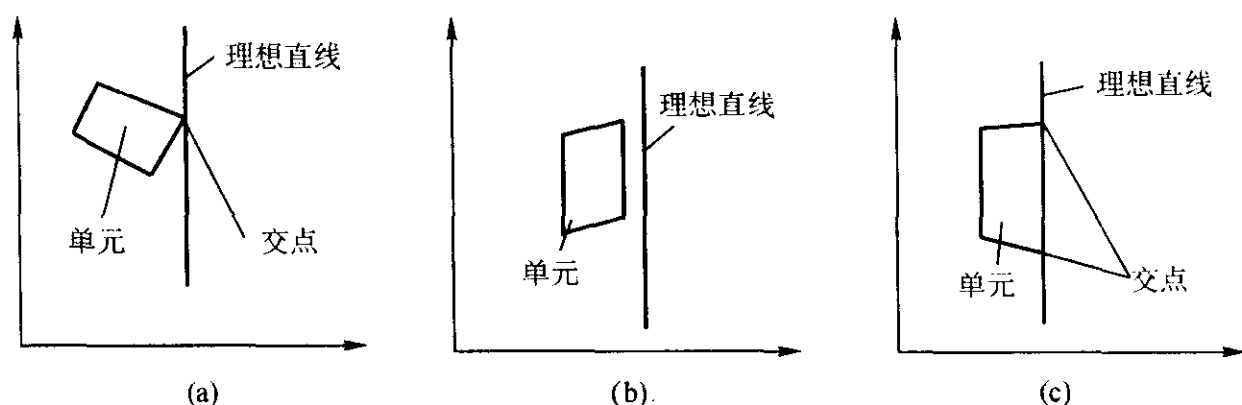


图 4.4-28 单元位于理想直线之内的情况

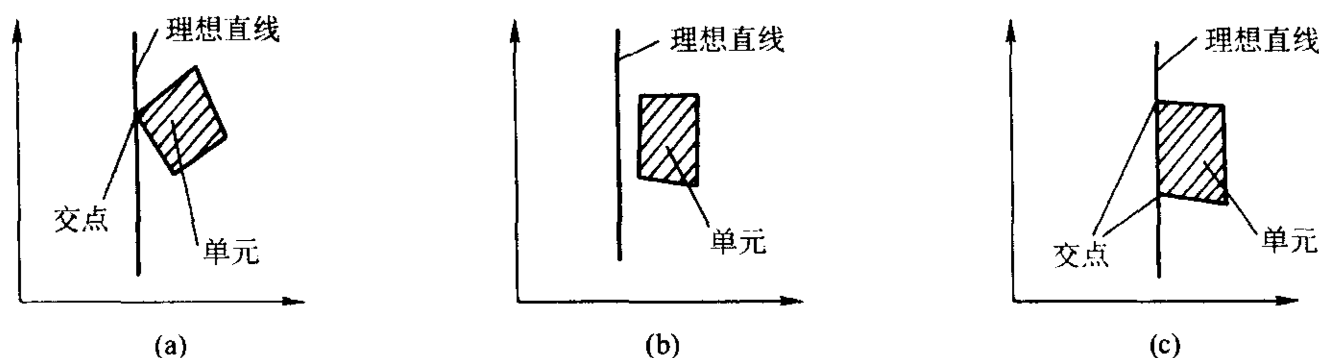


图 4.4-29 单元位于理想直线之外的情况

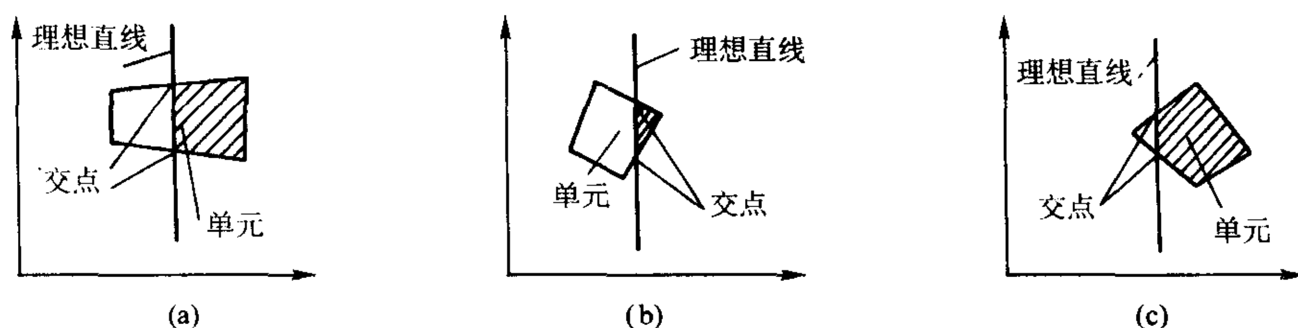


图 4.4-30 单元和直线相交的情况

形单元，再将新单元的节点坐标代入式（4.4-44）就可以求出新单元的面积。

② 三角形　如图 4.4-30（b）所示。此时两个交点与直线之外的一个单元节点形成一个三角形。将三角形的三个节点按逆时针编号，并设编号后的节点坐标为（x_i，y_i），$i=1$，2，3，则其面积计算如下：

$$S=(x_{32}y_{12}-x_{12}y_{32})/2 \tag{4.4-45}$$

③ 五边形　如图 4.4-30（c）所示。这种情况下，单元与直线的两个交点和单元位于直线之外的三个节点形成了一个五边形，同时和单元的另一个节点形成了一个三角形。此时，我们可以先按照上面的方法，计算出三角形的面积，再计算出整个四边形单元的面积，用单元的面积减去三角形的面积就可以得到五边形的面积。

综合考虑上述情况，则目标函数也就是超出部分的面积可表示如下：

$$\psi=\sum_{n=1}^{M}A_n \tag{4.4-46}$$

式中，A_n 为第 n 个需计算面积的单元的面积；M 为需计算面积的单元总数。

利用式（4.4-46）就可以计算出平面应变问题目标函数 ψ。当 ψ 越小，实际获得的最终成形件超出预定理想直线部分的面积就越小，终锻模具的充填状况也就越好。

（2）轴对称问题

对于轴对称问题，由于零件的形状是和体积相对应的，即使是相同的截面面积，离对称轴距离的不同其体积也不相同。因而再用截面面积表示形状的差异是不恰当的，所以本文采用截面的体积来表示形状的差异，也就是以超出部分的体积作为优化的目标函数。

轴对称情况下单元和直线之间的关系与平面变形情况是相同的，计算方法也类似，不同的是计算单元超出的部分是体积而不是面积。其目标函数可表示如下：

$$\psi=\sum_{n=1}^{M}V_n \tag{4.4-47}$$

式中，V_n 为第 n 个需计算体积单元的体积；M 为需计算体积的单元总数。

对于任意四边形轴对称单元 n，设单元节点的坐标为（x_i，y_i），$i=1$，2，3，4，其体积可计算如下：

$$V_n=\frac{1}{2}C_1(x_{32}y_{12}-x_{12}y_{32})+\frac{1}{2}C_2(x_{14}y_{34}-x_{34}y_{14}) \tag{4.4-48}$$

式中，$x_{ij}=x_i-x_j$，$y_{ij}=y_i-y_j$，x，y 为单元节点坐标，C_1，C_2 为两个系数，$C_1=\frac{2\pi}{3}(x_1+x_2+x_3)$，$C_2=\frac{2\pi}{3}(x_1+x_3+x_4)$。

计算出目标函数后，就可以利用黄金分割法进行锻件毛坯形状的优化。其优化设计步骤如图 4.4-31 所示。

9.2 优化设计实例

采用上述方法，分别对截面为 H 型轴对称锻件和平面应变形锻件的毛坯形状进行了优化设计。

（1）轴对称问题

本节以实际锻件超出预定理想直线部分的体积为目标函数，圆柱体毛坯的高径比为优化设计变量，目标是获得充填情况最好，同时飞边最小的终锻件。轴对称锻件和初始坯料尺寸分别如图 4.4-32 和图 4.4-33 所示。

设整个锻造过程等温，材料为 20# 钢，摩擦系数 $m=0.2$，上模速度为 -1.0 mm/s。由于锻件上下对称，故只取坯料的 1/4 进行优化。优化开始时优化设计变量（即圆柱体坯

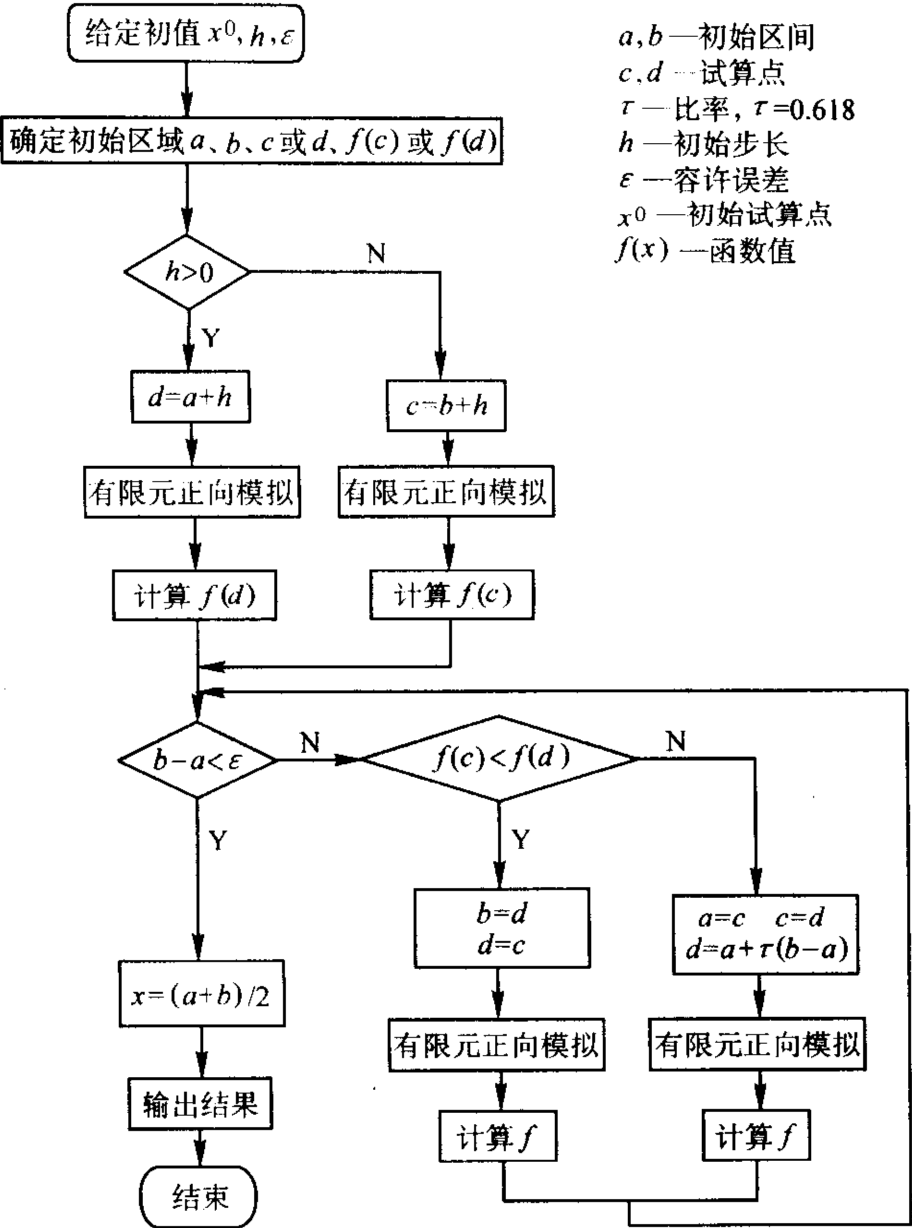


图 4.4-31　毛坯优化设计步骤流程图

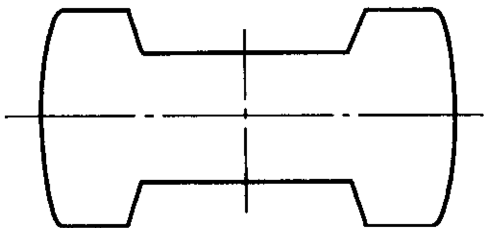


图 4.4-32　锻件图

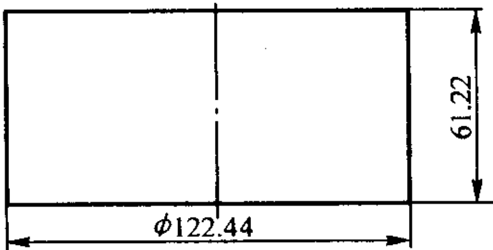


图 4.4-33　坯料尺寸

料的高径比）取为0.5，具体坯料尺寸为ϕ122.44 mm×61.22 mm。图4.4-34表示目标函数随优化迭代次数的变化情况。由图可看出，经过14次优化迭代，目标函数由2 200 mm³左右下降到最后的300 mm³，大约下降了80%，优化效果明显。图4.4-35显示了优化设计变量随优化迭代次数的变化情况。由图4.4-35可以看出，优化设计变量（即坯料的高径比）从初始值0.5开始升高，当到达最高点1.0左右后开始下降，然后在0.75左右振荡，最后收敛在0.75左右。优化趋势比较明显。

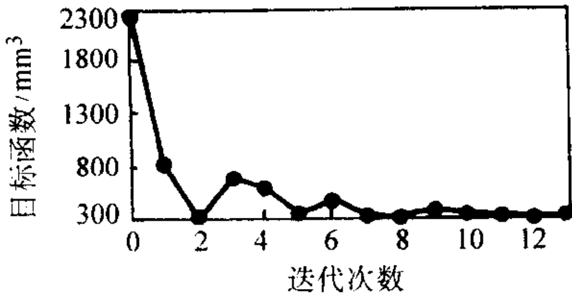


图 4.4-34　优化迭代中目标函数的变化

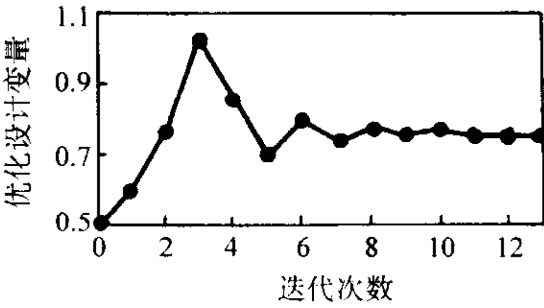


图 4.4-35　优化迭代中优化设计变量的变化

图4.4-36给出了不同优化迭代下的终锻件形状的比较，图4.4-37给出了优化过程中终锻件的变化情况，由图4.4-36和图4.4-37可以看出，第一次优化时得到的终锻件飞边较大，同时终锻模具型腔有很明显的充不满现象发生，而优化最后得到的终锻件不仅飞边减小了，而且终锻模具型腔完全充满，取得了明显的优化效果，达到了优化的目的。

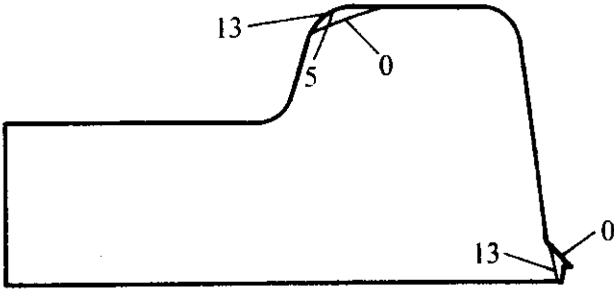


图 4.4-36　终锻件形状的比较

0，5，13—优化迭代次数

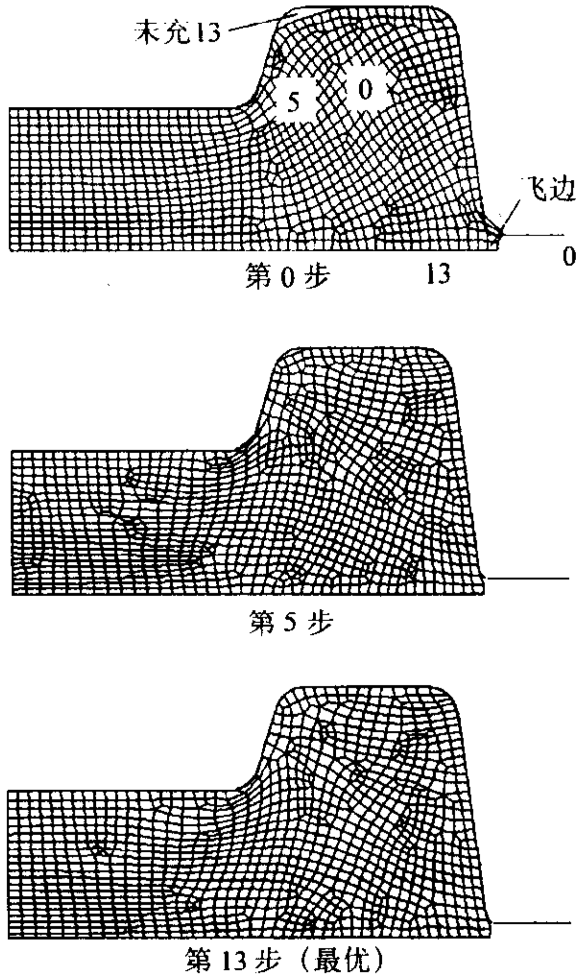


图 4.4-37　优化过程中终锻件形状的变化

（2）平面应变问题

以实际锻件超出预定理想直线部分的面积为目标函数，以毛坯的高宽比为优化设计变量，对二维平面应变变形问题的毛坯形状进行优化设计。如图4.4-38为终锻件形状，设整个过程等温，材料为20# 钢，摩擦系数$m=0.2$，上模速度为-1.0 mm/s。由于锻件上下对称，故只取坯料的1/4进行优化。优化开始时优化设计变量，即矩形截面坯料的高宽比取为0.8，具体坯料尺寸为26.909 98 mm×21.527 98 mm。图4.4-39显示了目标函数随优化迭代次数的变化情况。由图可看出，经过13次优化迭代，目标函数由9.5 mm²左右下降到最后的0.1 mm²左右，大约下降了90%，优化效果明显。图4.4-40显示了优化设计变量随优化迭代次数的变化

情况。可见优化设计变量（即坯料的高宽比）从初始值0.8开始升高，中间虽略有振荡，最终仍收敛在1.42左右。优化趋势比较明显。

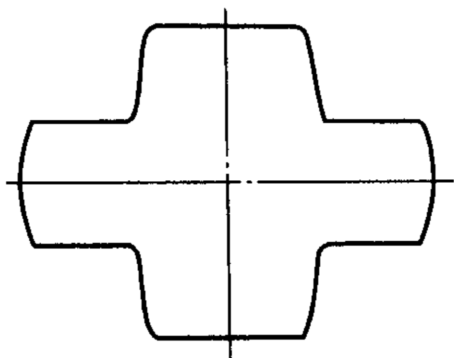


图4.4-38　终锻件形状

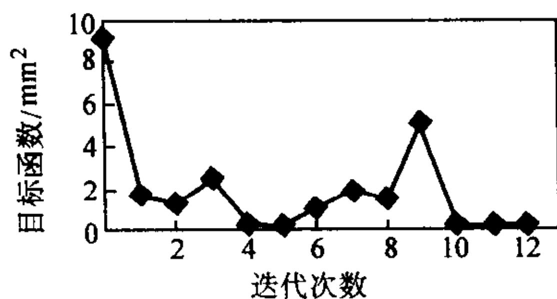


图4.4-39　优化迭代中目标函数的变化

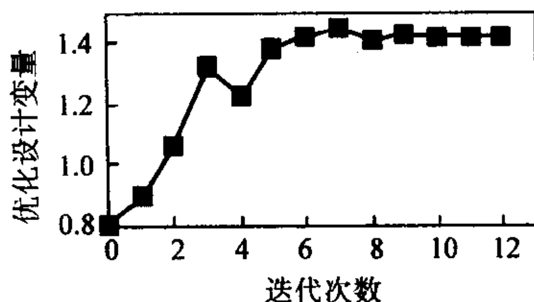


图4.4-40　优化迭代中优化设计变量的变化

图4.4-41给出了不同优化迭代下的终锻件形状的比较。可见第一次优化时得到的终锻件飞边较大，同时终锻模型腔有很明显的充不满现象发生，而优化最后得到的终锻件不仅飞边减小了，而且终锻模具型腔完全充满，取得了明显的优化效果，达到了优化的目的。

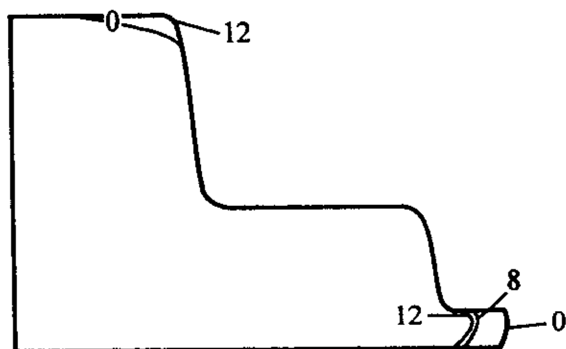


图4.4-41　终锻件形状的比较
0，8，12—优化迭代次数

10　锻件变形均匀性优化控制

锻件变形均匀性是衡量锻件质量的重要指标之一。锻件变形均匀性越好，其各部分的变形越趋于均匀，晶粒组织也越趋于一致，锻件的组织性能就越好。同时，锻件变形均匀性好，还可以避免锻件因局部变形剧烈而断裂。所以，提高锻件变形均匀性可以提高锻件的综合性能，提高锻件的使用寿命。

本节介绍以控制锻件变形均匀性为目标，通过优化预锻模具形状使锻件变形的不均匀性最小。用三次B样条曲线表示预成形模具的形状，以B样条曲线控制点的坐标作为优化设计变量，采用灵敏度分析的方法，对变形均匀性进行优化控制。

10.1　目标函数

以任意单元的等效应变与所有单元的平均等效应变的差的平方和作为衡量锻件变形均匀性的目标函数，其数学表达式如下：

$$\psi = \sum_{i=1}^{N}(\bar{\varepsilon}_i - \bar{\varepsilon}_{av})^2 \tag{4.4-49}$$

式中，$\bar{\varepsilon}_{av} = \dfrac{\sum_{j=1}^{N}\bar{\varepsilon}_j V_j}{\sum_{j=1}^{N} V_j} = \dfrac{\sum_{j=1}^{N}\bar{\varepsilon}_j V_j}{V_{tol}}$；$V_j$ 为单元 j 的体积；$\bar{\varepsilon}_j$ 为单元 j 的等效应变；$\bar{\varepsilon}_{av}$ 为所有单元的平均等效应变；N 为单元总数；V_{tol} 为所有单元的体积。

当 ψ 接近零时，实际获得的最终成形件的变形分布趋于均匀。优化问题就定义为通过使目标函数 ψ 取最小值的B样条曲线控制点的坐标来定义最佳的预成形模具形状。

10.2　灵敏度分析

根据式(4.4-49)给出的目标函数表达式，推导目标函数 ψ 对优化设计变量 p_1 的灵敏度如下：

$$\begin{aligned}\frac{\partial \psi}{\partial p_1} &= \sum_{i=1}^{N} 2(\bar{\varepsilon}_i - \bar{\varepsilon}_{av})\left(\frac{\partial \bar{\varepsilon}_i}{\partial p_1} - \frac{\partial \bar{\varepsilon}_{av}}{\partial p_1}\right) \\ &= \sum_{i=1}^{N} 2(\bar{\varepsilon}_i - \bar{\varepsilon}_{av})\left[\frac{\partial \bar{\varepsilon}_i}{\partial p_1} - \frac{1}{V_{tol}}\sum_{j=1}^{N}\left(\frac{\partial \bar{\varepsilon}_j}{\partial p_1}V_j + \bar{\varepsilon}_j\frac{\partial V_j}{\partial p_1}\right)\right]\end{aligned} \tag{4.4-50}$$

由式（4.4-50）可以看出，只要求出单元体积对优化设计变量的灵敏度$\partial V/\partial p_1$ 和单元等效应变对优化设计变量的灵敏度$\partial \bar{\varepsilon}/\partial p_1$，就可以得到目标函数对优化设计变量的灵敏度$\partial \psi/\partial p_1$。

（1）单元体积灵敏度$\partial V/\partial p_1$

单元体积对优化设计变量的灵敏度$\partial V/\partial p_1$ 可计算如下：

$$\frac{\partial V}{\partial p_1} = \frac{\partial V}{\partial \boldsymbol{X}}\frac{\partial \boldsymbol{X}}{\partial p_1} \tag{4.4-51}$$

式中，$\boldsymbol{X}$ 为单元节点坐标向量。

由式（4.4-51）可以看出，单元体积灵敏度的计算可以通过计算单元体积对单元节点坐标灵敏度和单元节点坐标对优化设计变量的灵敏度来获得。其中，单元节点坐标对优化设计变量的灵敏度的计算在本章第4节中已经给出，下面仅给出单元体积对单元节点坐标的灵敏度的计算。

对任意的四边形单元 i，其体积计算见式（4.4-48）。

下面分别计算轴对称问题和平面变形问题中的单元体积对节点坐标的灵敏度$\partial V/\partial \boldsymbol{X}$。

1）轴对称问题　根据式（4.4-48），轴对称单元体积对节点坐标的灵敏度计算如下：

$$\begin{aligned}\frac{\partial V}{\partial x_1} &= \frac{\pi}{3}[(x_{32}y_{12} - x_{12}y_{32}) - y_{32}(x_1 + x_2 + x_3) + \\ &\quad (x_{14}y_{34} - x_{34}y_{14}) + y_{34}(x_1 + x_3 + x_4)]\end{aligned}$$

$$\frac{\partial V}{\partial x_2} = \frac{\pi}{3}[(x_{32}y_{12} - x_{12}y_{32}) + (x_1 + x_2 + x_3)(y_{32} - y_{12})]$$

$$\begin{aligned}\frac{\partial V}{\partial x_3} &= \frac{\pi}{3}[(x_{32}y_{12} + x_{12}y_{32}) + y_{12}(x_1 + x_2 + x_3) + \\ &\quad (x_{14}y_{34} - x_{34}y_{14}) - y_{14}(x_1 + x_3 + x_4)]\end{aligned}$$

$$\frac{\partial V}{\partial x_4} = \frac{\pi}{3}[(x_1 + x_3 + x_4)(y_{14} - y_{34}) + (x_{14}y_{34} - x_{34}y_{14})]$$

$$\frac{\partial V}{\partial y_1} = \frac{\pi}{3}[x_{32}(x_1 + x_2 + x_3) - x_{34}(x_1 + x_3 + x_4)]$$

$$\frac{\partial V}{\partial y_2} = \frac{\pi}{3}(x_1 + x_2 + x_3)(x_{12} - x_{32})$$

$$\frac{\partial V}{\partial y_3}=\frac{\pi}{3}[-x_{12}(x_1+x_2+x_3)+x_{14}(x_1+x_3+x_4)]$$

$$\frac{\partial V}{\partial y_4}=\frac{\pi}{3}(x_1+x_3+x_4)(x_{34}-x_{14})$$

2) 平面变形问题　根据式（4.4-44），平面变形单元体积（面积）对节点坐标的灵敏度计算如下：

$$\frac{\partial V}{\partial x_1}=\frac{1}{2}(y_{34}-y_{32})$$

$$\frac{\partial V}{\partial x_2}=\frac{1}{2}(y_{32}-y_{12})$$

$$\frac{\partial V}{\partial x_3}=\frac{1}{2}(y_{12}-y_{14})$$

$$\frac{\partial V}{\partial x_4}=\frac{1}{2}(y_{14}-y_{34})$$

$$\frac{\partial V}{\partial y_1}=\frac{1}{2}(x_{32}-x_{34})$$

$$\frac{\partial V}{\partial y_2}=\frac{1}{2}(x_{12}-x_{32})$$

$$\frac{\partial V}{\partial y_3}=\frac{1}{2}(x_{14}-x_{12})$$

$$\frac{\partial V}{\partial y_4}=\frac{1}{2}(x_{34}-x_{14})$$

利用以上各式就可以计算出单元体积对单元节点坐标的灵敏度。

(2) 单元等效应变灵敏度

在有限元速度场迭代收敛后，单元等效应变的刷新公式如下：

$$\bar{\varepsilon}(t+\Delta t)=\bar{\varepsilon}(t)+\dot{\bar{\varepsilon}}(t)\Delta t \tag{4.4-52}$$

式中，$\bar{\varepsilon}(t+\Delta t)$为$(t+\Delta t)$时刻单元的等效应变；$\bar{\varepsilon}(t)$为 t 时刻单元的等效应变；$\dot{\bar{\varepsilon}}(t)$为 t 时刻单元的等效应变率。

将式（4.4-52）对优化设计变量 p_1 求偏导得

$$\frac{\partial\bar{\varepsilon}(t+\Delta t)}{\partial p_1}=\frac{\partial\bar{\varepsilon}(t)}{\partial p_1}+\frac{\partial\dot{\bar{\varepsilon}}(t)}{\partial p_1}\Delta t \tag{4.4-53}$$

上式就是有限单元等效应变对优化设计变量的灵敏度公式。当 $t=0$ 时，有$\partial\bar{\varepsilon}(0)/\partial p_1=0$。

对于任意一个单元，其等效应变率可表示如下

$$(\dot{\bar{\varepsilon}})^2=\dot{\boldsymbol{\varepsilon}}^{\mathrm{T}}\boldsymbol{D}\dot{\boldsymbol{\varepsilon}} \tag{4.4-54}$$

式中，$\dot{\boldsymbol{\varepsilon}}$ 为单元的应变率矢量，对于二维问题有 $\dot{\boldsymbol{\varepsilon}}=\{\dot{\varepsilon}_1,\dot{\varepsilon}_2,\dot{\varepsilon}_3,\dot{\varepsilon}_4\}^{\mathrm{T}}$；$\boldsymbol{D}$ 为数字对角线矩阵，对角线元素分别为 2/3，2/3，2/3 和 1/3。

将式（4.4-54）两边对优化设计变量 p_1 求偏导得

$$2\dot{\bar{\varepsilon}}\frac{\partial\dot{\bar{\varepsilon}}}{\partial p_1}=\frac{\partial\dot{\boldsymbol{\varepsilon}}^{\mathrm{T}}}{\partial p_1}\boldsymbol{D}\dot{\boldsymbol{\varepsilon}}+\dot{\boldsymbol{\varepsilon}}^{\mathrm{T}}\boldsymbol{D}\frac{\partial\dot{\boldsymbol{\varepsilon}}}{\partial p_1} \tag{4.4-55}$$

所以，欲求单元等效应变率灵敏度$\partial\dot{\bar{\varepsilon}}(t)/\partial p_1$，需先计算出单元各应变率分量的灵敏度$\partial\dot{\varepsilon}(t)/\partial p_1$。

对任意单元，其应变率矢量可表示如下

$$\dot{\boldsymbol{\varepsilon}}=\boldsymbol{B}\boldsymbol{V} \tag{4.4-56}$$

式中，$\boldsymbol{B}$ 为单元的应变率矩阵；$\boldsymbol{V}$ 为单元节点的速度矢量。在式（4.4-56）两边对优化设计变量 p_1 求偏导得

$$\frac{\partial\dot{\boldsymbol{\varepsilon}}}{\partial p_1}=\frac{\partial\boldsymbol{B}}{\partial\boldsymbol{X}}\frac{\partial\boldsymbol{X}}{\partial p_1}\boldsymbol{V}+\boldsymbol{B}\frac{\partial\boldsymbol{V}}{\partial p_1} \tag{4.4-57}$$

式中，$\boldsymbol{X}$ 为单元的节点坐标矢量。

可见，一旦确定了节点速度对优化设计变量的灵敏度$\partial V(t)/\partial p_1$和节点坐标对优化设计变量的灵敏度$\partial X(t)/\partial p_1$，即可得到单元等效应变率灵敏度$\partial\dot{\bar{\varepsilon}}(t)/\partial p_1$，其中$\partial V(t)/\partial p_1$和$\partial X(t)/\partial p_1$的计算参见本章第 3 节。最后根据式（4.4-50）、式（4.4-51）和式（4.4-53）求出目标函数对优化设计变量的灵敏度$\partial\psi/\partial p_1$。

锻件变形均匀性优化设计的步骤类似于本章第 6 节中介绍的模具形状优化设计步骤。

10.3　应用实例

图 4.4-42 所示为一典型的 H 形截面轴对称肋－腹件。在轴对称锻件中，肋－腹件大约占 50%，所以此类锻件具有一定的代表性。

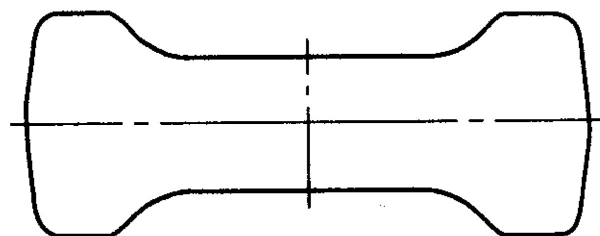


图 4.4-42　终锻件图

假设整个过程等温，材料为理想黏塑性体，其应力应变关系为$\bar{\sigma}=33.99\,\dot{\bar{\varepsilon}}^{0.145}$，摩擦系数 $m=0.2$，上模速度为 -1.0 mm/s。由于为轴对称变形，且上下对称，故只取坯料的 1/4 进行优化。在优化中，采用 10 个 B 样条控制点描述预锻模形状，因此就有 7 条 B 样条曲线和 14 个优化设计变量。

图 4.4-43 所示为优化迭代过程中预成形件、预成形模具形状和终锻件、终锻模具形状的变化情况。当优化结束后，目标函数值由最初的 1.471 降到了 1.167，降低了大约 20%，变形均匀性的优化效果比较明显。

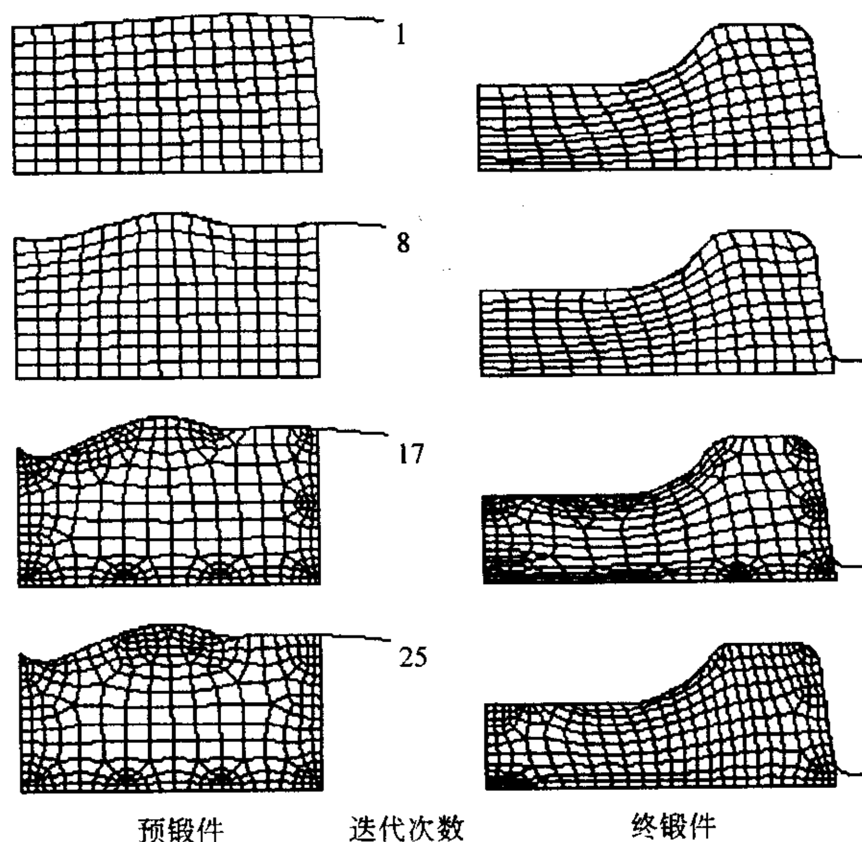


图 4.4-43　优化迭代过程中预锻件与终锻件形状的变化

编写：赵国群（山东大学）

第5章　金属塑性成形过程的微观组织优化

世界经济的发展使机械制造业的竞争日益激烈，高效、低耗、便捷地制造产品已成为企业提高自身竞争力的追求目标。在机械制造业中占据重要位置的金属塑性加工领域，由于传统的加工技术存在着严重的材料浪费和产品质量问题，因而近年来，少无切削精密成形技术受到高度重视，以净形锻造生产产品为目标的预成形模具优化设计理论和方法成为学术界研究的热点。但锻造过程是一个极其复杂的过程，由于其本身的工艺特点（高温、大变形）及大多数零件的几何形状相当复杂，在实际锻造生产中，锻出的产品可能会出现严重的质量问题，如充不满、折叠和局部变形严重等，从而将严重影响零件的力学性能。在工艺设计与控制方面，只有达到“形”与“神”的完美统一，即既要满足产品形状的基本要求，又能达到产品性能的更高要求，才能充分发挥金属塑性成形技术的优越性。而直接决定产品机械性能的便是金属材料在塑性成形过程中的微观组织演变及最终产品的组织结构。因此，近年来关于金属塑性成形过程的微观组织模拟与优化的研究已取得了重要进展。

众所周知，在热成形过程中，材料经历一系列的微观组织变化，如动态/静态恢复与再结晶及晶粒长大。了解影响产品微观组织演变的工艺因素对控制产品的力学性能是非常重要的，因为所有这些因素都会影响成形过程（如成形力、温度分布、应变及应变速率和金属的流动应力等）和产品的质量（包括强度、延展性和硬度等）。但是，由于大部分参量随着热变形的进行而随时发生变化，所以建立热成形过程参数和材料微观组织演变之间的量化关系及其基本算法是微观组织模拟和优化的难点。

有限元方法在金属塑性成形领域中，对给定的工艺条件下在预测金属流动规律和产品宏观形状方面已经取得了广泛应用，使传统的依赖于设计者经验的工艺和模具设计从经验走向了科学。将工程优化理论和方法引入到金属塑性成形数值模拟技术中，又一次使金属塑性成形工艺和模具设计方法学发生了巨大变化。数值模拟与优化技术相结合，能够自动地快速获得模拟的最终目标，提高了数值模拟的实效性。在完成工艺模拟、模具优化两级跳跃后，金属塑性成形研究领域的学者又在尝试微观组织模拟和优化的第三级跳跃，使金属塑性加工技术真正能够实现在控制产品形状的同时进一步提高产品的性能。因此，热成形过程的微观组织模拟与优化是在金属塑性成形领域中有限元数值模拟技术的广泛应用和工程优化理论在金属塑性成形数值模拟技术中取得重要进展的基础和前提下发展起来的。

1　微观组织演变模型

金属热变形过程中，伴随着材料的变形，微观组织也会发生变化，发生动/静态再结晶和晶粒长大。

（1）动态再结晶

动态再结晶发生在变形过程当中，在流动应力达到峰值之前，动态再结晶发生的临界应变与对应于最大流动应力的应变之间的差值相对较小。如果发生再结晶，就会形成没有预应变的晶粒。描述动态再结晶机理的关系式如下：

$$\left.\begin{aligned} \varepsilon_c &= 4.76\times10^{-4}\exp\left(\frac{8\,000}{T}\right)\\ d_{dyn} &= 22\,600Z^{-0.27}\\ X_{dyn} &= 1-\exp\left[-\ln2\left(\frac{\varepsilon-\varepsilon_c}{\varepsilon_{0.5}}\right)^2\right]\\ \varepsilon_{0.5} &= 1.144\times10^{-5}d_0^{0.28}\dot{\varepsilon}^{0.05}\exp\left(\frac{6\,420}{T}\right) \end{aligned}\right\}\quad(4.5\text{-}1)$$

式中，ε_c 为动态再结晶开始时的临界应变；T 为绝对温度；Z 为 Zenor - Hollomon 参数，它符合如下关系 $Z=\dot{\varepsilon}\exp(Q/RT)$；$d_{dyn}$为独立于原始晶粒大小 d_0 的动态再结晶晶粒大小，只由 Z 参数决定；Q 为从实验中得到的活动能量；R 为气体常数；X_{dyn}为动态再结晶的体积分数；$\varepsilon_{0.5}$为再结晶过程到50%时的应变；d_0 为原始晶粒大小。

（2）静态再结晶

将变形组织置于高温中，一段时间后发生静态再结晶。理论上，当变形材料回火处理时静态再结晶才会发生，但起初动态再结晶与静态再结晶实际上没有什么不同。唯一不同之处在于动态再结晶与变形能量有关，而静态再结晶与前一过程中的变形程度、回火温度等有关。在卸载状态下发生的静态再结晶的微观组织演变如下：

$$\left.\begin{aligned} X_{st} &= 1-\exp\left[-\ln2\left(\frac{t}{t_{0.5}}\right)^2\right]\\ t_{0.5} &= 2.2\times10^{-12}S_v^{-0.5}\dot{\varepsilon}^{-0.2}-\exp\left(\frac{30\,000}{T}\right)\\ S_v &= \frac{24}{\pi d_0}\ (0.419e^{\varepsilon}+0.155e^{-\varepsilon}+0.143\,3e^{-3\varepsilon})\\ d_{st} &= \frac{5}{(S_v\varepsilon)^{0.6}} \end{aligned}\right\}\quad(4.5\text{-}2)$$

式中，X_{st}为静态再结晶的体积分数；$t_{0.5}$为静态再结晶的容积度为50%所用时间；S_v 为单位体积的晶界面积；d_{st}为静态再结晶的晶粒大小。

在高温下静态再结晶也很迅速，较大的变形程度和原始晶粒大小也影响着静态再结晶的过程。

（3）晶粒生长

再结晶过程完成后的新组织是亚稳定的。为获得更稳定的组织，需要释放晶粒内部的潜能。因此，再结晶完成（X_{dyn}或 $X_{st}>0.95$）后，要发生进一步的晶粒长大来减少单位体积内的晶界面积。晶粒长大的等式如下。

动态再结晶晶粒长大

$$d=d_{dyn}+1.1\gamma\ (d_{pd}-d_{dyn})\qquad(4.5\text{-}3)$$

式中，$\gamma=1-\exp[-295\dot{\varepsilon}^{0.1}\exp(-8\,000t/T)]$，$d_{pd}=5\,380\exp(-6\,840/T)$。

静态再结晶晶粒长大

$$d^2=d_{st}^2+1.44\times10^{12}t\exp\left(-\frac{Q}{RT}\right)\qquad(4.5\text{-}4)$$

式中各符号意义如前。

2　金属塑性成形过程微观组织模拟与优化的基本算法

热成形过程的微观组织模拟是在金属塑性成形领域中有限元数值模拟技术的广泛应用和工程优化理论在金属塑性成形数值模拟技术中取得重要进展的基础和前提下发展起来

的。最初，关于微观组织的模拟通常是采用热力耦合的分析方法，通过计算热力学参数来间接地模拟材料的微观结构。后来，在形状优化设计中较多采用的微观遗传算法和灵敏度分析方法被引入微观组织模拟中，直接以微观结构参量作为目标函数的设计变量，直接实现微观组织的优化。由于热力耦合分析方法在前面在章节中已经做了介绍，本节只简要介绍遗传算法和灵敏度分析方法。

2.1　遗传算法

遗传算法（genetic algorithm）是由 John Holland 等人提出，首先应用于自动控制、神经网络等学科，后来作为一种优化方法被引入到金属塑性成形领域中。与常规优化方法不同，基因遗传算法将优化问题视为一种生存环境，将优化过程视为一种进化过程，是一种模拟生命进化机制的搜索和优化算法。群体、染色体、适应值和模式是遗传算法中的常用概念。

群体（population）是指每一个迭代步中参与操作的解的集合。与一般的优化算法不同，遗传算法不是从单个个体的属性来确定搜索方向，而是从群体的总体属性中确定搜索方向。群体中解的个数称为群体规模，每一个迭代步中的群体称为一代。

染色体（chromosome）是指将设计变量根据某种规则转换成的代码串，这与自然界中的过程一样，遗传过程中实际操作的是染色体，而非其显式形式。代码串中的每一位称为基因。

适应值函数（fitness function）是指目标函数经过某种变换后得到的数值。适应值反映了对应于此目标函数的染色体对环境的适应程度，适应值高，表明染色体对环境的适应程度高，反之亦然。

模式（schema）表示基因中某些特定位相同的结构。简单的遗传操作实际包含了并行进行的、丰富的模式隐含操作，使得遗传算法得以有效进行。

遗传算法利用“优胜劣汰”的生物进化机制选优，优良的物种进化，而低劣的物种则随着种群的繁衍而灭绝，其基本操作是对优化设计变量进行编码、复制、杂交和变异，如图 4.5-1 所示。

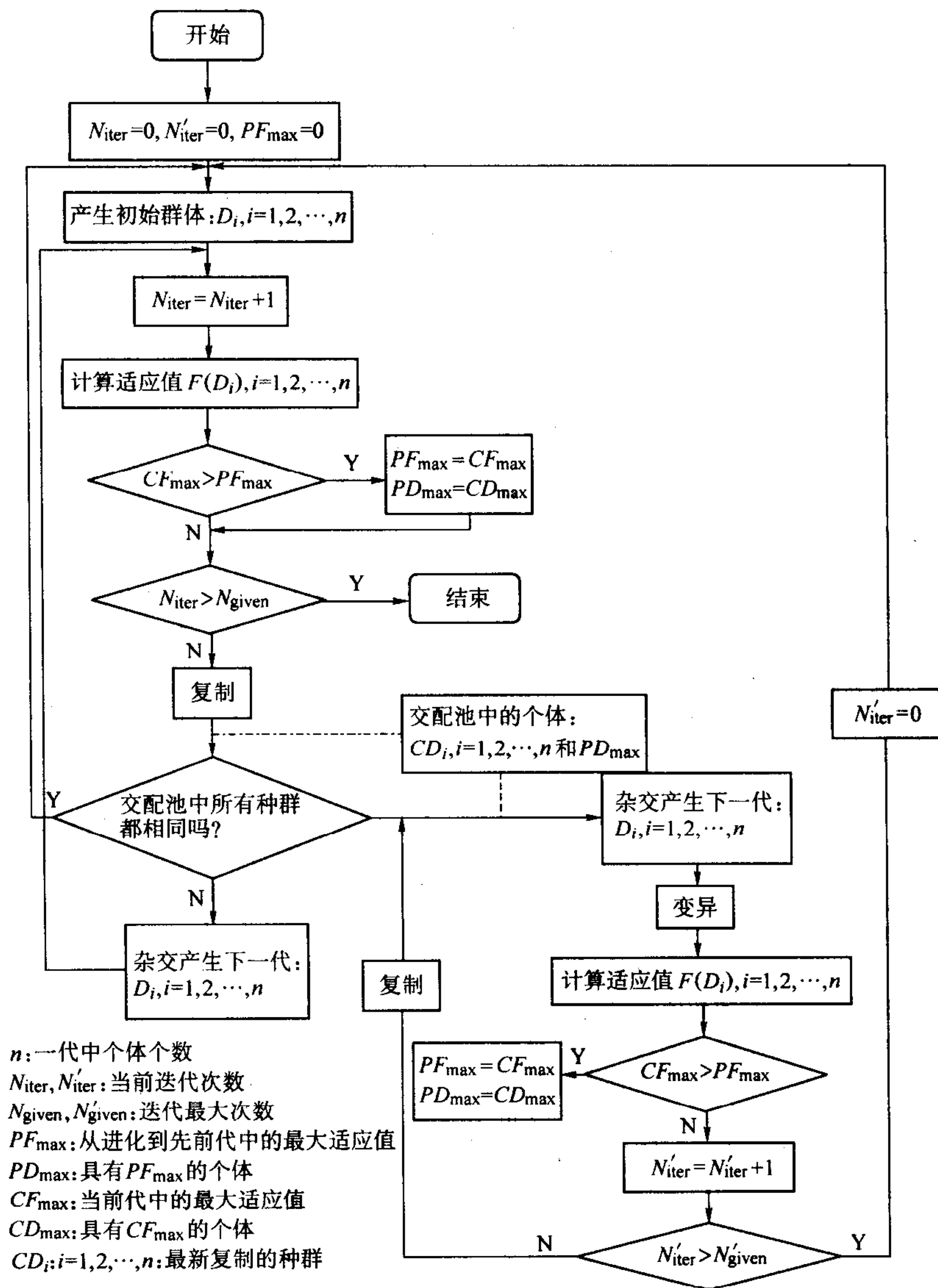


图 4.5-1　遗传算法的实施过程

1）编码　编码是将设计变量转化为染色体的过程。一般来说，是将设计变量转化为二进制字符串。在进行二进制符号编码时首先应确定每个设计变量的编码长度 k，一般取 $k=4$。如果设计变量个数为 n，则染色体的代码长度为 $k*n$。另外，编码前还需确定每个设计变量的变化范围。

2）复制　能繁殖后代的几列字符串由当前种群中选择，一个字符串 $D_i(i=1,2,\cdots,n)$ 是否被选取取决于 $P(D_i)$ 的数值。$P(D_i)$ 的数值由下式确定：

$$P(D_i)=\frac{F(D_i)}{\sum_{i=1}^{n}F(D_i)} \tag{4.5-5}$$

式中，F 为标度适应值，定义如下：

$$F(D_i)=\frac{f_{avg}}{f_{max}-f_{avg}}f(D_i)+\frac{f_{max}-2f_{avg}}{f_{max}-f_{avg}}f_{avg} \tag{4.5-6}$$

式中，$f(D_i)=C-\psi_0(D_i)$，$f_{avg}=\frac{1}{n}\sum_{i=1}^{n}f(D_i)$，$f_{max}=\max f(D_i)$，$\psi_0$ 是目标函数值，C 是一个远超过任何目标函数可能值的一个较大的正数。

需要指出的是，上述适应值的定义对 $f(D_i)$ 中差距的扩大有一定影响。当 ψ_0 差距很小时，这种影响可防止同一种群中的系列串得出相似的 $P(D_i)$ 值，在它们中建立起一个比较明显的继承关系。被选中而进行复制的字符串表示为 $CD_i(i=1,2,\cdots,n)$，如图4.5-1所示。

3）杂交　杂交分两步。首先，在整个进化过程中，直到当前的一代，从交配池中的 $n+1$ 个字符串即最新复制的 n 个字符串加上具有最大适应值的一个字符串中选出 $n/2$ 对字符串；其次，每一对字符串进行杂交而产生下一代的两个后代，即：对每一对字符串，二进制数的位置 k 由第一位与第（$m-1$）位之间选择（m 是该字符串的位数），然后通过在 $k+1$ 与 m 间的所有位进行对换而创建两个新字符串。

4）变异　任何一个串中的二进制位在经过杂交后会依据一个给定的每个位变换的可能性而随意变换，变换过程采用随机数生成器来实现。变异被用来防止设计变量的过早集中，即防止所有的交配池中 $n+1$ 个字符串的二进制位的结构在进化早期变为一致，但变异可导致快速收敛，因此，有些研究学者在应用微观遗传算法时不采用变异手段，而在算法系统中发生这种过早收敛时会触发一个也由杂交、变异、复制组成的次进化圈，以寻找更好的 PD_{max} 值。这样，进化就可由最新生成的一代和 PD_{max} 而得以延续。

2.2　灵敏度分析方法

基于灵敏度分析的优化设计方法属于梯度型优化设计方法。该方法将预成形设计和模具设计问题处理为优化问题，用严密的数学公式进行描述，将优化问题的目标函数定义为一组给定设计变量中所希望的最终状态和数值计算状态之间的误差的某种程度。灵敏度分析方法需要计算目标函数对预成形参数的灵敏度信息（导数），然后采用高效的优化算法（如BFGS算法）进行优化设计。求解目标函数对设计变量的灵敏度通常有两种方法：一种是根据有限元刚度方程推导灵敏度的值；另一种是使用直接差分法求解灵敏度的值。前一种方法的推导过程复杂，但求解的灵敏度值精度高，优化迭代收敛速度高，优化次数少；后一种方法计算简单，但灵敏度值精度低，优化迭代次数多，有时会出现优化不收敛的现象。

有关灵敏度分析方法的基本理论和在优化中的实施步骤详见本篇的第4章。

3　微观组织模拟与优化实例

3.1　H形锻件成形过程组织模拟与优化

在热成形中，对热传导特性的控制对于成功地生产产品是十分重要的。也就是说，在金属流动控制中必须考虑热传导的影响。下面以采用热力学有限元耦合分析模型通过遗传算法来预测发生在工件与模具中的热传导和金属流动行为为例来介绍如何通过工艺优化设计来控制微观组织结构的。

（1）热力耦合有限元分析模型

在修正的拉格朗日有限元方法中，假设工件材料呈现刚（黏）塑性变形行为，模具是刚性体，则优化问题的独立状态变量可减少为位置矢量（$\boldsymbol{X}$）、等效应变（$\bar{\varepsilon}$）、节点速度矢量（$\boldsymbol{V}$）、单元体静水压力矢量（$\boldsymbol{P}$）、工件节点温度（T）、模具节点温度（T_D）等。从动量与能量守恒的角度来讲，可以得出以下有限元的平衡公式。

1）从工件的受力平衡角度得出如下平衡公式

$$f_1^{(i)}(\boldsymbol{X}^{i-1},\varepsilon^{i-1},\boldsymbol{V}^i,\boldsymbol{P}^i,T^i,\boldsymbol{D})=0 \tag{4.5-7}$$

2）从工件的能量守恒角度得出如下平衡公式

$$f_2^{(i)}(\boldsymbol{X}^{i-1},\varepsilon^{i-1},\boldsymbol{V}^i,T^i,T_D^i,\boldsymbol{D})=0 \tag{4.5-8}$$

3）从模具的能量守恒角度得出如下平衡公式

$$f_3^{(i)}(\boldsymbol{X}_D^{i-1},\boldsymbol{V}^i,T^i,T_D^i,\boldsymbol{D})=0 \tag{4.5-9}$$

其中 $i=1,2,\cdots,k$ 表示成形过程中的第 i 个时间步，$\boldsymbol{D}$ 表示指定的工艺过程参数，可由设计变量决定，包括机械工艺参数（预锻及终锻模具形状、接触面摩擦、模具速度、模具行程及工件材料的流动应力等）和热工艺参数（锻件与模具的初始温度、模具与工件界面间及环境温度与冷却介质间的热传导系数、工件与模具的热传导率及比热容、成形过程中的时间间隔及热边界条件等）。三个等式形成了关于状态变量的一个耦合系统，表达了工件热力学行为与机械行为之间以及工件与模具热性能之间的密切关系，即：工件材料流动应力随温度不断变化，同时材料的塑性变形和材料与模具在接触面上相对滑动会产生热量，通过工件表面就会发生热传导。该耦合系统中状态变量参数值的求解可采用迭代法进行。

（2）优化设计方程

获得所要求的产品质量和生产效率等问题可以转化为确定有关的设计变量值而使某一特定的函数值最小化即目标函数最小化问题。目标函数一般可以用状态变量与设计变量的函数来表示，即

$$\psi_0=\psi_0(\boldsymbol{V}^1,\boldsymbol{P}^1,T^1,T_D^1,\boldsymbol{V}^2,\boldsymbol{P}^2,T^2,T_D^2,\cdots,\boldsymbol{V}^k,\boldsymbol{P}^k,T^k,T_D^k,\boldsymbol{D}) \tag{4.5-10}$$

此外，金属塑性成形过程需满足某些特定的条件和要求，即某些设计变量的值必须限制在一个允许的范围内，如成形载荷不能超出设备的负载能力等，这些限制被称为约束条件，可大致表述如下：

$$\left.\begin{array}{l}\boldsymbol{D}=\{d_j,d_{j\min}<d_j<d_{j\max}\}(j=1,2,\cdots,r)\\ \psi_l=0(l=1,2,\cdots,p)\\ \psi_k\leqslant 0(k=p+1,p+2,\cdots,q)\end{array}\right\} \tag{4.5-11}$$

在这里，ψ_l、ψ_k 与 ψ_0 一样，是设计变量的函数。

由于相应的状态变量可由等式（4.5-7）～式（4.5-9）求得，所需解决的优化问题可以描述如下：在所有的设计变量 D_i 中，能满足设计约束等式（4.5-11）而使目标函数式（4.5-10）取最小值的即为所要求的解。

（3）优化结果

金属在高温下成形的主要特点是：因位错密度增加而产生加工硬化，同时，又因大量的动态与静态的再结晶得以软化，成形过程影响着产品最终所具有的冶金与机械特性，因此，在成形设计中对再结晶性能进行控制是十分重要的。由于最终产品中再结晶晶粒尺寸的分布与产品中的温度分布密

切相关，因此，可以通过控制成形工艺来获得产品中均匀的温度分布以实现均匀的微观组织结构。

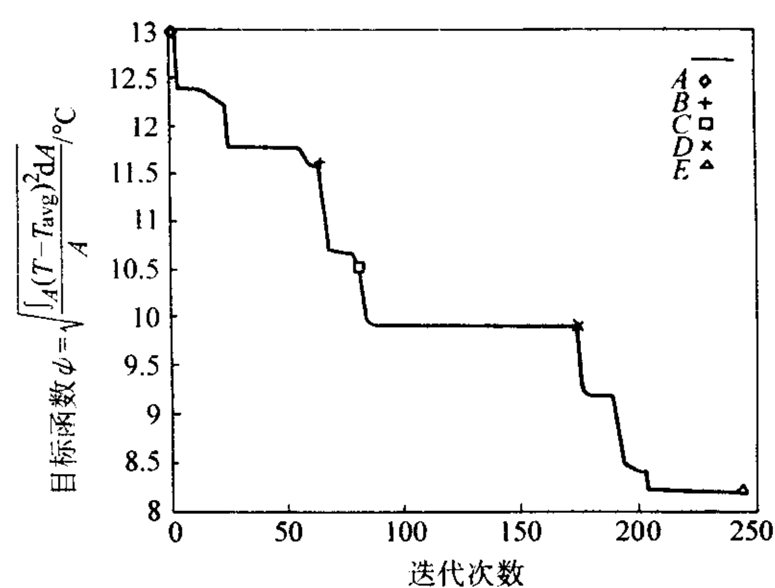
以圆柱形坯料加工 H 形轴对称产品为例，对直接终锻和先预成形再终锻两种成形工艺进行对比分析。在工艺参数中考虑的设计变量为：材料特性参数、产品及模具的形状尺寸、模具速度及行程、坯料及模具的初始温度、环境温度等。目标函数定义为产品中温度分布的不均匀程度，即

$$\psi = \sqrt{\frac{\int_A (T - T_{avg})^2 \mathrm{d}A}{A}}$$

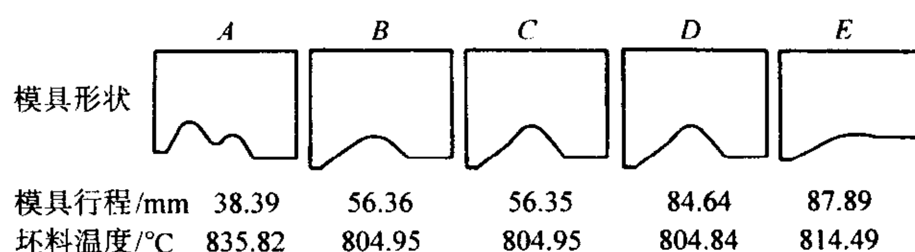
$$T_{avg} = \frac{1}{A}\int_A T\mathrm{d}A \tag{4.5-12}$$

式中，T 代表整个成形过程中横截面上的温度分布；A 代表横截面的面积。

目标函数的收敛过程及设计变量值（预成形模具形状、加载行程及温度等）的优化进程如图 4.5-2 所示。与未经预锻的成形相比较，温度分布的一致性得到明显提高，特别是在横截面的中心部分，如图 4.5-3 所示。



(a) 目标函数的收敛过程



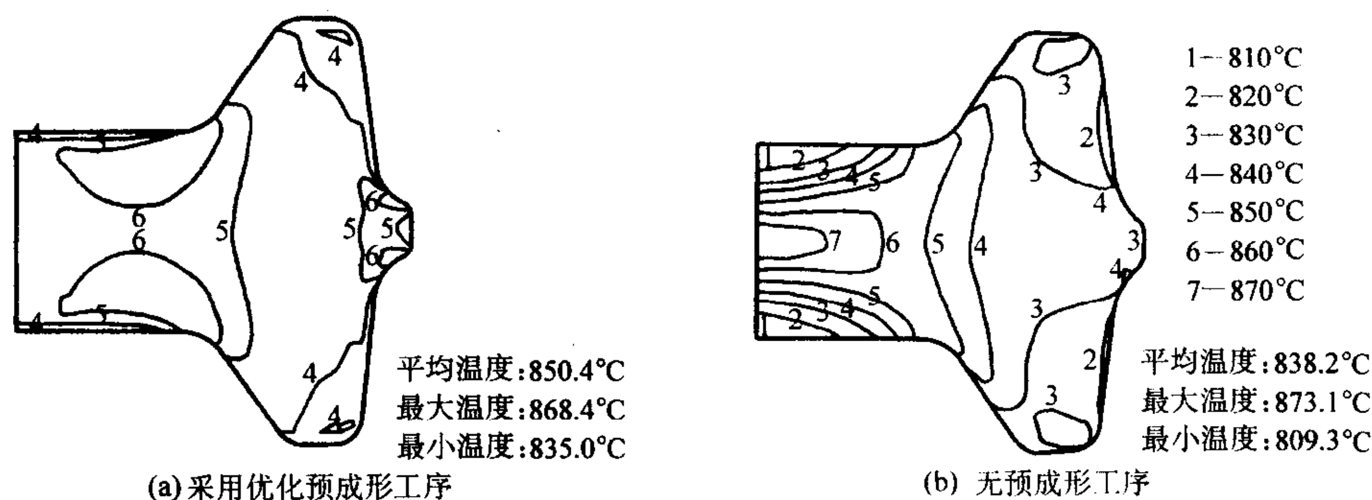
	A	B	C	D	E
模具行程/mm	38.39	56.36	56.35	84.64	87.89
坯料温度/°C	835.82	804.95	804.95	804.84	814.49

(b) 设计变量的优化进程

图 4.5-2　目标函数收敛及设计变量值的优化进程

3.2　热挤压成形过程的微观组织模拟与优化

挤压工艺在金属塑性成形领域中的应用十分广泛，在实际生产应用中对挤压产品性能的要求也越来越高。因此，通过控制宏观工艺参数实现微观组织优化以提高挤压产品的性能具有重要的应用价值。文献［20］利用刚（黏）塑性有限元法及微观组织演变模型以热挤压过程为例研究了通过挤压模具形状优化设计实现挤压产品均匀而良好的微观组织。具体过程如下。



(a) 采用优化预成形工序　(b) 无预成形工序

图 4.5-3　H 形产品成形后的温度分布

（1）微观组织演变预测算法

在热成形过程中，为了预测材料的微观组织演变，就要了解材料的冶金性能和热力学性能对材料的影响。为此，将刚（黏）塑性有限元与微观组织模型进行耦合，其计算流程图如图 4.5-4 所示。如果等效应变大于临界应变（$\varepsilon_{eff} \geqslant \varepsilon_c$），就会发生动态再结晶；若等效应变小于临界应变（$\varepsilon_{eff} < \varepsilon_c$），就会发生静态再结晶。晶粒的长大取决于再结晶体积百分比。如果再结晶体积分数大于 95%，就必须考虑晶粒长大。

当时间间隔过小或在给定温度下的应变值不足以产生完全再结晶时，就会发生部分再结晶。就部分再结晶而言，其微观组织由两部分组成。一是 X，代表发生动态或静态再结晶的部分，另一部分是（$1-X$），代表未发生再结晶的部分。则平均晶粒尺寸为

$$d_{avg} = d_{avg,i-1}(1 - X_i) + d_{avg,i}X_i \tag{4.5-13}$$

式中，$d_{avg,i-1}$ 代表再结晶前的平均晶粒尺寸；$d_{avg,i}$ 代表再结晶后的平均晶粒尺寸；X_i 为再结晶的体积分数。

（2）模具形状优化设计步骤

采用有限元法和优化设计技术对热挤压成形过程的设计变量进行反复计算，直至得到一个能够使最终产品具有均匀的微观组织的最优解。优化设计变量为控制挤压模具形状的位置点的坐标，模具形状以这些控制点的 B 样条曲线描述。其基本过程及步骤如下：

1）首先假定初始的设计变量［$P_{(1,1)}, P_{(1,2)}, \cdots, P_{(1,n)}$，其中 1 代表初始步，$n$ 代表设计变量的数目］，并由 B 样条曲线定义初始模具形状。然后计算初始目标函数值 G_1。

2）计算新的一组设计变量的目标函数值 G_i。

3）将目标函数值与前一步的目标函数值进行比较：如果 $|G_i - G_{i-1}| \leqslant \delta$（$\delta$ 是给定的任意小的一个正数），则目标函数收敛，程序结束；如果 $|G_i - G_{i-1}| > \delta$，用优化设计方法刷新设计变量值，返回到第二步中，重复计算第三步，直到目标函数收敛为止。

（3）目标函数

优化设计过程中的主要问题是找到合适的目标函数。为了获得性能良好的锻件，应以终锻件的微观组织为目标。因此，使得平均晶粒尺寸与实际晶粒尺寸之差最小定义为目标函数。其表达式如下：

$$\left.\begin{aligned} G &= \frac{\sqrt{\sum (d_i - d_{avg})^2 V_i}}{V_{total}} \\ d_{avg} &= \frac{\sum d_i V_i}{V_{total}} \end{aligned}\right\} \tag{4.5-14}$$

式中，V_{total}代表挤压件的体积；d_i 和 V_i 分别代表第 i 个单元体的晶粒尺寸和单元体体积。因此，使得目标函数 G 取最小值的模具形状就是可获得均匀微观组织的优化的模具形状。

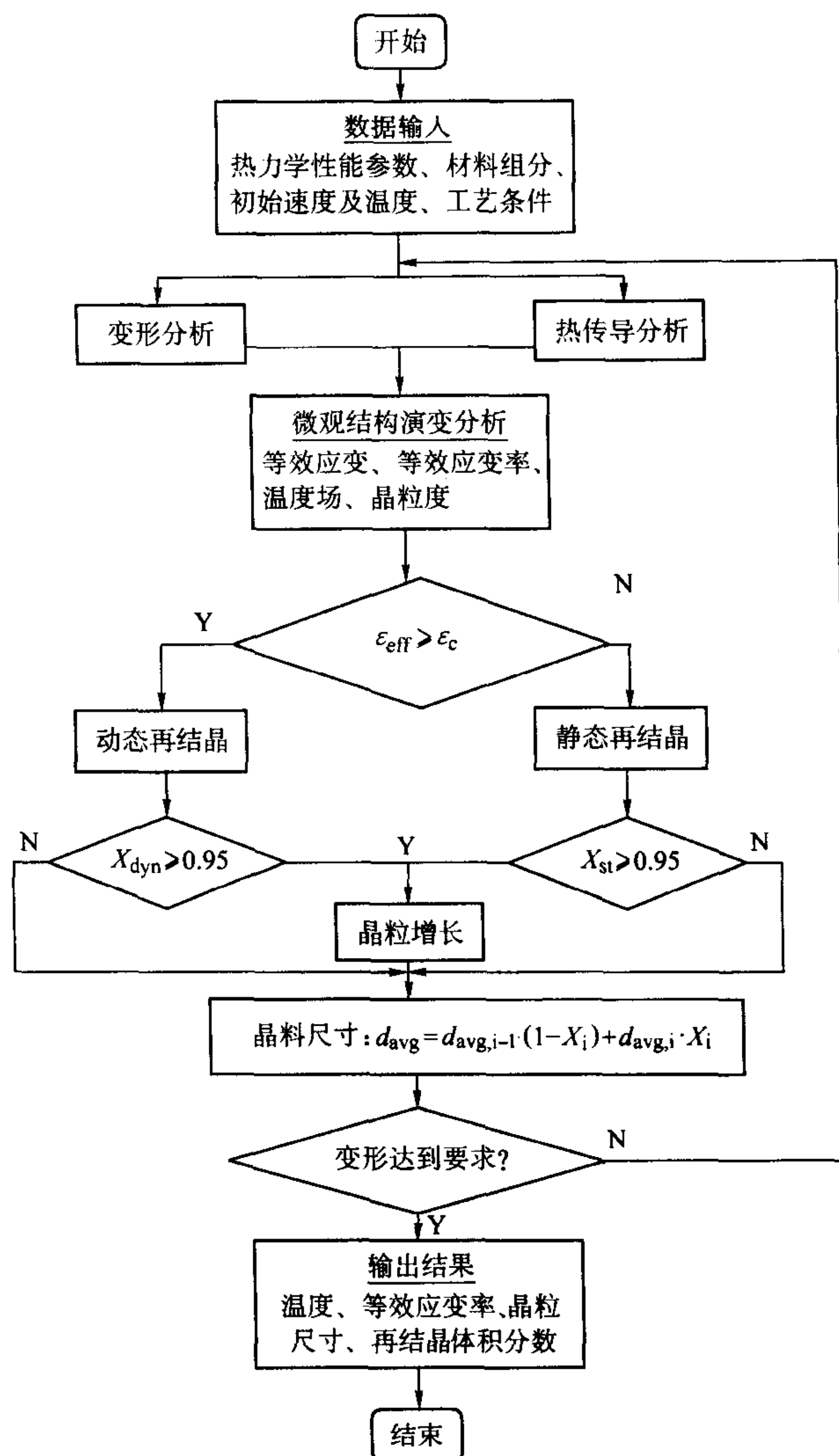


图 4.5-4　有限元法预测微观组织流程图

（4）数值模拟条件及结果

为了预测微观组织演变过程，采用了刚（黏）塑性有限元方法。材料的流动应力是应变、应变速率和温度的函数。挤压速度为 30 mm/s，摩擦系数为 0.3，初始晶粒尺寸约为 300 μm，坯料材料为 AISI1030（含碳量 0.28%～0.34%，含锰量 0.6%～0.9%），模具材料为 3Cr2W8V，模具初始温度为300℃，初始坯料加热到 1 200℃并保温 10 min 后进行挤压。

随着迭代次数的增多，晶粒尺寸的平均值逐渐减小，当第 15 次迭代后该值成为一个常数。因此，第 15 次迭代生成的模具形状即为最优化的结果，其晶粒尺寸的优化过程如图 4.5-5 所示。从模拟和优化迭代过程看，迭代收敛时的应变速率分布更加均匀。可见，动态再结晶的晶粒尺寸主要取决于应变速率的大小。虽然在挤压过程中坯料的外部温度较低，但由于该处具有较高的应变和应变速率，因此再结晶体积分数在外部区域比内部区域大。而内部温度虽然较高，但因其应变和应变速率较低，所以再结晶进行的并不充分。结果，外部区域的晶粒尺寸要小于中心部分。尽管在迭代过程中，模具形状发生变化时，对温度场几乎没有什么影响，而应变和应变速率的分布却发生了很大变化。因此，影响热挤压过程中再结晶的主要因素并不是温度，而是应变和应变速率。晶粒尺寸在整个迭代过程中从 195 μm 到 207 μm 之间的变化规律也反映出同样的结果，即模具形状的改变很难影响平均晶粒尺寸的大小，若要控制平均晶粒尺寸，不仅要改变模具形状，而且要控制其他工艺条件，如挤压速度、挤压温度等。

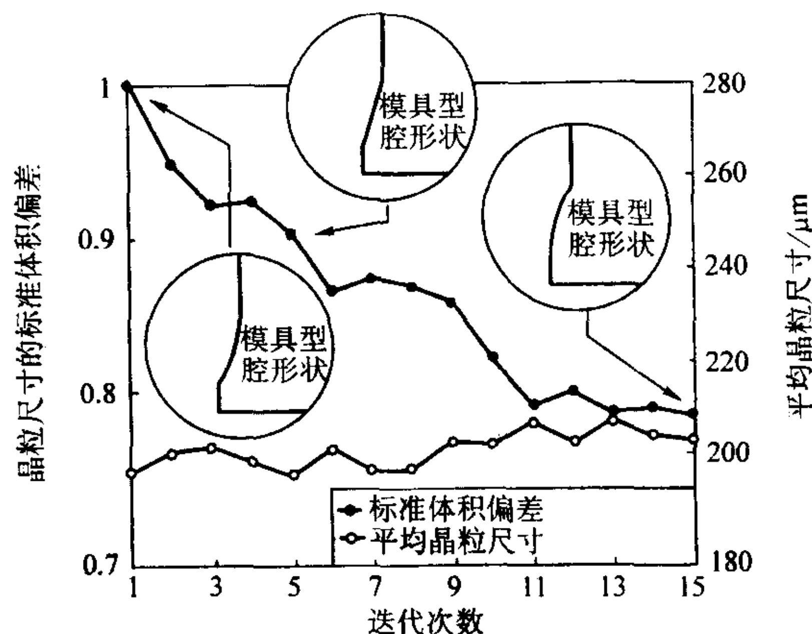


图 4.5-5　迭代过程中晶粒尺寸偏差及平均晶粒尺寸的变化

3.3　基于灵敏度分析方法的微观组织模拟与优化

由于灵敏度分析方法属于梯度型优化设计方法，其收敛速度较快，因此，Gao Z.Y. 等将该方法应用于微观组织模拟与优化中，考虑了复杂的再结晶机理如再结晶、部分再结晶和完全再结晶等，推导了关于设计变量的再结晶体积分数和动态再结晶晶粒大小的灵敏度分析方程。优化的目标为使单元晶粒尺寸与最终产品的平均晶粒尺寸的均方差描述的目标函数取最小值，将模具未充满及多余的飞边金属浪费作为优化过程的两个约束条件。优化设计变量分为状态参数（坯料的原始形状）和过程参数（模具速度）两类，以涡轮圆盘非等温锻造过程为例进行了微观组织模拟与优化，得到了能够获得良好微观组织结构的坯料原始形状和模具速度的优化解。

（1）优化设计问题

1）目标函数　微观组织的变化，尤其是晶粒尺寸的变化对微观组织性能具有十分重要的影响。理论上，获得均匀良好的晶粒尺寸及较大的再结晶体积分数是最理想的状态。但是，由于有限元模拟的离散化和金属塑性成形的非均匀变形行为，在每一个单元体内的晶粒尺寸都不相同。因此，所有的亚结构都会影响终锻件的性能。为此，引入每一单元体的平均晶粒度概念，优化建模时不仅要考虑再结晶晶粒尺寸和再结晶的程度，还要考虑未发生再结晶部分的晶粒尺寸及体积分数。因此，将目标函数定义为所有单元体的晶粒尺寸与整个终锻件的平均晶粒尺寸的均方值差，即

$$\min\left[\frac{\sum_{i=1}^{m}\int(D_i - D_{ave})^2 \mathrm{d}V_i}{\sum_{i=1}^{m}\int \mathrm{d}V_i} + \beta_i(D_{ave} - D_{des})^2\right] \tag{4.5-15}$$

式中，m 为坯料的单元总数；D_i 为每个单元体的实际晶粒尺寸；β_i 为权因子；D_{des}为所要求获得的晶粒尺寸；D_{ave}为最终锻件的平均晶粒尺寸，$D_{ave} = \dfrac{\sum_{i=1}^{m}\int D_i \mathrm{d}V_i}{\sum_{i=1}^{m}\int \mathrm{d}V_i}$。

2）约束条件 优化设计的目的是获得原始坯料形状和模具速度的优化参数以便得到终锻时高精度少飞边的并具有优良组织的锻件。因此，在建立上述优化设计目标函数的同时，添加如下约束条件：

$$\frac{V_{actual}-V_{desire}}{V_{desire}}<\psi \tag{4.5-16}$$

式中，V_{desire}和V_{actual}为终锻件的理想体积和实际体积；ψ为控制形状约束而给定的较小的常数。

3）设计变量 在锻造过程中，有两种不同的设计变量（状态变量和过程变量）在最终产品的机械和冶金性能方面起着重要作用。初始坯料的形状设计变量为状态设计变量，而模具速度变量为过程设计变量。

① 形状设计变量 微观组织的变化与初始坯料的形状有密切关系。初始坯料的形状会影响成形过程，如影响应变、应变速率及温度的分布，进而进一步影响微观组织的性能。

② 速度设计变量 许多材料对应变速率十分敏感。锻造过程中应变速率的分布在很大程度上影响着微观组织结构的特点，尤其是再结晶晶粒尺寸和再结晶体积分数。应变速率的控制常由对模具速度的控制得以实现。因此，模具速度的优化在成形过程中十分重要。在模具速度优化设计中，应尽可能用较少的参数来描述模具速度，典型的是上模速度给定为常数、指数或表格数据，而下模固定。

（2）灵敏度分析

基于锻造过程优化设计的灵敏度分析包括以下三个方面：微观组织的灵敏度分析、流动行为的灵敏度分析及温度的灵敏度分析。

1）微观组织行为及其灵敏度分析 为了优化锻造过程中的微观组织，尤其是计算微观组织参数和微观组织的灵敏度信息，需要得到微观组织行为和热机械变量的关系。用临界应变、再结晶体积分数和动态再结晶晶粒尺寸来描述动态再结晶过程，用平均晶粒尺寸描述复杂的再结晶机理。

① 临界应变 动态再结晶通常在应变值达到动态再结晶的临界应变后发生，当微观结构发生演变时，必须确定当前点的应变是否大于临界应变。在到达临界应变值之前，是不会发生再结晶的。

② 再结晶体积分数 再结晶体积分数是再结晶晶粒的体积与总体积之比。

③ 再结晶晶粒尺寸 再结晶晶粒尺寸是发生动态再结晶后的晶粒尺寸，是应变速率和温度的函数。

④ 平均晶粒度及其梯度 微观结构优化设计的目的是获得均匀良好的晶粒尺寸和大的再结晶体积分数。完全再结晶发生的温度范围很小，并且非等温锻造过程中大多数区域并不会发生完全再结晶。因此，在终锻中可能会发生以下几种情况：无再结晶，部分再结晶或完全再结晶。

2）流动行为的灵敏度分析 在非等温有限元数值模拟中，计算每一个设计变量的等效应变和等效应变速率的方法比等温过程模拟的公式要复杂许多。形状设计变量与模具速度变量的灵敏度分析也有一些不同。对于刚（黏）塑性材料，流动应力$\bar{\sigma}$（$\bar{\varepsilon}$，$\dot{\bar{\varepsilon}}$，T）是等效应变、等效应变速率和温度的函数。

对每一个设计变量进行求导，得到如下流动应力对设计变量的灵敏度方程：

$$\frac{\partial\bar{\sigma}}{\partial p_1}=\frac{\partial\bar{\sigma}}{\partial\bar{\varepsilon}}\frac{\partial\bar{\varepsilon}}{\partial p_1}+\frac{\partial\bar{\sigma}}{\partial\dot{\bar{\varepsilon}}}\frac{\partial\dot{\bar{\varepsilon}}}{\partial p_1}+\frac{\partial\bar{\sigma}}{\partial T}\frac{\partial T}{\partial p_1} \tag{4.5-17}$$

3）温度灵敏度分析 金属塑性成形过程的热平衡方程可表述如下：

$$C\dot{T}+K_cT=Q \tag{4.5-18}$$

式中各参数的含义见本篇第4章。

4）优化设计过程 给定初始坯料形状和模具速度，在初始速度场建立起来以后，就可进行热力耦合有限元分析。首先，在每一步模拟中进行坯料和模具的热分析。坯料的热分析以模具的温度不变为前提。同样，在模具的热分析过程中坯料的温度是不变的。然后进行流动分析。当热分析（坯料和模具）和流动分析都收敛时，再进行形状设计变量和模具速度设计变量的节点速度和节点温度的灵敏度计算。设计变量的节点坐标、等效应变和温度的灵敏度由于是依靠过程的参数，所以可用迭代求解进行修正。当终锻行程结束时，再结晶体积分数、动态再结晶晶粒尺寸、残余应力、平均晶粒度及其梯度便能够计算出来。由此，便可调用优化程序为初始坯料形状和模具速度提供设计变量的刷新值。

（3）计算实例

采用上述理论及方法对轴对称涡轮圆盘的锻造过程进行了微观组织模拟与优化。坯料的原始形状为$H=58.0$ mm、$R=70.0$ mm的圆柱体，坯料的原始晶粒尺寸为65 μm，摩擦系数为0.25，热力学参数为：坯料与模具的热导率分别为25 W/（m·K）和19 W/（m·K）、坯料与模具的体积热容为3.77 MJ/（m³·K）、热辐射和传热系数为8.5×10^{-9} W/（m²·K⁴）和20 W/（m²·K）、坯料与模具的初始温度分别为1 100℃和300℃、环境温度为20℃。当初始坯料加热到1 100℃时直接放入下模中，热锻结束后，立即放入水中冷却。

未对初始坯料和模具速度进行优化设计时，坯料不能完全充满模具型腔而且飞边体积很大，如图4.5-6所示。终锻后，动态再结晶体积分数、再结晶晶粒度、平均晶粒尺寸及残余应变等微观结构参数的分布如图4.5-7。从图中可以看出，晶粒尺寸方差值为87.30 μm²，平均晶粒度的方差值为45.01 μm²。终锻后再结晶体积分数在锻件中部区域较大，其原因是由于该处具有较大的应变值和较高的温度，产生了较多的位错，使再结晶行为进行得充分。在与模具接触的中心区，由于应变速率较小并且模具与坯料界面处的温度较低，再结晶体积分数并不是很大。而由于模具深腔处的等效应变值较小，材料的再结晶体积分数较小。由于缺少充分的再结晶晶粒长大时间，所以飞边处的再结晶晶粒尺寸也不是太大。所有具有较大再结晶体积分数的区域，其平均晶粒尺寸均较小，这是因为再结晶使得初始晶粒发生了细化。在再结晶行为充分的区域，其残余应变较小，并且大多数组织得到了一定程度的软化。所有这些都证明上述所建立的微观结构模拟方法是可行和有效的。

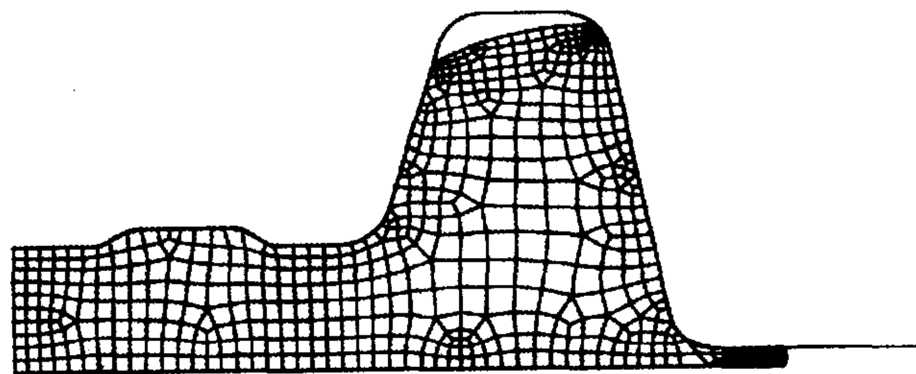


图4.5-6 未经优化设计的终锻件形状

在上述优化过程中，根据优化目标函数不同实施了两种优化方案。优化设计方案Ⅰ的目标函数为使等效应变的偏差最小以获得更均匀的终锻件等效应变分布，优化设计方案Ⅱ的目标函数为使晶粒尺寸与终锻件平均晶粒度的方差值最小。上述两种方案均采用相同的设计变量和约束条件。第一种优化设计迭代5次后收敛，终锻件的等效应变方差值从0.096减小到0.056，晶粒尺寸方差值从87.30 μm²减小到80.14 μm²，平均晶粒度的方差值则从45.01 μm²减小到42.03 μm²，其他的微观组织参数由于未经优化所以改变不大。

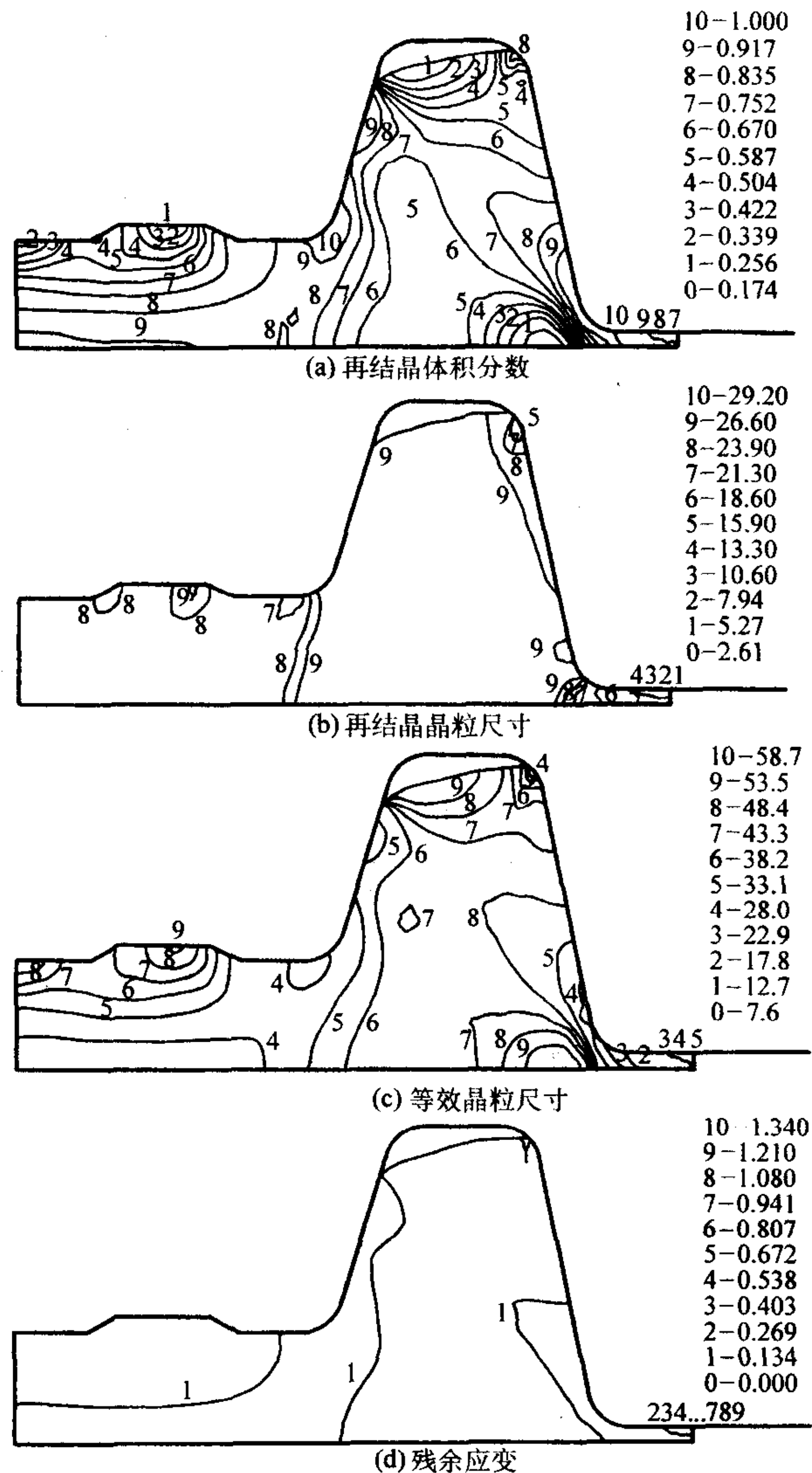


图 4.5-7　未经优化设计的终锻件微观结构

第二种优化设计方案中，优化设计经 7 次迭代后收敛。优化过程中模具速度逐渐增大，这对减少坯料与模具间界面上的热损失并产生足够的再结晶是十分有益的。优化后初始坯料的半径减小、高度增大，飞边的体积也减小了，型腔充填完好，如图 4.5-8 所示。型腔体积的形状约束函数值由 0.016 减小到 0.004 5，飞边体积的形状约束函数值由 0.013 减小到 0.006，经优化，终锻后的微观结构参数分布如图 4.5-9 所示。大多数区域的动态再结晶体积分数要比第一种方案相应区域大，而中心区域再结晶体积分数达到了 1.0，说明中心区域发生了完全再结晶。在模具与坯料界面处的中心区域，因温度较低，再结晶体积分数相对较小。平均晶粒尺寸方差值从 87.30 μm^2 减少到 58.31 μm^2，而晶粒度的方差值也由 45.01 μm^2 减少到 32.64 μm^2。这表明采用优化设计比未采用优化设计的终锻件平均晶粒尺寸更加均匀、良好。由于再结晶体积分数较大，优化设计中大多数区域中的残余应变比未经优化设计时要小得多。

将设计方案Ⅰ与设计方案Ⅱ的结果加以比较可以看出，不同的设计目标对可获得的优化结果影响很大。设计方案Ⅰ中，等效应变值变化很大，从 0.096 减小到 0.056，但晶粒度与平均晶粒度的方差值却变化不大，从 312.60 μm^2 减小到 224.86 μm^2。在设计方案Ⅱ中，晶粒尺寸与平均晶粒度的方差值变化很大，从 312.60 μm^2 减小到 65.28 μm^2，但等效应

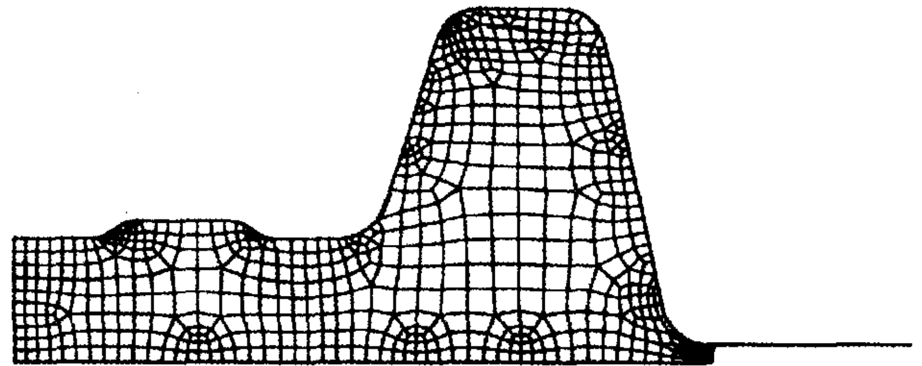


图 4.5-8　对坯料尺寸和模具速度优化设计后的终锻件形状

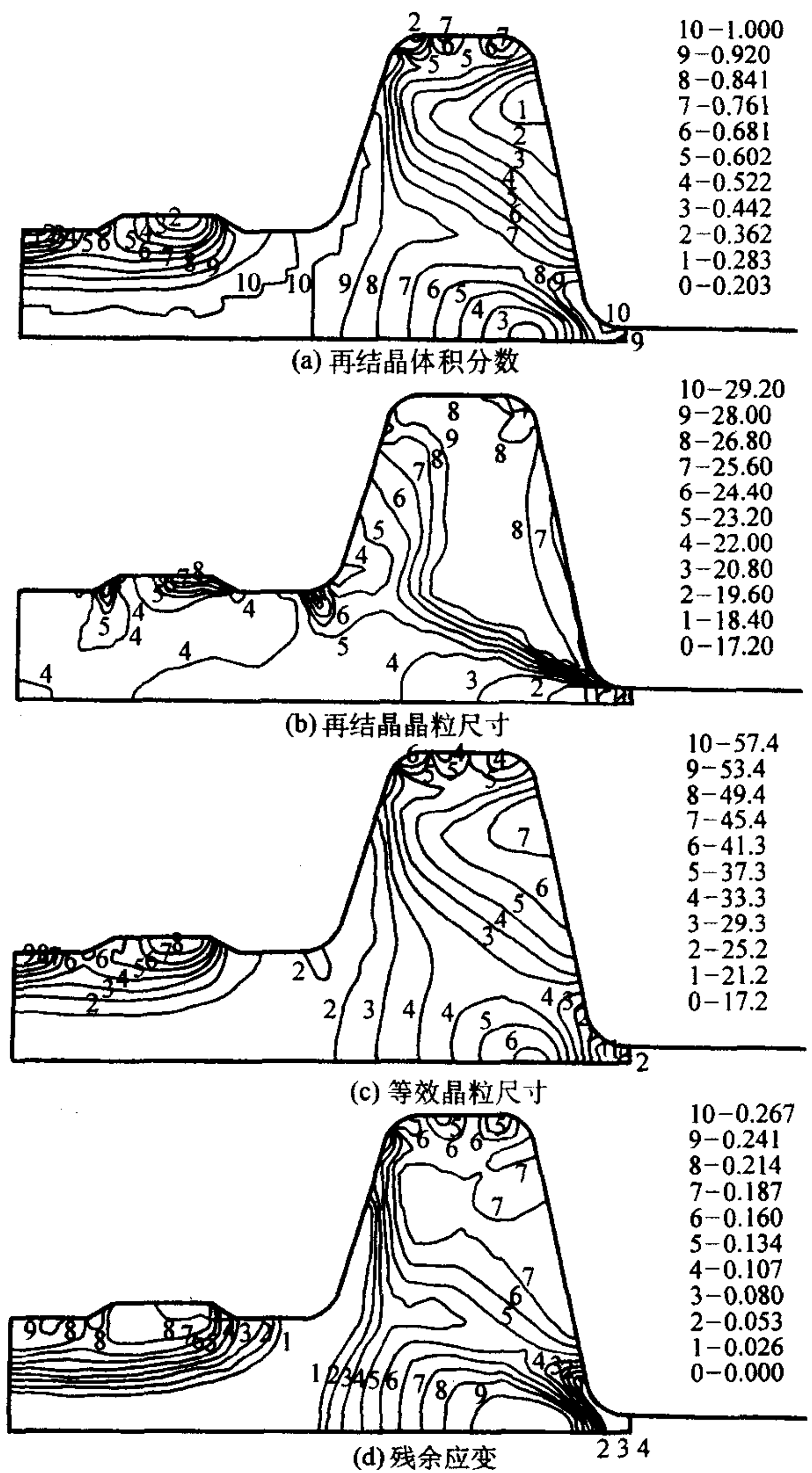


图 4.5-9　经优化设计后终锻件的微观结构

变值变化却不大，从 0.096 减小到 0.089。上述对比结果可以看出，微观组织的变化不仅受应变的影响，还同时受应变速率、温度及其他因素的影响。同时，从两种不同优化目标的优化结果的对比还可以看出，要想得到要求的微观组织，就必须将设计目标直接描述为预期的微观组织结构。

3.4　基于遗传算法的微观组织模拟与优化

采用遗传算法对典型的圆柱体镦粗工艺过程进行微观组织模拟与优化。以镦粗成形后晶粒度的细小化和均匀化为目标，目标函数式如式（4.5-15）。分别进行单一的摩擦因素和多种工艺因素的优化。

（1）摩擦因素优化

首先对摩擦因素进行优化。坯料材料为45[#]钢，ϕ4.0 mm×4.0 mm的圆柱体，压下量为40%，由于圆柱镦粗是轴对称变形，取四分之一进行模拟。假定初始晶粒度为280 μm，锻件始锻温度1 050℃，模具的初始预热温度为300℃，摩擦系数取值在0~1.0范围之间，其随机的初始值为0.5。

优化程序循环迭代5代之后收敛，此时，优化出的摩擦系数收敛为0.03。图4.5-10为目标函数在收敛前随迭代次数变化曲线。图4.5-11为摩擦系数在收敛前随迭代次数变化曲线。图4.5-12为摩擦系数采用优化前数值与优化后数值分别进行镦粗之后的晶粒度分布的比较。

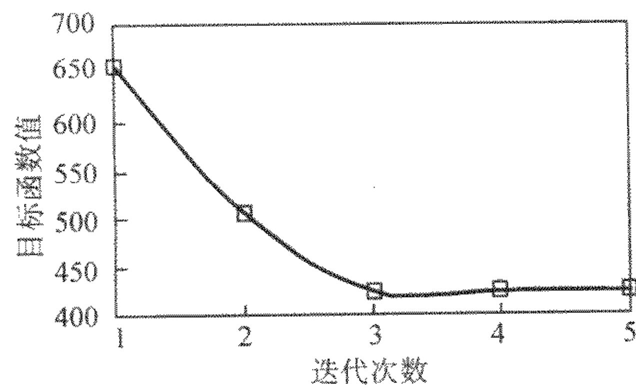


图4.5-10 目标函数随迭代过程的变化曲线

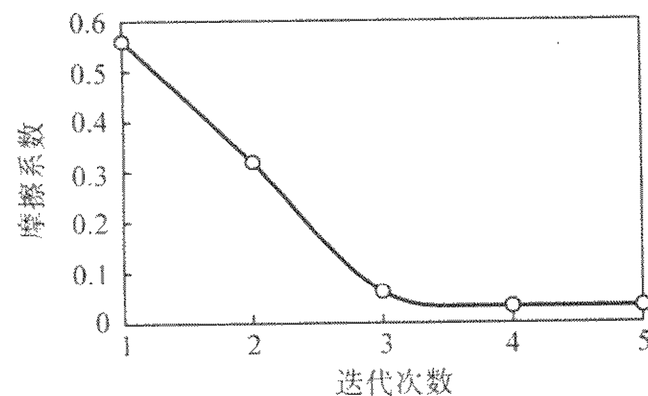
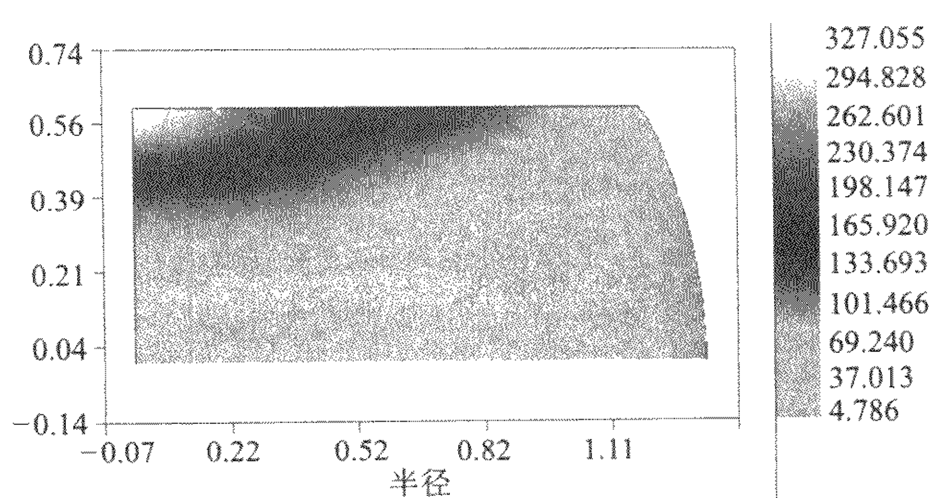


图4.5-11 摩擦系数随迭代过程的变化曲线

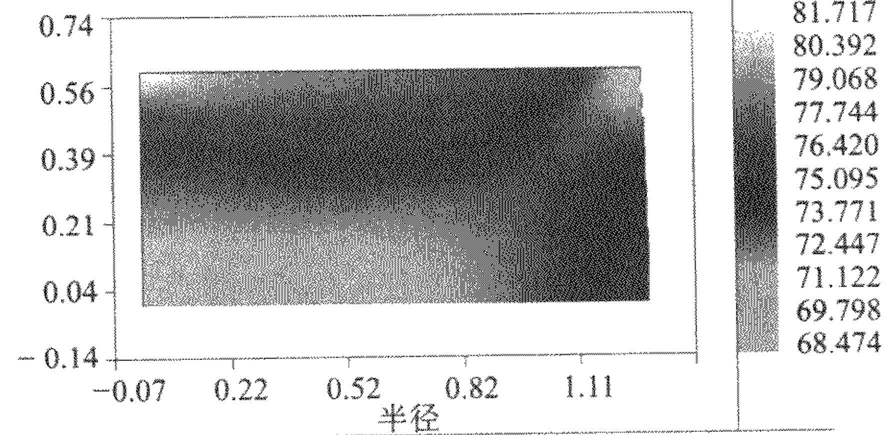
由图4.5-12可以看出，初始状态下，晶粒的大小由4.79~327.06 μm不等，整个锻件的平均晶粒度为87.31 μm。优化后，晶粒的大小集中在68.47~81.72 μm之间，此时，整个锻件的平均晶粒度为73.87 μm。由此可以看出，随着摩擦系数的减小，晶粒分布更均匀了，而且晶粒得到了细化，平均晶粒度下降了15.4%。

(2) 多因素优化

除了摩擦能够显著影响晶粒度分布的均匀性以外，模具压下速度和模具的预热温度等宏观工艺参数以及材料本身固有的特性如材料本身的热导率、比热容以及模具材料与工件



(a) 优化前m=0.5



(b) 优化后m=0.03

图4.5-12 摩擦系数优化前后晶粒度分布的比较

材料之间的传热系数等也会影响锻件晶粒度的分布。坯料的材料、形状尺寸和压下量等与单一摩擦因素优化分析相同。表4.5-1为考虑包括摩擦在内的多个参数随迭代次数增加的变化值。图4.5-13为多因素影响的微观组织优化目标函数值随迭代次数的增加而逐渐减小的变化曲线。图4.5-14为经过8次迭代后晶粒度分布的结果。

由图4.5-14可以看出，优化前，晶粒度的大小在34.18~341.24 μm之间变化，平均晶粒度为99 μm。优化后，晶粒度变化的范围在63.39~107.56 μm之间，而平均晶粒度为79 μm。平均晶粒度下降了20 μm。经过优化，微观组织更均匀更细小，与初始目标达成一致。从表4.5-1中各参数在优化过程的变化幅度和优化结果看，摩擦仍然是影响镦粗工艺微观组织晶粒度的主要因素，而模具预热温度等与热状态相关的工艺因素对最终目标函数值（即晶粒度）具有一定的影响，但是相对于摩擦的影响来看是比较小的。这种对比结果说明，热塑成形过程中的变形均匀性比热状态对变形后微观组织均匀性和细小化的影响要显著得多。

表4.5-1 各优化参数随迭代次数的变化值

迭代次数	1	2	3	4	5	6	7	8
摩擦系数	0.71	0.04	0.38	0.037	0.10	0.09	0.09	0.09
模具速度（相对值）	-2.28	-1.75	-1.75	-1.25	-1.65	-1.25	-1.25	-1.25
模具预热温度/℃	342	414	441	333	441	369	370	332
模具与工件间的传热系数/kW·m^{-2}·K^{-1}	2.16	3.44	1.18	1.18	1.17	2.86	2.88	2.87
模具的表面传热系数/kW·m^{-2}·K^{-1}	1.12	4.06	3.22	4.34	4.35	2.95	2.96	2.95
模具的比热容/W·m^{-1}·K^{-1}	49.26	39.04	29.04	59.06	39.26	29.26	29.23	24.35

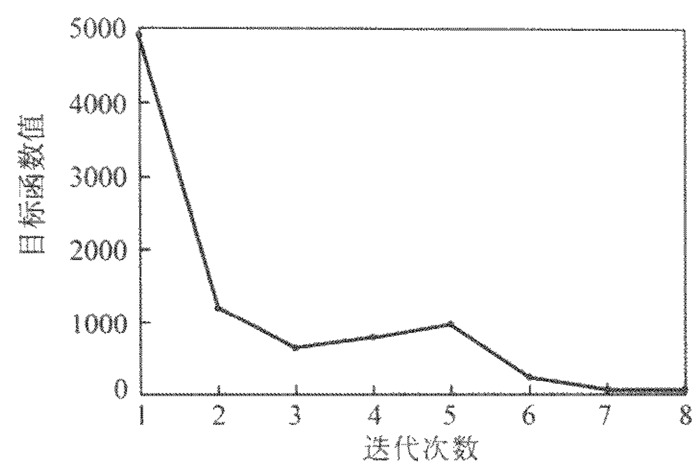
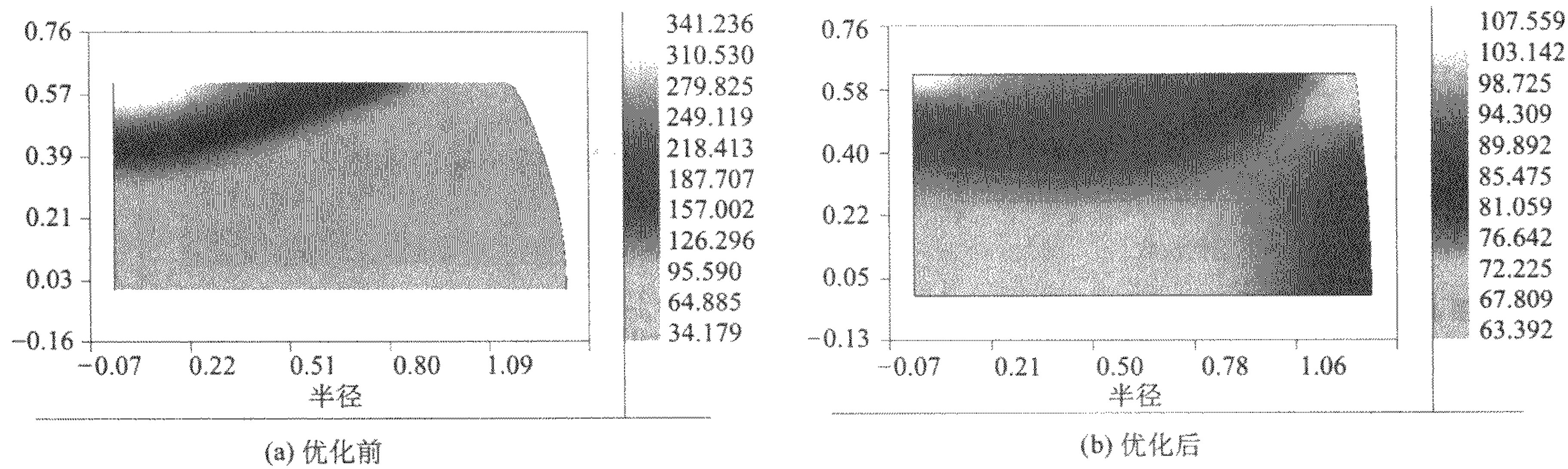


图 4.5-13　目标函数随迭代次数的变化曲线



(a) 优化前　　(b) 优化后

图 4.5-14　优化前后晶粒度分布比较

编写：王广春（山东大学）

第 6 章　稳态金属成形过程优化设计

挤压是一种少无切削加工的金属塑性加工工艺，它是根据金属塑性成形原理，利用装在压力机上的模具，在相当大的单位压力以及一定速度下，使金属在模具型腔内产生塑性变形，从而获得所需形状尺寸及其具有一定性能的零件。根据挤压金属流动方向与凸模运动方向之间的关系，挤压工艺可分为正挤压、反挤压、复合挤压和径向挤压等。正挤压工艺流程如图 4.6-1 所示。

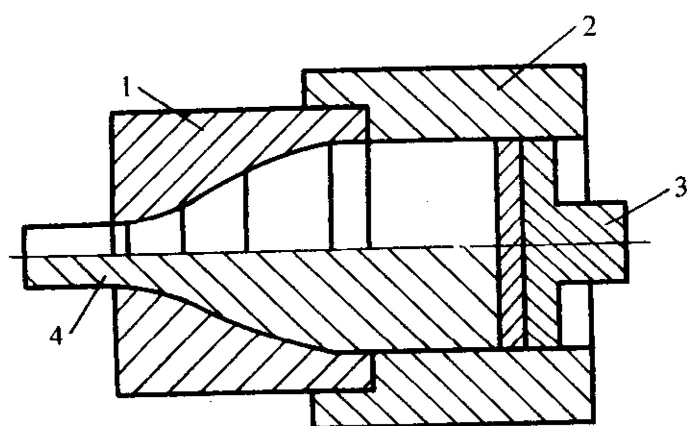


图 4.6-1　正挤压工艺示意图
1—凹模；2—凹模；3—冲头；4—工件

从本质上讲，金属在挤压过程中的变形属于非稳态流动。一般情况下，可将整个挤压过程分为三个阶段：挤压初期，塑性区由小逐渐变大，金属流动属于非稳态流动；挤压中期，塑性变形区达到最大且保持不变，由弹性区（或称刚性区）进入塑性区的金属量等于由塑性区流出工作带成为已成形区的金属量，变形体内的速度场变化很小，因此可将此时期的金属流动视为稳态流动；挤压后期，由于凹模腔内所剩金属量很少，凸、凹模与变形体逐渐接触，使摩擦对金属流动的影响增大，致使由弹性区流入塑性区的金属少于由塑性区流出工作带而成为已成形区的金属，塑性区变小，挤压缺陷有可能在此阶段产生，金属流动属于非稳态流动。属于稳态流动的挤压中期在整个挤压过程中所占的比例取决于坯料原始高径比和模具几何形状。在模具几何形状固定的情况下，坯料的原始高径比越大，稳态流动所占的比例越高。当坯料的原始高径比很大时，可忽略挤压前、后两期的非稳态流动，将整个挤压过程视为稳态流动。本章以稳态挤压为例，介绍稳态金属塑性成形过程的优化设计方法。

1　稳态成形过程优化设计问题

优化设计问题是求解在某些约束条件下目标函数的最优问题。所谓目标函数就是设计者对设计目标的数学表示，可根据不同的设计需要构造不同的目标函数。对于金属塑性成形问题，通常以节省材料和能量等作为设计目标，因而设计的目标函数通常是材料消耗和成形力或成形能量等。所谓约束就是设计者除了目标函数外还必须考虑的一些其他因素或条件，是对目标函数的限制。在金属成形问题中一般要考虑到成形设备的吨位、模具费用、材料允许的最大应力和应变率、生产率等因素，因而常把这些因素作为设计约束。

对于正挤压工艺设计而言，要达到的目标有许多，其中最基本的要求就是所选择的工艺应能成形出形状精确且无缺陷的工件。其他设计目标有变形均匀、载荷或能量最小、模具磨损最小、微观结构控制等。凹模轮廓的形状对控制金属流动、模具的应力分布、应变分布以及挤压力具有明显的影响。不当的模具轮廓设计会导致过多的材料内部剪切和模具出口附近金属流动方向的急剧变化，因而在挤压合金材料时，会导致纤维断裂和第二相粒子与主体材料分离等问题。因此，凹模轮廓设计就成为挤压工艺设计中最重要的方面之一。

另一方面，金属在挤压过程中是在三向不均匀压应力作用下产生塑性变形，在变形过程中金属流动剧烈，其挤压变形力一般要大于模锻力，消耗的总能量也多，因而经常以挤压能量最小作为挤压凹模轮廓优化设计的目标。同时，使挤压能量最小也有助于提高工件的变形均匀、防止变形缺陷产生和提高模具使用寿命。

对于一个给定的正挤压工艺，在冲头速度或生产率给定的情况下，使能量最小就等价于使成形力最小，因此可以选用成形力作为正挤压工艺优化的目标函数。在罚函数有限元法中，成形力可以表示为模具和工件接触表面 Γ_{cf} 上速度分布的函数，即

$$P = \int_{\Gamma_{cf}} \xi(\boldsymbol{u}_n - \bar{u}_n)\mathrm{d}\Gamma \tag{4.6-1}$$

式中，P 为成形力；$\boldsymbol{u}_n$ 为材料速度的法向分量，$\boldsymbol{u}_n = (N_{iL},\ V_L,\ \boldsymbol{n}_i)$；$\bar{u}_n$ 为冲头速度；ξ 为一个非常大的正常数；N_{iL} 为形状函数；V_L 为节点速度；$\boldsymbol{n}_i$ 为工件和模具接触面上的单位法向量。

一般用两种模型来描述正挤压凹模的轮廓形状。第一种模型是用由直线段组成的曲线来表示模具的轮廓形状，因而轴对称正挤压过程就可以视为一个多工序的凹模呈锥形的正挤压问题。图 4.6-2 中的凹模被划分成 n 条直线段，因此，凹模形状可表示为一个包含 $n+1$ 个直线段端点的设计模型。每一个端点的 x_1 和 x_2 坐标作为确定凹模轮廓形状的设计变量。由图 4.6-2 可知，设计变量的总数为 $2n+1$。若以三次样条曲线代替上述直线段，便产生了如图 4.6-3 所示的描述凹模轮廓形状的第二种模型。该模型的设计变量为每段三次样条曲线端点的 x_1 和 x_2 坐标。

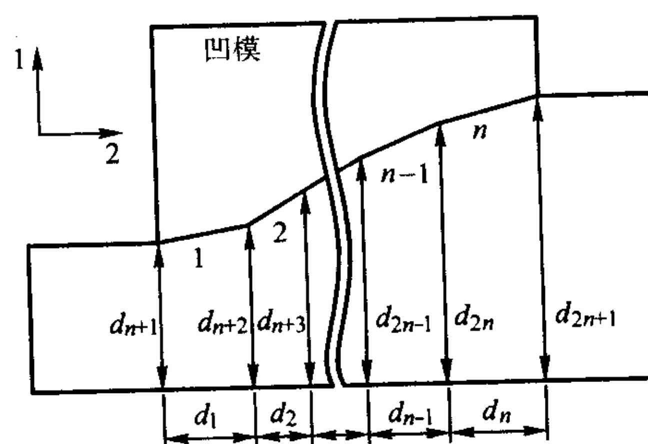


图 4.6-2　正挤压凹模轮廓设计的直线模型

上述正挤压凹模轮廓形状优化设计问题的数学表示如下：

$$\text{寻求设计变量 } d \text{ 使 } \psi_0 = P \text{ 最小} \tag{4.6-2}$$

且满足

$$\psi_1 = d_{2n+1} - r_i = 0 \tag{4.6-3}$$

$$\psi_2 = d_{n+1} - r_o = 0 \tag{4.6-4}$$

$$\psi_i = \psi_i\ (s,\ d)\ \leqslant 0,\ i = 3,\ \cdots,\ \beta \tag{4.6-5}$$

以及有限元方程：

$$h(s, d) = 0 \tag{4.6-6}$$

式中，r_i，r_o为挤压件横截面的初始半径和结束半径；d为设计变量；s为状态变量。

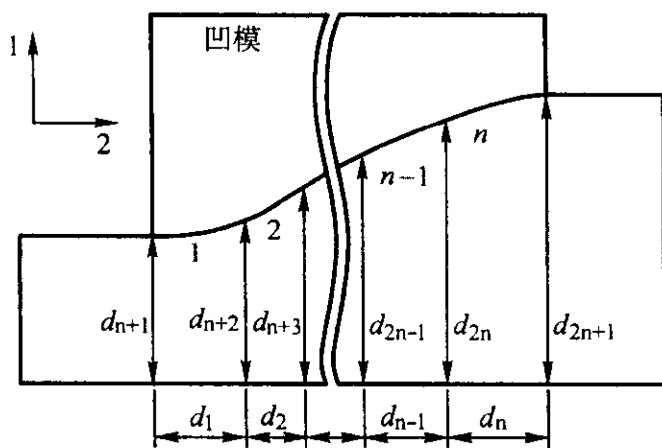


图4.6-3　正挤压凹模轮廓设计的三次样条曲线模型

由式（4.6-3）和式（4.6-4）给出的约束可保证获得一个给定面积上的断面缩减率。式（4.6-5）表示各种不同的设计约束，这些约束包括工艺设备约束、不均匀变形程度约束、静水压力的非负值大小约束等等。

在机械设计中，设计系统的力学特性可以用一系列称为状态变量的变量来描述，与这些变量相关的基本方程称为状态方程。此外，还有一系列描述系统的变量，这些变量称为设计变量。设计变量必须由设计者自己选定。

在应用有限元模拟技术进行稳态金属成形分析和设计时，其设计变量和状态变量包括如下。

1）设计变量 d　包括模具几何形状、工艺过程条件等。

2）状态变量 s　包括有限元节点速度向量、节点力向量等等。

如果设计变量确定，状态变量即可通过求解有限元方程而确定。

2　设计灵敏度分析

优化设计问题的迭代算法一般需要计算目标函数和约束函数对设计变量的灵敏度，即目标函数和约束函数相对于设计变量的微分。在数值计算过程中，大部分的时间用来寻找设计灵敏度的信息。

为了确定设计灵敏度，考虑状态变量和设计变量的微小变化，认为$\delta\psi_i$是$\psi_i(s,d), i=0,1,2,\cdots,\beta$增量的线性近似：

$$\delta\psi_i[s^0,d^0]=\frac{\partial\psi_i}{\partial s}[s^0,d^0]\delta s+\frac{\partial\psi_i}{\partial d}[s^0,d^0]\delta d \tag{4.6-7}$$

式中，d^0为优化设计迭代中设计变量的当前值；s^0为与d^0相关的状态变量。

考虑到$h(s^0,d^0)=0$，δs和δd满足方程$h(s^0+\delta s,d^0+\delta d)=0$，可得下列方程：

$$J\delta s+\frac{\partial h}{\partial d}\delta d=0 \tag{4.6-8}$$

式中雅可比矩阵$\boldsymbol{J}=\frac{\partial h}{\partial s}[s^0,d^0]$是采用Newton－Raphson迭代法求解有限元方程的刚度矩阵。对式（4.6-8）求解可得：

$$\delta s=-\boldsymbol{J}^{-1}\frac{\partial h}{\partial d}\delta d \tag{4.6-9}$$

把式（4.6-9）代入式（4.6-7），并用δd来表示所有的变量，可以得到

$$\delta\psi_i=\left[-\frac{\partial\psi_i}{\partial s}\boldsymbol{J}^{-1}\frac{\partial h}{\partial d}+\frac{\partial\psi_i}{\partial d}\right]\delta d \tag{4.6-10}$$

为了避免直接对雅可比矩阵$\boldsymbol{J}$求逆，定义相关变量λ^i，$i=0，1，2，\cdots，\beta$，作为以下相关方程的解：

$$\boldsymbol{J}^{\mathrm{T}}\lambda^i=\frac{\partial\psi_i}{\partial s} \tag{4.6-11}$$

从式（4.6-11）和式（4.6-10），可以得到：

$$\delta\psi_i=l^{i^{\mathrm{T}}}\delta d \tag{4.6-12}$$

式中l^i（$i=0，1，2，\cdots，\beta$）为点（s^0，d^0）处的设计灵敏度向量，且有

$$l^i=\frac{\partial\psi_i^{\mathrm{T}}}{\partial d}-\frac{\partial h^{\mathrm{T}}}{\partial d}\lambda^i \tag{4.6-13}$$

通常，函数ψ_0，ψ和h相对于设计变量d的微分的显式表达式很难求解，因此，采用有限差分法来计算其相对于设计变量的偏微分：

$$\frac{\partial f_i}{\partial d_j}=\frac{f_i(s^0,d^0+\Delta d_je^j)-f_i(s^0,d^0)}{\Delta d_j} \tag{4.6-14}$$

式中，Δd_j为步长向量Δd的第j个分量，即d_j的微小摄动量；e^j为单位向量，在第j个位置为1，其余地方为0。

差分方法比较通用和简单，但必须选择合适的步长Δd_j以保证设计灵敏度的精度。若选择不当，会增加数值计算的难度。

3　二维挤压模具形状优化设计

如果不考虑挤压过程中温度变化的影响，即认为挤压过程为等温过程。采用式（4.6-1）定义的目标函数，对理想刚塑性材料、无摩擦的平面应变等温正挤压工艺的凹模形状轮廓进行优化设计。

为了测定算法的精度，将优化得到的凹模轮廓形状与Richmond给出的理论凹模轮廓形状进行比较。优化中采用的正挤压凹模轮廓形状由5段三次样条曲线来描述。

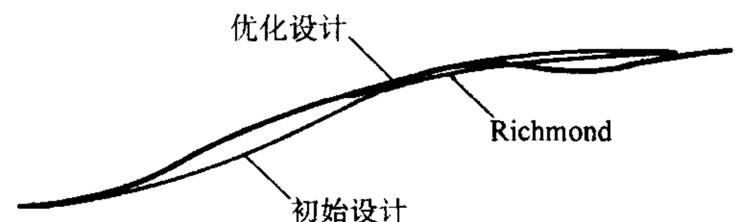


图4.6-4　优化凹模形状与理想凹模形状的比较

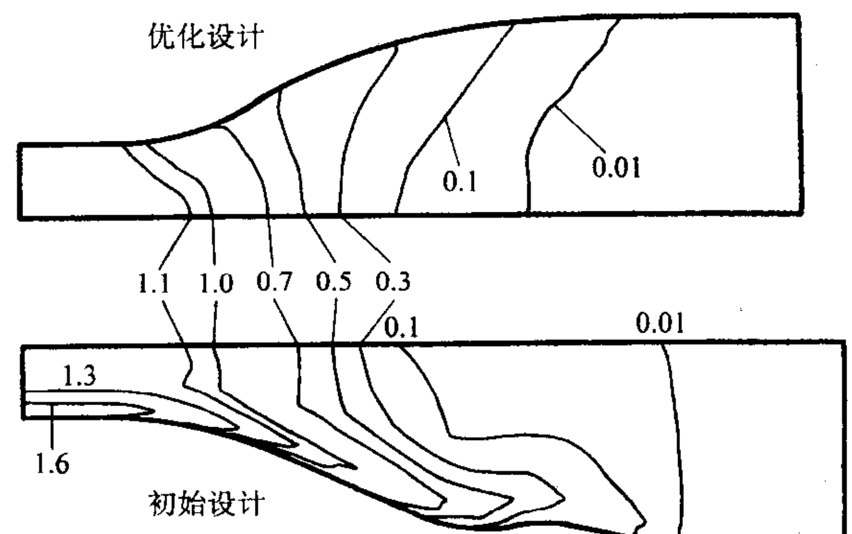


图4.6-5　初始设计凹模和优化设计凹模中的等效应变分布

图4.6-4所示为某一断面缩减率为$R_A=(F_0-F_1)/F_0\times100\%$，其中$F_0$、$F_1$分别为挤压变形前、后毛坯的横截面面积，优化所得的凹模轮廓形状和理论凹模轮廓形状的比较。由图4.6-4可以看出计算结果与理论结果吻合很好。由图4.6-5中的等效应变分布也可以看出，对于初始设计的凹模轮廓形状，材料在凹模表面经历了一个相对较大的变形，导致在模具出口处沿径厚向出现较大的应变梯度。与此相反，采用优化的凹模轮廓形状，凹模出口处的应变分布较均匀，其分布值范围在1.150～1.159之内，非常接近于理论值1.154 7。

下面介绍不同工艺条件和材料特性对模具轮廓形状的影

响，其中工艺条件和材料如下。

1）工艺　轴对称正挤压。

2）材料　理想刚塑性材料，$\bar{\sigma}=100$ MPa，应变率硬化材料$\bar{\sigma}=c\dot{\bar{\varepsilon}}^n$，其中，$c=65.95$ MPa，应变率硬化指数 $n=0.08$，0.15，0.195。

3）入口半径　$r_i=30$ mm。

4）断面缩减率（%）　$R_A=55.6$，33.3，16.7。

5）摩擦系数　$\mu=0.05$，0.1，0.2。

6）冲头速度　$\bar{u}_n=-100$ mm/s。

对于理想刚塑性材料，优化得到的凹模轮廓形状如图 4.6-6 所示。在优化过程中，采用由八个线段组成的直线模型来描述凹模的轮廓形状。由图 4.6-6 可以看出，不管断面缩减率如何，随库仑摩擦系数的变化，凹模表面轮廓沿轴向的长度变化明显，这表明摩擦是影响优化凹模轮廓的一个重要因素。同时，可以看出在给定库仑摩擦系数的情况下，凹模轮廓长度随断面缩减率的增大而变长。由图 4.6-7 和图 4.6-8 可以看出，应变率硬化材料（$n=0.195$）和理想刚塑性材料的优化凹模轮廓形状有明显的差别。凹模轮廓的长度随应变率硬化指数 n 的增大而变长，这意味着材料的应变率硬化特性越高，用于完成剪切变形的断裂能量越多。

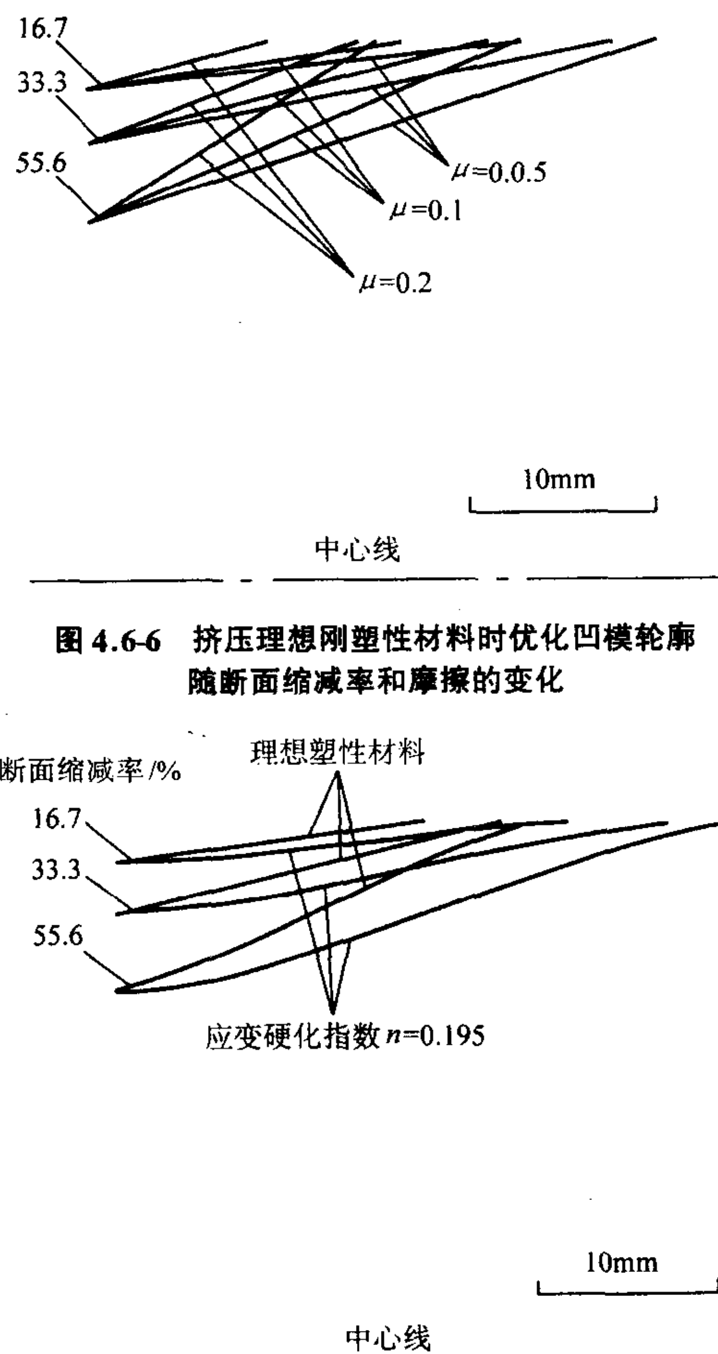


图 4.6-6　挤压理想刚塑性材料时优化凹模轮廓随断面缩减率和摩擦的变化

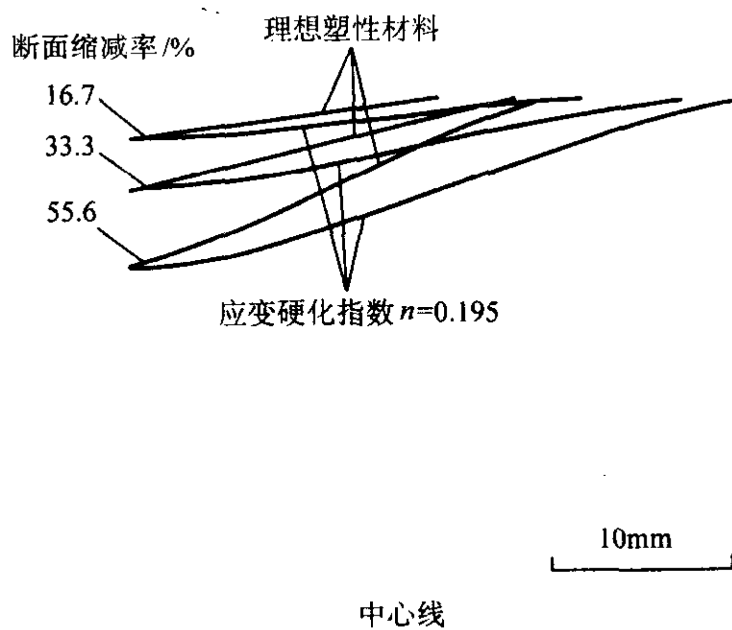


图 4.6-7　挤压理想塑性材料和应变率硬化材料的优化凹模轮廓形状（$\mu=0.1$）

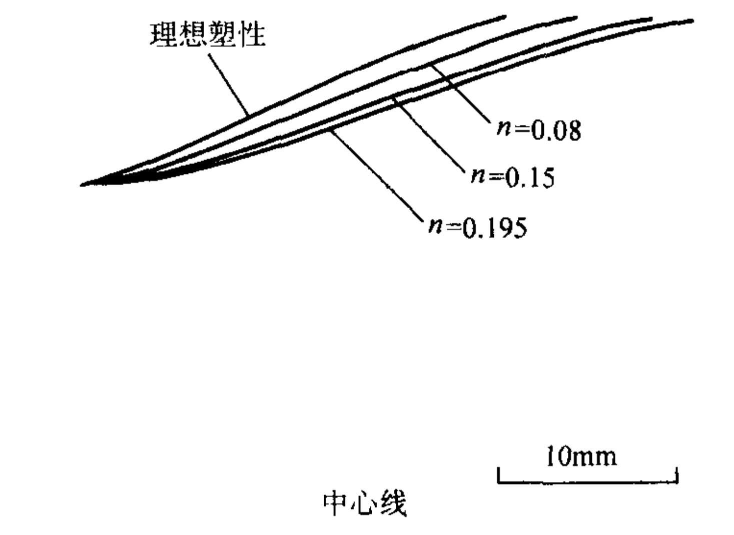


图 4.6-8　优化凹模轮廓随应变率硬化指数的变化（$R_A=55.6\%$）

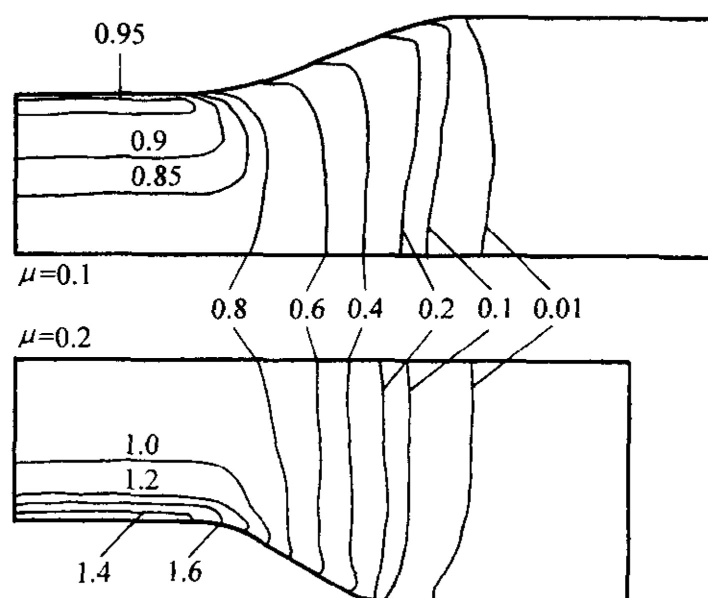


图 4.6-9　优化凹模形状对应的挤压件的等效应变分布（$n=0.195$，$R_A=55.6\%$）

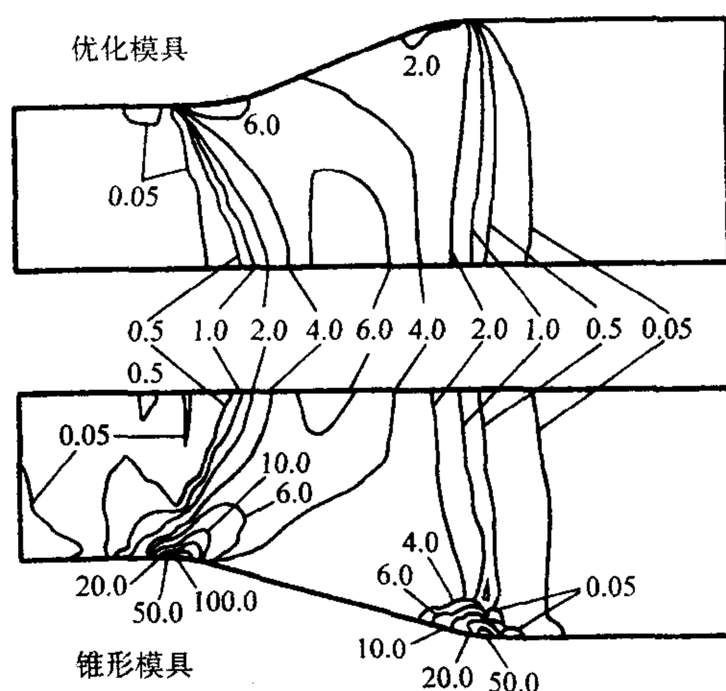


图 4.6-10　锥形模具与优化模具对应的挤压件内等效应变率分布的比较

由图 4.6-9 给出的等效应变分布图可以看出，摩擦对挤压件的应变分布均匀性的影响很大。等效应变值在模具表面附近最大，而在挤压轴附近最小。值得注意的是，当 $\mu=0.2$ 时，等效应变的最大值不是出现在表面那一层，而是出现在材料的内部。由图 4.6-10 所示的等效应变率的分布可以推断出更详细的金属流动规律：采用同样长度的模具进行挤压，应用锥形模具时应变率的峰值出现在模具附近，而应用优化后的模具轮廓时，应变率峰值则出现在变形区的中心区域，这对获得均匀的最终应变分布十分有利。挤压中心裂纹的出现与其中心区域的拉应力有关，由图 4.6-11 可以看出，优化得到的平均应力在中心区域均为负值（压应力），因而可以避免中心裂纹的产生，提高工件的质量。图 4.6-12 所示为优化模具和锥形模具在两种摩擦条件下的模具压力分布。优化模具表面的分布压力较小，这表明以能量最小准则进行优化设计有利于减少模具表面的磨损。

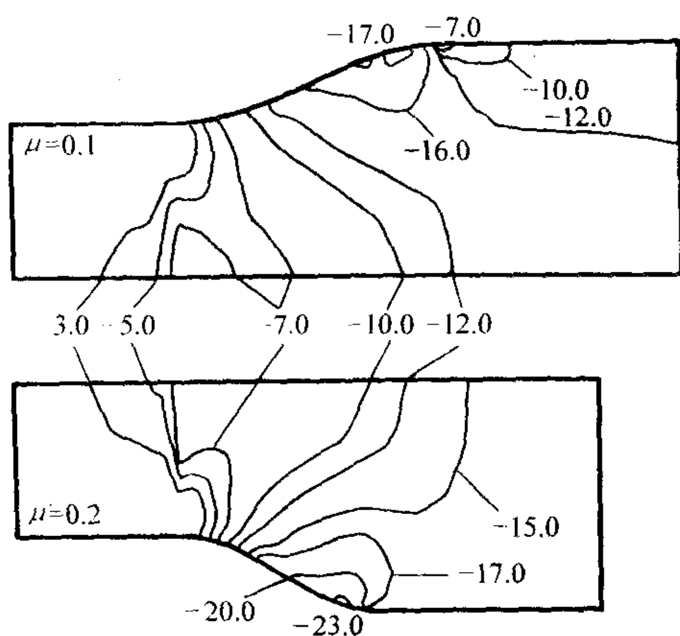


图 4.6-11　优化模具对应的挤压件内平均应力分布

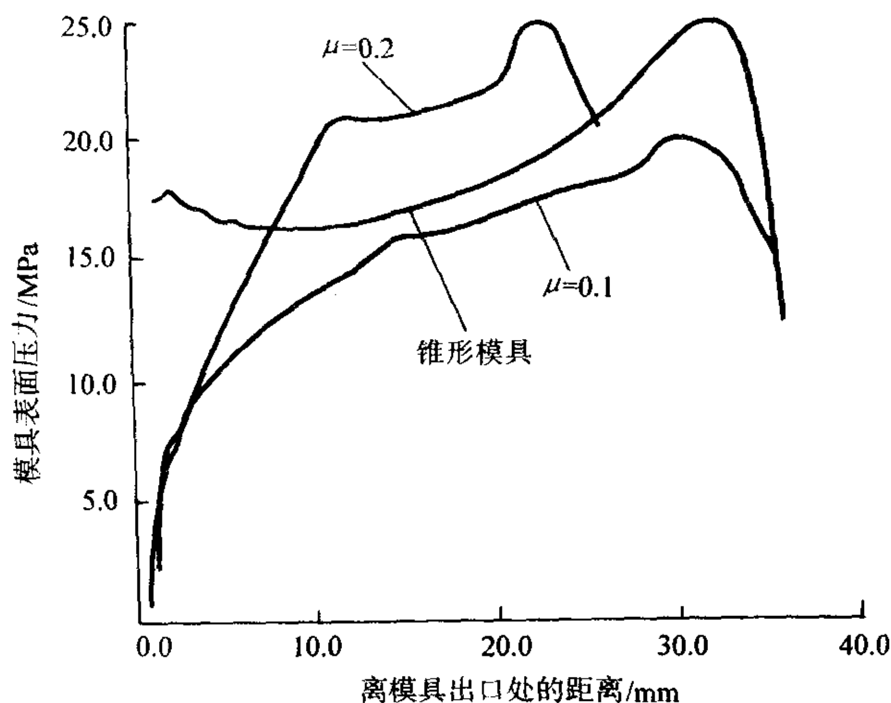


图 4.6-12　锥形模具和优化模具的压力分布比较

4　三维挤压模具形状优化设计

4.1　设计模型

如图 4.6-13 所示，假设模具几何形状至少有两个对称面，并且令 z 轴与挤压件的中心线重合，则模具表面上任意点的位置可用向量表示为：

$$r = R(\theta,z)\cos\theta_i + R(\theta,z)\sin\theta_j + zk \tag{4.6-15}$$

式中$R(\theta,z)$为沿中心线任意位置上模具横截面的轮廓形状，所以$R(\theta,0)$就表示挤压件横截面的轮廓形状，$R(\theta,L)$表示模具挤压入口处横截面的轮廓形状。

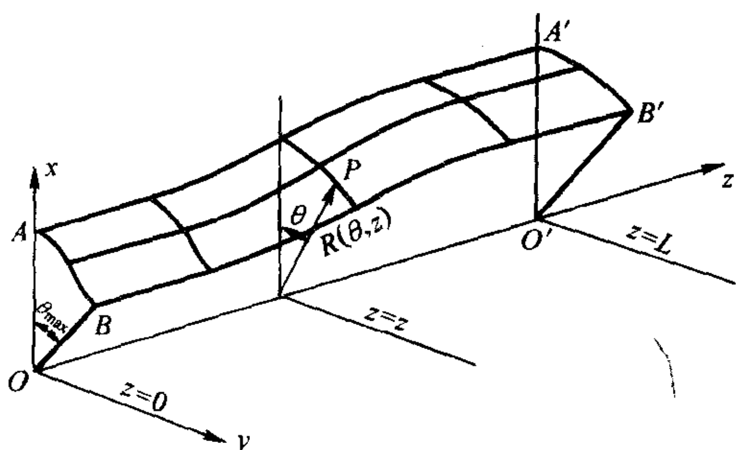


图 4.6-13　至少有两个对称面的挤压工艺的模具形状示意图

在这种情况下，$R(\theta,z)$可表示如下：

$$R(\theta,z) = R(\theta,0)[1 - Z(\theta,z)] + R(\theta,L)Z(\theta,z) \tag{4.6-16}$$

式中，复合函数$Z(\theta,z)$满足：

$$Z(\theta,z) = \begin{cases} 0 & z=0 \\ 1 & z=L \end{cases} \tag{4.6-17}$$

为了使凹模曲线在出口和入口处能够光滑过渡，函数$Z(\theta,z)$还应满足以下附加约束：

$$\frac{\partial Z}{\partial z} = \begin{cases} 0 & z=0 \\ 0 & z=L \end{cases} \tag{4.6-18}$$

很明显，符合上述条件的函数$Z(\theta,z)$是无限的，每一个函数都会通过式（4.6-16）而获得特定的模具形状。此处只考虑下列形式的函数：

$$Z(\theta,z) = Z_1(z)f_1(\theta) + Z_2(z)f_2(\theta) \tag{4.6-19}$$

式中Z_1(和 Z_2)是满足$Z_1(0)=0$，$Z_1(L)=1$ 和 $Z'_1(0)=Z'_1(L)=0$ 和以下约束的函数：

$$f_1(\theta) = [1+\cos(\pi\theta/\theta_{max})]/2 \tag{4.6-20}$$

$$f_2(\theta) = [1-\cos(\pi\theta/\theta_{max})]/2 \tag{4.6-21}$$

式中 θ_{max}为两相邻对称平面的夹角，如图 4.6-13 所示。

所选择的$f_1(\theta)$和$f_2(\theta)$的表达式，使$Z(0,z)=Z_1(z)$和$[Z(\theta_{max},z)]=Z_2(z)$成立。

用有三个子区间的三次样条曲线来重新表示函数$Z_1(z)$[和 $Z_2(z)$]。三次样条曲线的定义如下：

$$F(z) = \frac{g_i(z_{i+1}-z)^3}{6h_i} + \frac{g_{i+1}(z-z_i)^3}{6h_i} + \left[\frac{F_i}{h_i} - \frac{g_i h_i}{6}\right](z_{i+1}-z) + \left[\frac{F_{i+1}}{h_i} - \frac{g_{i+1}h_i}{6}\right](z-z_i) \tag{4.6-22}$$

在子区间 $z_i < z < z_{i+1}$ 中，满足 $h_i = z_{i+1} - z_i$，$F_i = F(z_i)$ 和 $g_i = F''(z_i)$。g_i 的值可以用函数 F 在数据点(z_i, F_i)，$i=1,2,3,4$ 的 C^1 连续来确定。然后，用这些数据点的坐标定义函数 Z_1（和 Z_2）。因此，它们也可视为设计变量。

忽略那些可以预先确定的设计变量，即假定表示函数 Z_1（和 Z_2）的两条三次样条曲线的子区间的长度完全相等，这样，该设计模型中独立设计变量的数目就可以降低为 7 个，分别用 $d_1 \sim d_7$ 表示，如图 4.6-14 所示。

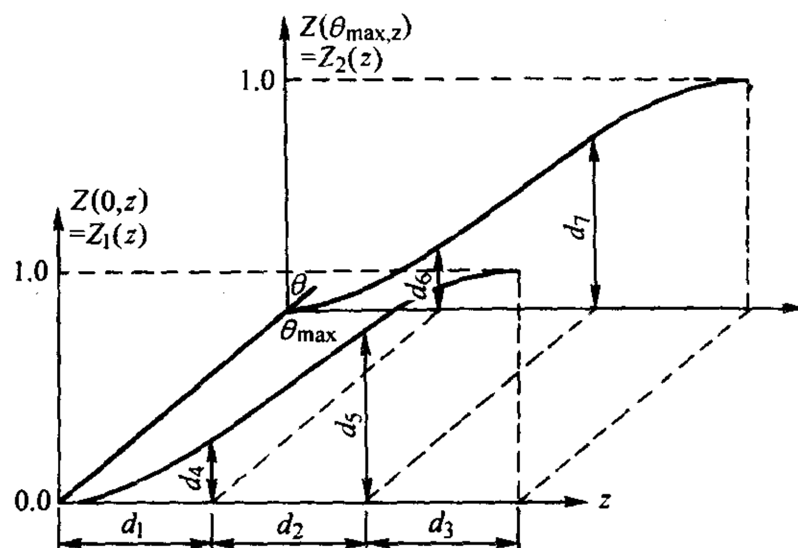


图 4.6-14　三维模具形状优化设计变量的选择

如果只有一个对称面或根本没有对称面，那么模具几何形状就没有如图 4.6-13 所示的一条中心线。在这种情况下，在垂直于 z 轴的平面上相对于模具出口横截面$R(\theta,0)$的横截面$R(\theta,z)$的位置未知，因此，必须把确定$R(\theta,z)$位置的函数也包括在式（4.6-15）中。如果存在一个与 z 轴一致的对称面，式（4.6-15）可以修改为：

$$r = \{R(\theta,z)\cos\theta + p(z)\}i + R(\theta,z)\sin\theta_j + zk \tag{4.6-23}$$

式中，函数$p(z)$为沿 x 方向模具横截面$R(\theta,z)$的变换函数，如图 4.6-15 所示。

此处函数$p(z)$可用一个有三个子区间的三次样条曲线来近似表示。由于$p(0)=0$，所以函数$p(z)$的引入可给出其中的六个设计变量。

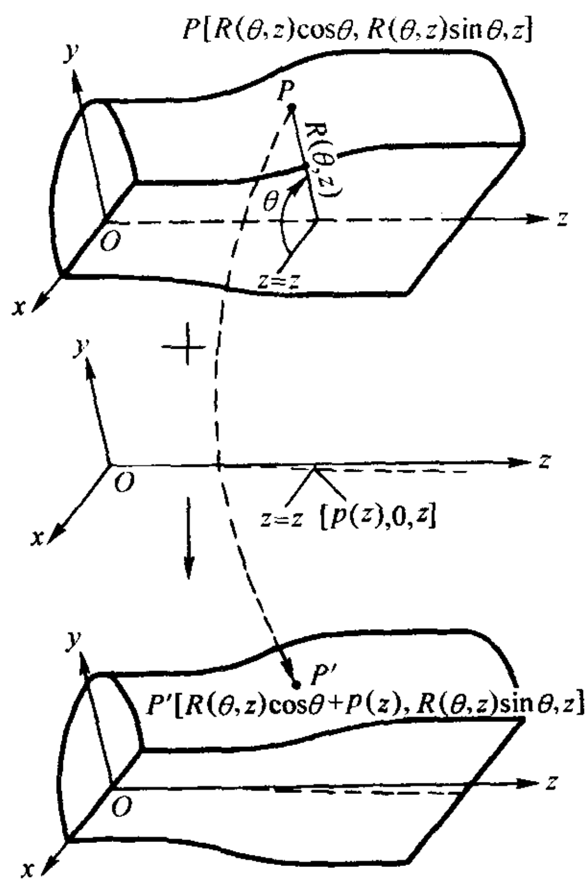


图 4.6-15　只有一个对称面的模具形状示意图

与二维挤压问题相同，三维挤压问题的优化设计也同样需要计算目标函数和约束函数的设计灵敏度，即目标函数和约束函数相对于设计变量的微分。由于三维问题设计灵敏度的计算与二维问题基本类似，在此不再一一叙述，请参见前面相关章节，下面仅给出优化设计实例。

4.2　优化实例

本节介绍多边形、T形和半圆半椭圆形等模具截面形状的挤压工艺优化设计。表 4.6-1 给出了各种形状的设计模型。假定坯料为圆形棒料，以挤压力为优化设计目标对模具形状进行优化，其工艺过程条件如下：

$\bar{\sigma} = 65.95\,\bar{\varepsilon}^{0.195}$ MPa;

冲头速度 = 100 mm/s;

模具和材料之间的接触摩擦系数 $\mu = 0.1$;

毛坯直径为 40 mm。

图 4.6-16 给出了各种模具形状时挤压件的有限元模型，其中采用的单元为六面体单元。进行有限元模拟和优化设计即可对各种多边形截面的挤压模具形状进行优化。三维模具的优化形状如图 4.6-17 所示。图 4.6-18 表示多边形截面挤压件内部的等效应变分布，由图 4.6-18 可以看出，优化得到

表 4.6-1　不同横截面的设计模型

形状	$R(\theta,0)$/mm
多边形	（图：y, x, o, $R(\theta,0)$, θ_{max}, θ） $R(\theta,0) = \dfrac{-c}{a\cos\theta}$ $a = 1$ $b = 0$ $c = -\dfrac{1}{2\tan\theta_{max}}$ $\theta_{max} = \dfrac{\pi}{n}$

形状	边长 l	边数 n	拐角半径 r
三角形	30	3	3
四边形	25	4	2.5
五边形	15	5	0
六边形	16	6	2.5
八边形	11.48	8	4.8

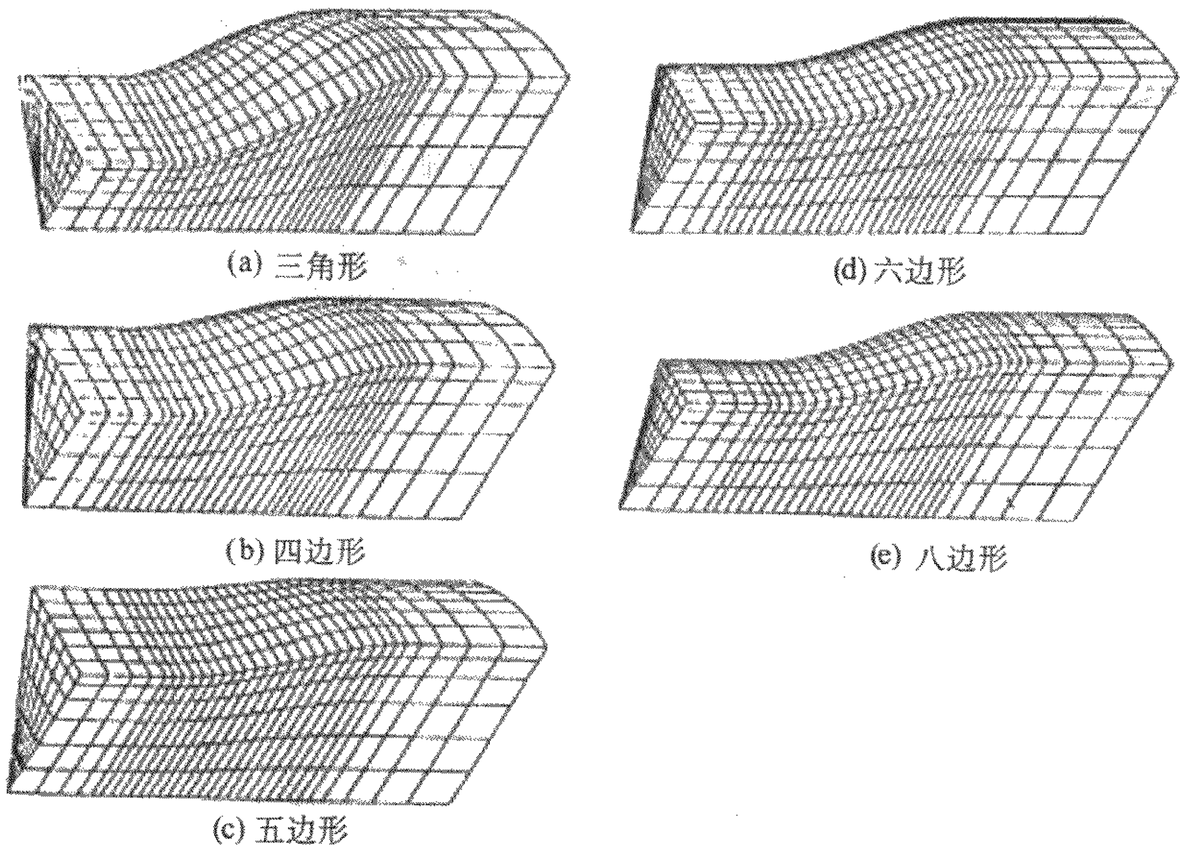
形状	$R(\theta,0)$/mm
T形截面	（图：y, x, o, 1, 2, 3, 4, 5, θ, $R(\theta,0)$） $R(\theta,0) = \dfrac{-c}{a\cos\theta + b\sin\theta}$

形状	线段 1 $0 < \theta < \theta_1$	线段 2 $\theta_1 < \theta < \theta_2$	线段 3 $\theta_2 < \theta < \theta_3$	线段 4 $\theta_3 < \theta < \theta_4$	线段 5 $\theta_4 < \theta < \theta_{max} = \pi$
长 T形($R=3$)	$a=1$ $b=0$ $c=-25$ $\theta_1=11.3°$ $l=15$	$a=0$ $b=1$ $c=-5$ $\theta_2=45°$ $l=20$	$a=1$ $b=0$ $c=-5$ $\theta_3=71.6°$ $l=10$	$a=0$ $b=1$ $c=-15$ $\theta_1=108.4°$ $l=10$	$a=1$ $b=0$ $c=5$ $l=15$
短 T形($R=3$)	$c=-15$ $\theta_1=18.4°$ 其余同上	$l=10$ 其余同上	同上	同上	同上

续表 4.6-1

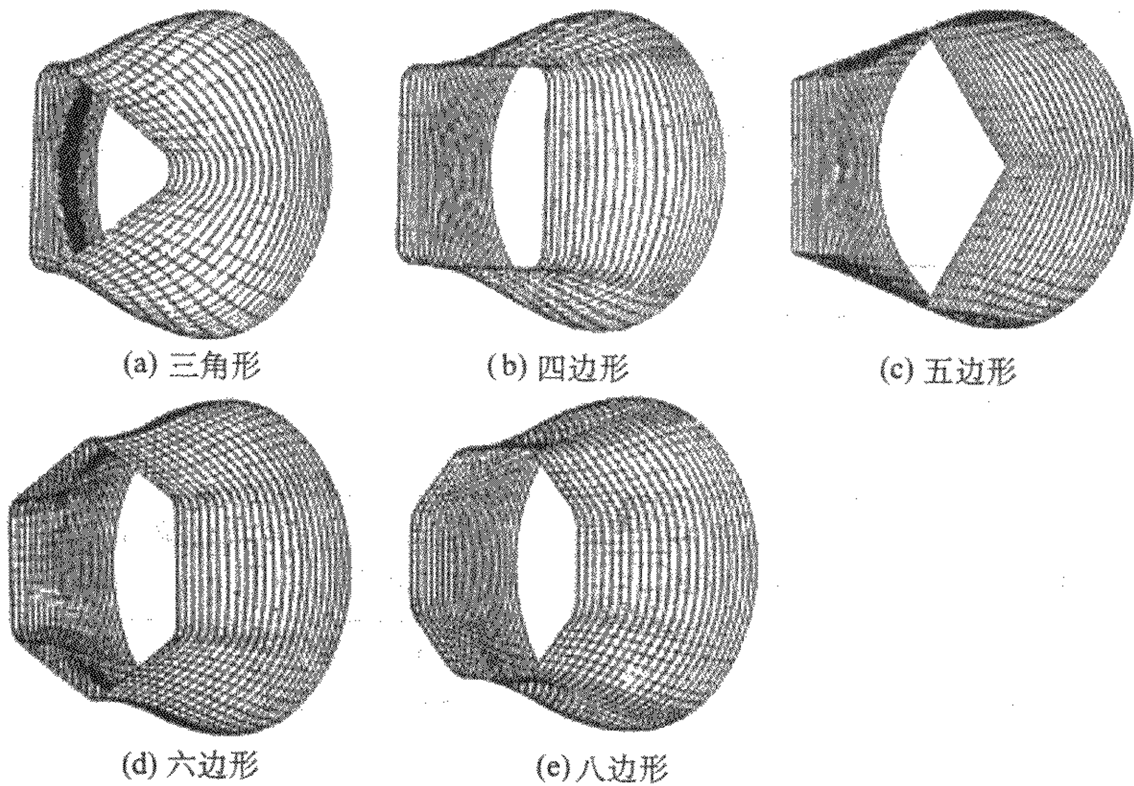
形状	$R(\theta,0)$/mm	
半圆半椭圆形截面	(图：y、$R(\theta,0)$、o、θ、x)	$R(\theta,0)=15, 0<\theta<\pi/2$ $R(\theta,0)=\dfrac{ab}{\sqrt{a^2\sin^2\theta+b^2\cos^2\theta}}, \dfrac{\pi}{2}<\theta<\theta_{\max}=\pi, a=15, b=10$

的挤压件内最终应变分布的均匀性得以明显提高，尤其是在与模具边界平行的方向上。采用上述方法，也可对T型截面的挤压工艺进行优化设计。对于该种情况，模具几何形状只有一个对称面。这时，就必须采用式（4.6-23）给出的设计模型，这种模型可以允许模具横截面$R(\theta,z)$沿x轴相对于出口横截面有相对位置的变化。对由函数$p(z)$以及函数$Z_1(z)$和$Z_2(z)$定义的设计变量进行优化，优化得到的模具形状如图4.6-20所示。由图4.6-21所示的挤压件的等效应变分布图可以看出，等效应变的最大值出现在T形截面的凹表面处，应变的最小值出现在T形截面的底部。



(a) 三角形　(b) 四边形　(c) 五边形　(d) 六边形　(e) 八边形

图 4.6-16　挤压模拟中的有限元网格



(a) 三角形　(b) 四边形　(c) 五边形　(d) 六边形　(e) 八边形

图 4.6-17　优化的挤压模具形状

为了验证优化结果，将半圆半椭圆横截面的挤压工艺的优化结果与无优化的结果进行比较，优化结果如图4.6-22所示。无优化结果可通过分别用$p(z)=0$和$-p(z)$代替优化设计中的$p(z)$来获得，如图4.6-23所示。图4.6-24所示的

等效应变比较表明优化设计对于变形均匀性的提高具有明显的作用。由图 4.6-25 还可以看出，优化设计还可以防止产品在脱离模具后的弯曲现象。而未经优化的模具会使得挤压件在出口处的轴向速度分布不均匀，从而导致挤压件出现弯曲现象，这自然就需要后续工序，对挤压件进行一定的弯曲或扭转矫形，导致额外能量的消耗。结果表明，以挤压力最小为目标，即在给定的冲头速度下使成形能量最小，可避免挤压件出现弯曲或扭转缺陷。

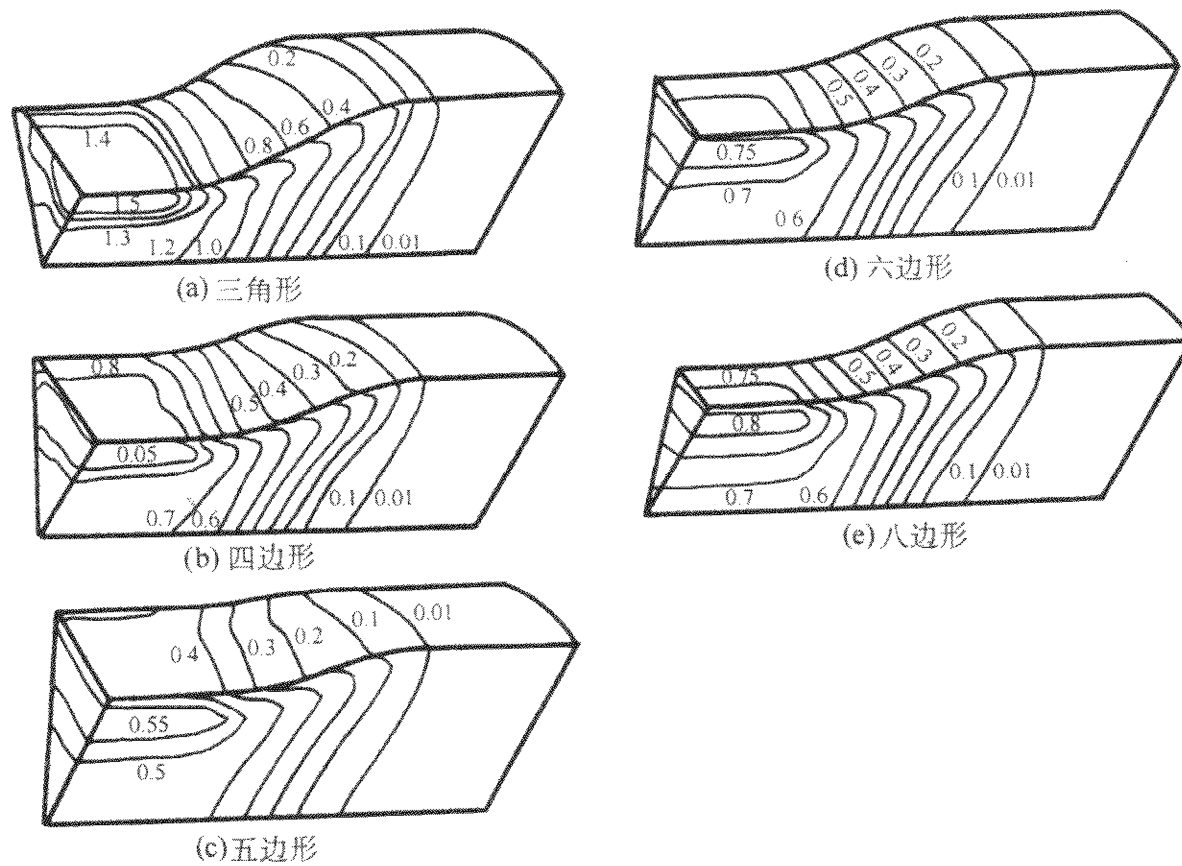


图 4.6-18　多边形截面挤压件的等效应变分布

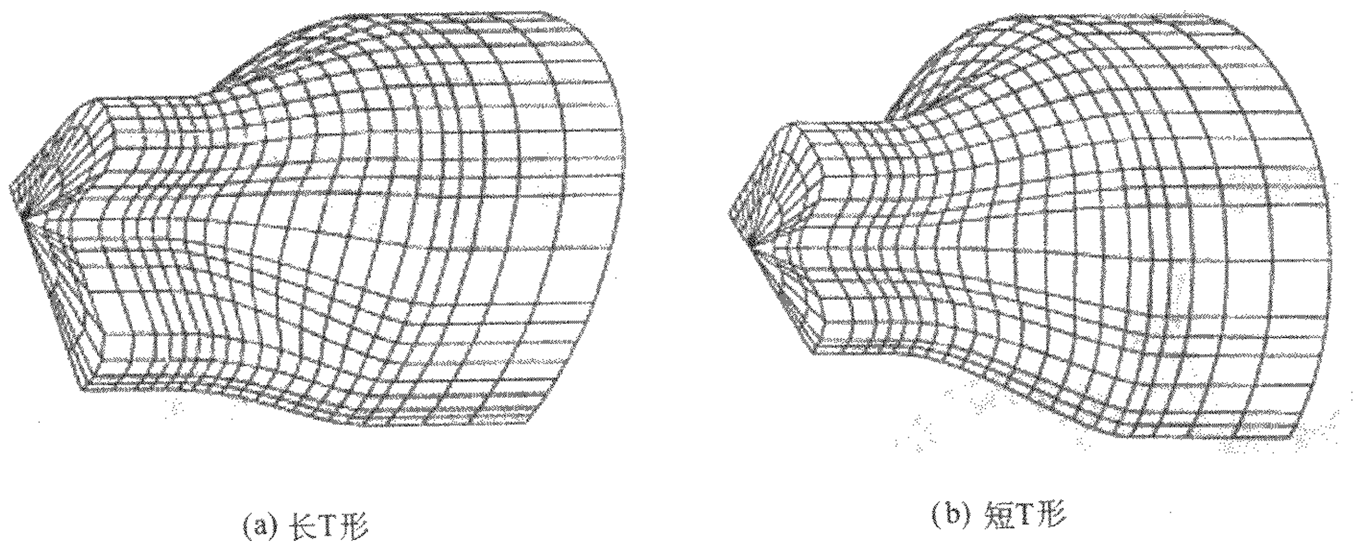


图 4.6-19　T 形截面挤压过程的有限元网格

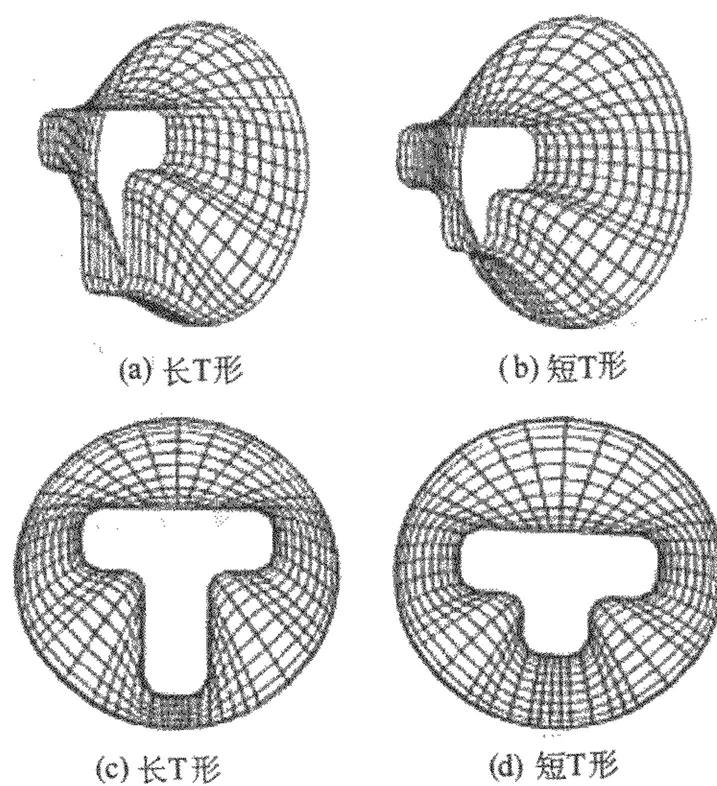


图 4.6-20　T 形截面挤压的优化模具形状

（c）和（d）相当于出口横截面内的进口横截面的相对位置

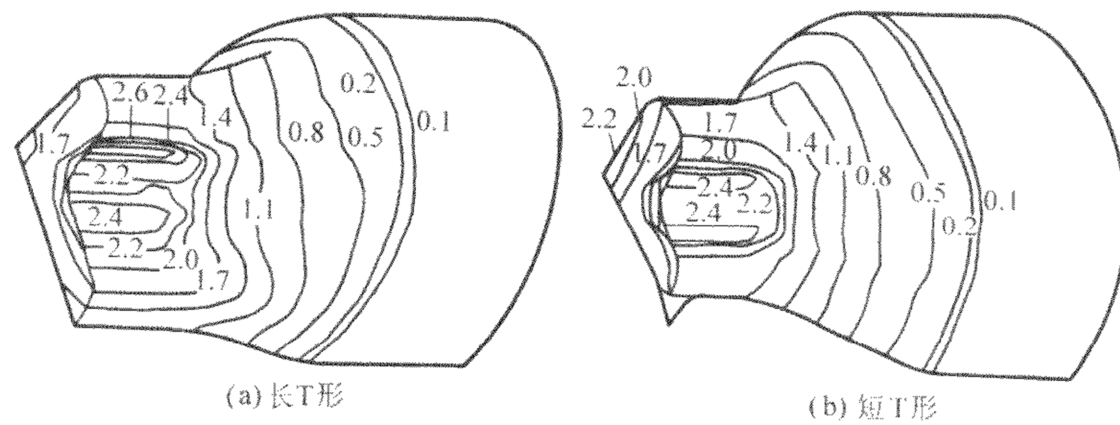


图 4.6-21　T形截面挤压件的等效应变分布

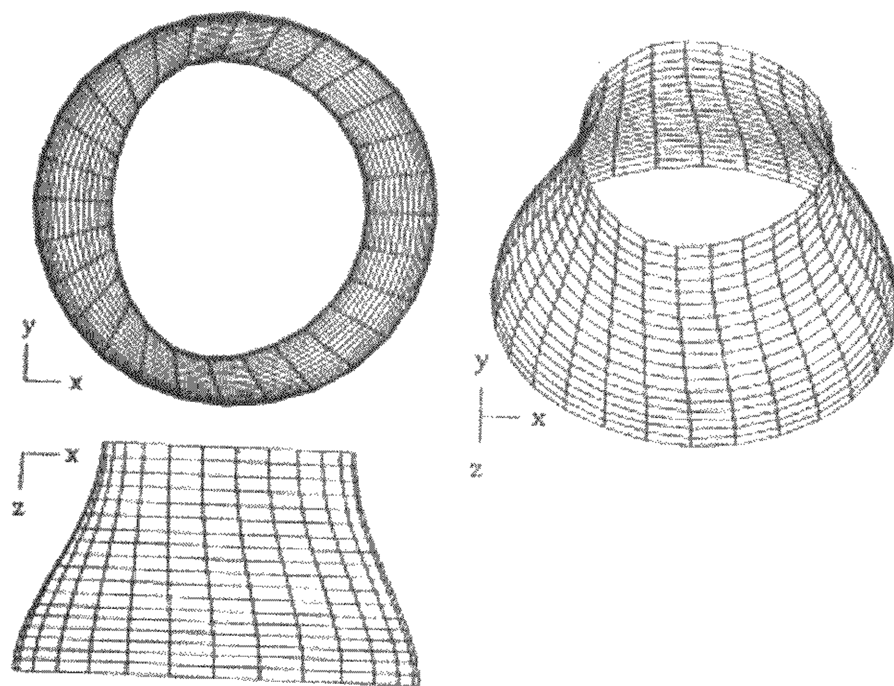


图 4.6-22　半圆半椭圆形挤压的优化模具形状

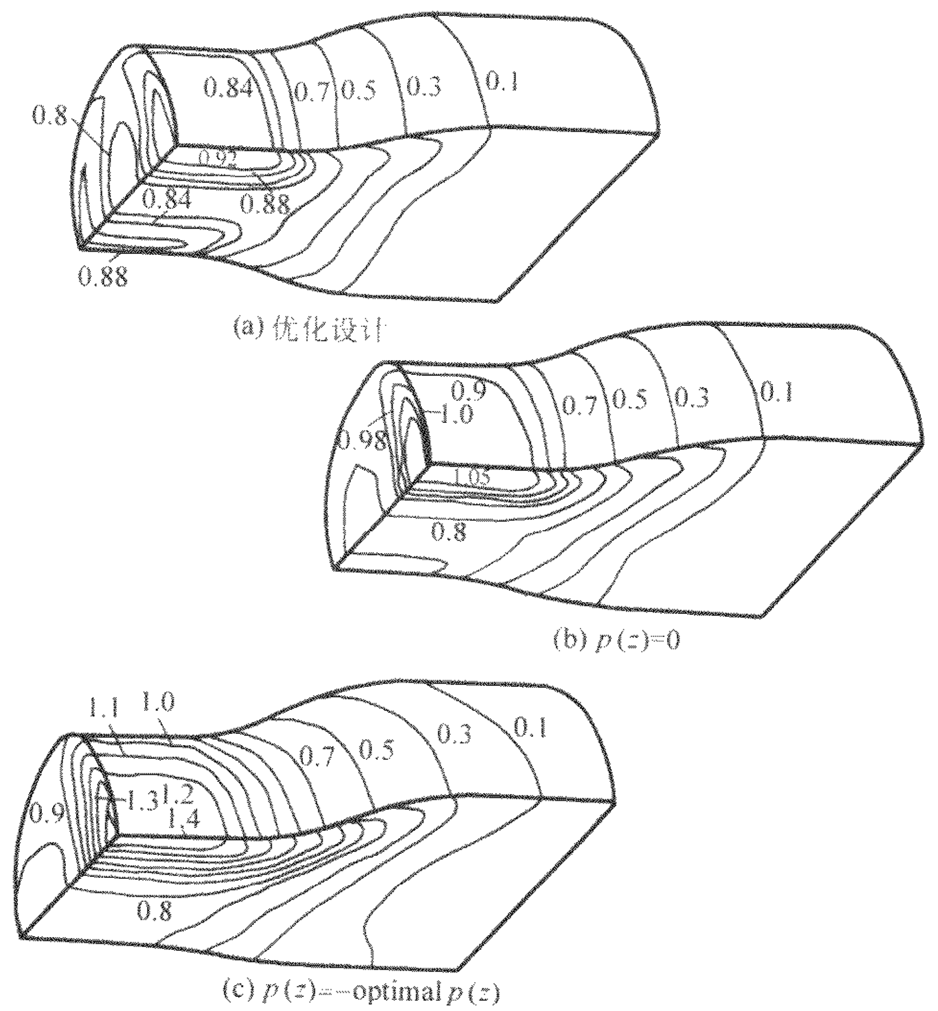


图 4.6-24　优化和无优化情况下等效应变分布的比较

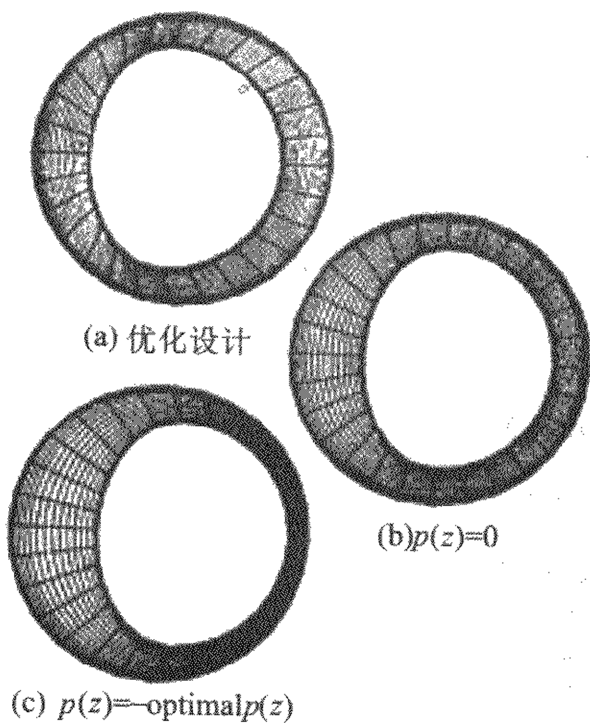


图 4.6-23　半圆半椭圆形挤压优化模具形状与无优化时的比较

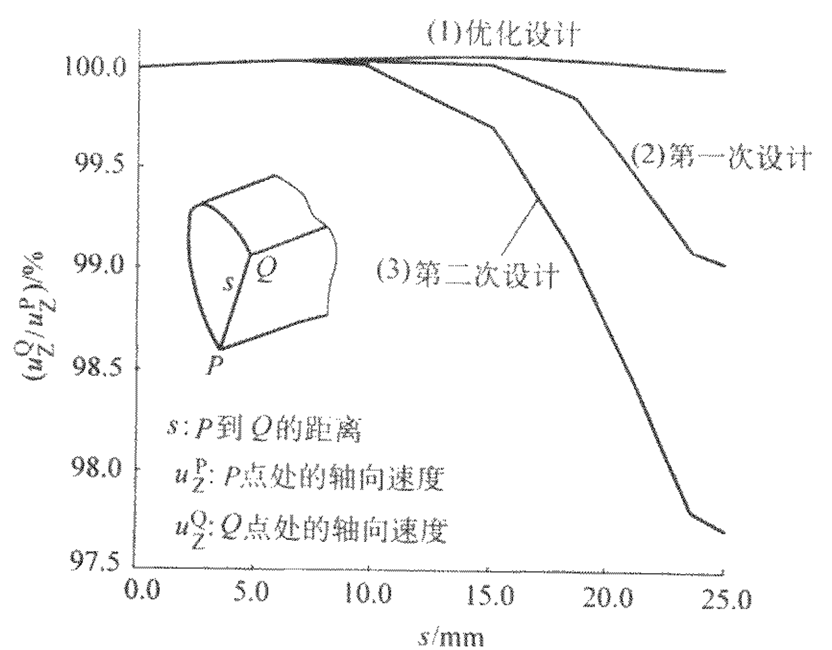


图 4.6-25　优化和无优化出口处轴向速度分布的比较

编写：赵国群（山东大学）

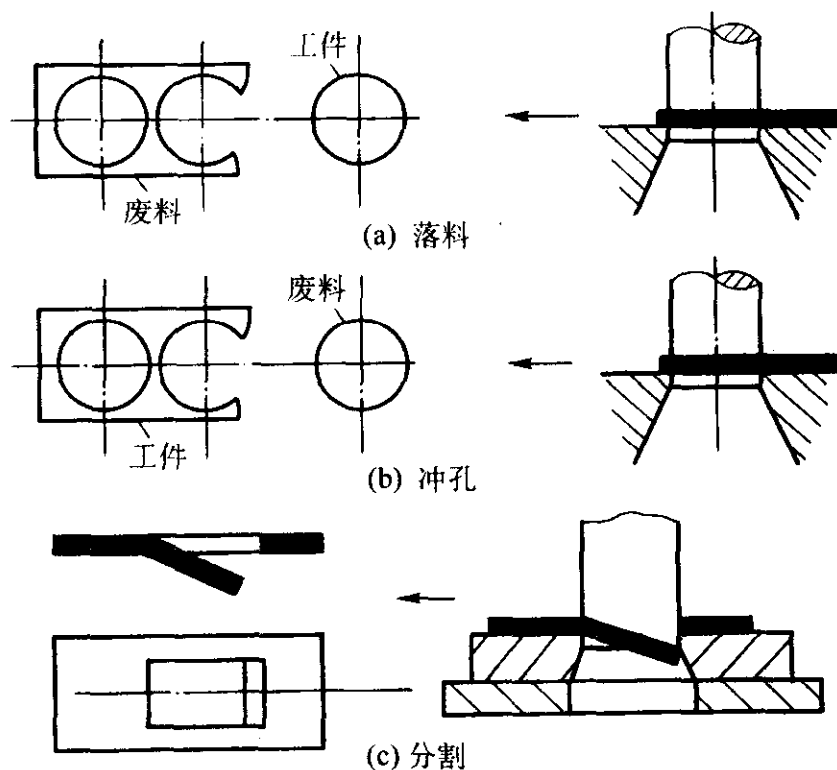
第7章 板料冲压工艺优化设计

金属板料冲压工艺优化设计，就是采用现代计算机技术、现代数值计算和优化方法，对板料冲压成形工艺安排（如冲裁、弯曲和拉深工艺等）及其各工序的模具设计进行最佳合理方案搜索，获得最佳设计方案，从而减少实际的物理试模次数，实现少无缺陷冲压成形，达到提高材料利用率、提高生产率和降低产品开发成本的目标。

1 板料冲压工艺及其质量要求

板料冲压工艺是一种先进的金属制品少无切削制造方法，以金属塑性变形为基础，利用模具和冲压设备对金属板料进行加工，以获得所需要的零件形状和尺寸。冲压工艺与切削加工相比，具有生产率高、加工成本低、材料利用率高、产品尺寸精度稳定和容易实现机械操作的自动化等一系列优点，尤其适用于大批量生产。

冲压工艺大致可以分为分离工序和成形工序两大类。分离工序又分为落料、冲孔和切割等，如图 4.7-1 所示。成形工序则可以分为弯曲、拉深、翻边、缩口、胀形和旋压等，如图 4.7-2 所示。根据零件的形状、尺寸精度和技术要求，可以采用以上工序的组合对板料毛坯进行加工，获得要求的制品。

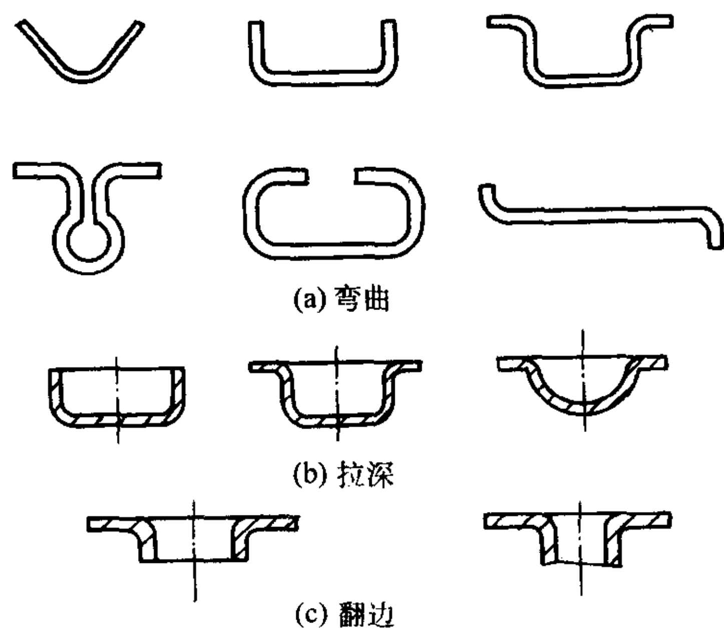


(a) 落料

(b) 冲孔

(c) 分割

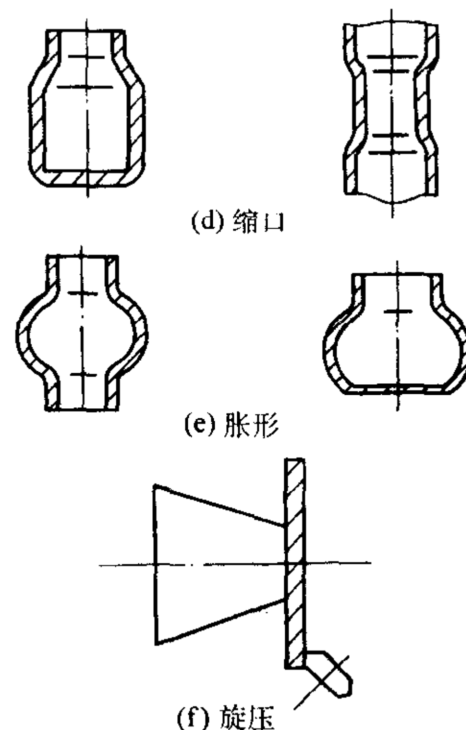
图 4.7-1 分离工序


(a) 弯曲

(b) 拉深

(c) 翻边

图 4.7-2 成形工序


(d) 缩口

(e) 胀形

(f) 旋压

续图 4.7-2

不同的冲压工序对冲压件具有不同的质量要求。例如，冲裁工艺要求冲压件切断面应平直、光洁，即无裂纹、撕裂、夹层、毛刺等缺陷，零件表面应尽可能平整，即穹弯小，同时还要求较高的材料利用率。弯曲工艺则要求制品横截面的畸变和翘曲小，避免弯曲角部外侧产生拉裂，重点控制弹性回复。拉深工艺要求合理的下料尺寸，避免产生拉裂和起皱。对于复杂形状的冲压件，如汽车结构件，因其复杂的拉深成形过程可能包含了弯曲效应，还要求有效地控制回弹。

2 板料冲压工艺优化的意义

冲压模具设计是一个反复试错的过程，传统的冲模设计采用实际试模的方法，导致模具的开发周期长，成本高。例如，日本本田汽车公司某车型由于模具开发时间滞后 3 天所带来的损失为 800 万美元，而丰田汽车公司滞后 18 天所带来的经济损失则高达 5 000 万美元。随着现代计算机技术、数值计算方法和有限元数值模拟技术的发展，模具设计的合理性检验可以通过计算机分析快速实现，从而使得设计人员可以及时地对不合理的模具设计进行修改。通过工艺优化，可以利用计算机代劳大量的工程计算工作，在设计的合理空间进行高效的搜索，最终获得比较合理的模具设计。因此，板料冲压工艺的设计优化，可以大大降低试模次数，显著缩短冲模开发时间，降低开发成本。

冲裁工艺优化的核心是在获得断面质量符合要求的冲裁件的前提下进一步提高材料利用率；弯曲工艺优化的核心主要是有效地控制下料尺寸，同时，有效控制和补偿回弹角度。拉深成形工艺优化的核心主要是获得合理的下料尺寸、合理的变形力控制，避免起皱和拉裂，对于复杂的成形过程，还要控制和补偿回弹。

3 板料成形常用数学优化方法

除了本篇第 2 章介绍的工程优化设计方法，在板料冲压工艺优化领域，还常使用正交实验法、人工神经网络法、遗传算法、响应面法等，本篇第 5 章第 3 节对遗传算法的基本原理已经进行了介绍，此处不再赘述。

3.1　正交实验法

对于各种板料冲压工艺，一般很难建立冲压质量与各工艺参数的函数关系，有时甚至还不清楚各工艺参数影响冲压质量的灵敏度。因此，常采用正交设计的方法，以数值模拟为工具，进行设计实验（Design of Experiment，DOE），从而可以发现冲压质量关于各工艺参数的灵敏度，为工艺优化奠定基础。

一般地，称影响试验指标的因素为因子，每个因子可能处的状态为水平。正交设计方法是一种研究多因子试验问题的重要数学方法，通过使用正交表进行整体设计、综合比较、统计分析。这种方法的最大优点是从许多实验条件中选择最具有代表性的少数实验就能获得可靠的实验结果，即发现哪个因子的影响是主要的，哪个因子的影响是次要的。

正交表是一些已经制作好的规范化的表，是正交实验设计的基本工具。最常用的是水平数相等的正交表和水平数不相等的正交表（也称为混合水平正交表）。正交表所共有的特性称为正交性，正是由于这种正交性，才使得用正交表安排的试验，具有均匀分散性和整齐可比性。如表4.7-1所示为一水平数相等的正交表 L_8（2^7）。其中 L 代表正交表；8为表的行数，即试验方案数；2为水平数，用数据1和2表示；7为列数，即因子数目。如表4.7-2所示为一混合水平正交表 L_8（$4^1 \times 2^4$）。正交表的第一列为4水平列，用1、2、3和4表示，其余列为2水平列，用1和2表示。该表共有5列，即5个因子。当各因子间存在相互作用时，还需要使用正交作用表。

表4.7-1　正交表 L_8（2^7）

序号＼列号	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

表4.7-2　混合水平正交表 L_8（$4^1 \times 2^4$）

序号＼列号	1	2	3	4	5
1	1	1	1	1	1
2	1	2	2	2	2
3	2	1	1	2	2
4	2	2	2	1	1
5	3	1	2	1	2
6	3	2	1	2	1
7	4	1	2	2	1
8	4	2	1	1	2

正交试验结果的分析可以采用直观分析或方差分析的方法。如表4.7-3所示为4因子3水平的正交实验的直观分析表。先计算出试验结果，然后根据极差的大小排序，极差越大说明该因子对试验结果的影响越大，该因素越重要。画出各因子与试验结果的关系图后，即可以看出各因子的水平值对试验结果的影响趋势，从而可以确定最佳的参数组合。关于正交表的方差分析，此处不再赘述，详见文献［29］。

表4.7-3　直观分析表

试验序号＼试验因子		1（A）	2（B）	3（C）	4（D）	试验结果
1		1	1	1	1	Y_1
2		1	2	2	2	Y_2
3		1	3	3	3	Y_3
4		2	1	2	3	Y_4
5		2	2	3	1	Y_5
6		2	3	1	2	Y_6
7		3	1	3	2	Y_7
8		3	2	1	3	Y_8
9		3	3	2	1	Y_9
直观分析计算	水平和 T_1	T_{1A}	T_{1B}	T_{1C}	T_{1D}	
	T_2	T_{2A}	T_{2B}	T_{2C}	T_{2D}	
	T_3	T_{3A}	T_{3B}	T_{3C}	T_{3D}	
	水平均值 R_1	R_{1A}	R_{1B}	R_{1C}	R_{1D}	
	R_2	R_{2A}	R_{2B}	R_{2C}	R_{2D}	
	R_3	R_{3A}	R_{3B}	R_{3C}	R_{3D}	
	极差 R	R_A	R_B	R_C	R_D	

表中，T_{ij} 表示第 j 列水平值为 i 时的试验结果之和；$i=1,2,3$；$j=A,B,C,D$；$R_{ij}=T_{ij}/N$；N 为水平数目，此处 $N=3$。

$$R_j = \max\{R_{1j}, R_{2j}, R_{3j}\} - \min\{R_{1j}, R_{2j}, R_{3j}\} \quad (4.7\text{-}1)$$

3.2　人工神经网络

人工神经网络是一种大规模的分布式并行处理系统，它可以模拟人脑的记忆和联想功能，不需要了解过程的输入与输出参数之间的变化规律。通过对给定的样本数据进行学习，从大量的数据中提取规则，以一组权重系数的形式形成一种网络的稳定状态，然后通过联想记忆和推广能力来获取所需数据。人工神经网络的特点：并行性、稳健性、容错性和自学习能力。在人工神经网络中，神经元是最基本的处理单元，其数学模型如图4.7-3所示，其中 X_{ji} 为第 j 个神经元对第 i 个神经元的输入，W_{ij} 为第 j 个神经元对第 i 个神经元的连接权值，θ_i 为第 j 个神经元的阈值，f 为传递函数，y 为该神经元的输出。神经元的输入、输出关系如下：

$$Y = f\left(\sum W_{ij}X_{ji} - \theta_i\right) \quad (4.7\text{-}2)$$

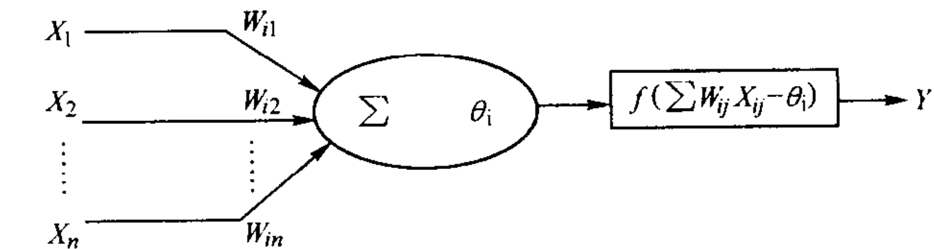


图4.7-3　神经元的数学模型

根据神经元之间连接方式的不同，人工神经网络分为：没有反馈的前向网络和有反馈的前后向结合型网络。前向网络（如图4.7-4所示）由输入层、输出层和中间层（隐层）组成，中间层可以有若干层，每一层的神经元只接受来自紧邻的前一层神经元的输出，一般不接受超越层次的输出。反馈型网络（如图4.7-5所示）的任意两个神经元之间都可能有连接，包括神经元到自身的反馈。在众多的人工神经网络模型中，以基于误差反向传播（Back Propagation，BP）的前向网络应用最广。

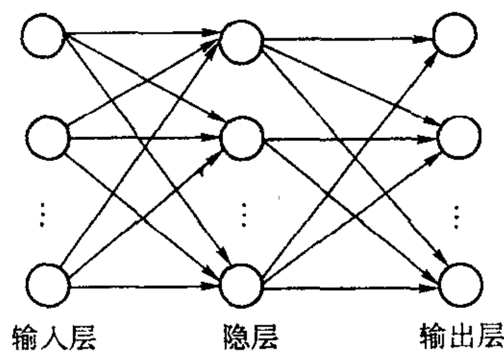


图4.7-4 具有一层隐层的前向网络

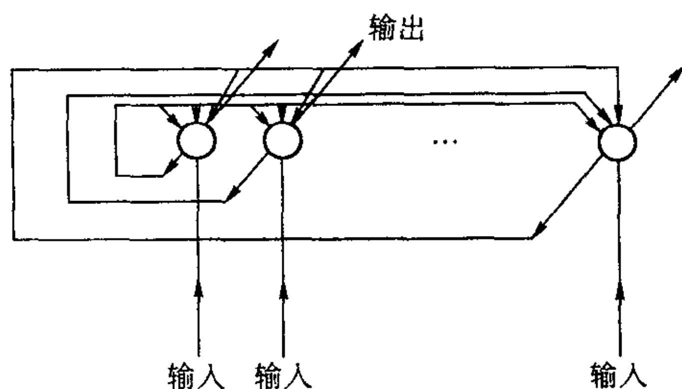


图4.7-5 反馈型网络

在人工神经网络结构确定后，如何确定网络各神经元之间的连接权值是一个重要问题。对于非线性变换单元所组成的BP网络，变换关系是一个非线性连续可微函数，它可以严格利用梯度法进行推算，权的学习解析式十分明确。一般情况下，BP算法包含两个阶段：一是前向传播阶段，当给定输入量时，通过神经网络向前计算每一个神经元的输出，然后计算期望输出值 T_j 与实际输出值 X_j 之间的误差，均方差 E 表示为：

$$E = \frac{1}{2}\sum (T_j - X_j)^2 \tag{4.7-3}$$

二是误差反向传播阶段，将误差从输出层向输入层的方向进行反向传播以修正相应的连接权值，修正量可表示为：

$$\Delta W_{ji}^k(m+1) = -\eta \partial E/\partial W_{ji} \tag{4.7-4}$$

为了防止解的振荡，加一个惯性项。因此，新的权值修正量 ΔW_{ji}^k 可以表示为：

$$\Delta W_{ji}^k(m+1) = -\eta \partial E/\partial W_{ji} + a\Delta W_{ji}^k(m) \tag{4.7-5}$$

式中，η 为学习率；a 为动量因子；m 为迭代次数。

神经网络学习的步骤如下：

1）确定BP网络的结构，即网络层数及各层神经元数量；

2）给出学习样本的输入、输出向量，如有必要，做归一化处理；

3）初始化，随机设置各层的权系数和阈值；

4）向网络提供输入样本和理想输出值；

5）计算实际输出值；

6）计算误差；

7）若不满足误差要求，则返回继续学习；否则，则认为计算收敛。

3.3 响应面法

响应面法（Response Surface Method，RSM）是以试验设计为基础的用于处理多变量问题建模和分析的统计处理技术。最早由G.E.P.Box和K.G.Wilson提出，但初期的响应面法都没有考虑噪声因素，直至二十世纪80年代R.H.Myers把噪声因素引入响应面法后才在工业界得到广泛的应用。

例如，若某产品的输出特性与温度 x_1 和压力 x_2 有关，则可以把观察到的响应量 y 写成温度和压力的函数，即：

$$y = f(x_1, x_2) + \varepsilon \tag{4.7-6}$$

式中，ε 为随机误差。

如果用 $E\{y\} = \eta$ 表示期望的响应值，则由 $\eta = f(x_1, x_2)$ 表示的曲面称为响应面，如图4.7-6所示为响应面在 x_1 和 x_2 坐标平面上的等值线投影。

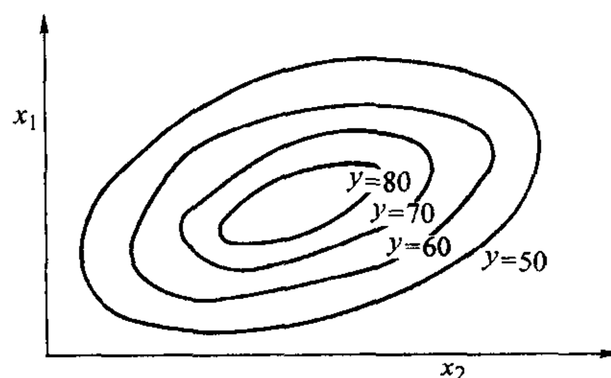


图4.7-6 响应面的等值线

在大部分设计问题中，响应量和设计变量之间的关系形式是未知的。因此，RSM的第一步是求出 $f(x)$ 的适当近似。通常在自变量的某个范围内用低次多项式近似，如用自变量的线性函数（一次模型）：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon \tag{4.7-7}$$

或用较高次的多项式，如二次模型表示：

$$y = \beta_0 + \sum_{i=1}^{n}\beta_i x_i + \sum_{i=1}^{n}\beta_{ii}x_i^2 + \sum_{i<j}\sum \beta_{ij}x_i x_j + \varepsilon \tag{4.7-8}$$

响应面法用于优化设计一般分为三个阶段：因素的筛分、寻域和优化。先用少数几次试验筛分出影响产品质量特性（响应量）或者与噪声因素相互影响的主要设计参数，从这些试验中拟合出线性模型。通过模型分析查明参数在设计空间的变动范围，并应确定按哪个方向去寻找最佳的设计参数值。当设计参数的变动区域确定后，进行最后的试验，拟合出二阶的响应面模型，通过模型分析，确定最佳的参数组合。

4 冲裁工艺中的毛坯排样优化技术应用

4.1 冲裁工艺中的排样技术及排样方法

在大批量生产的冲压零件的成本中，材料费用占60%以上，因此，通过合理的冲裁工艺降低材料消耗（尤其是有色金属和贵重材料），对提高冲压生产的经济效益具有重要意义。

冲裁件在条料或板料上的布置方法称为排样。排样的合理与否不仅直接影响材料利用率，而且还会影响模具寿命、冲压生产率、工件精度和操作的安全性等。排样的材料利用率可以表示如下：

单个零件的材料利用率：$\eta_1 = \dfrac{n_1 F}{Bh} \times 100\%$ (4.7-9)

条料的材料利用率：$\eta_2 = \dfrac{n_2 F}{LB} \times 100\%$ (4.7-10)

板料的材料利用率：$\eta_3 = \frac{n_3 F}{L_0 B_0} \times 100\%$　　(4.7-11)

式中，F 为冲裁件面积，mm^2；B 为条料宽度，mm；h 为送进料距，mm；η_1 为一个进距内冲压件数；η_2 为一条条料上的冲压件数；η_3 为一张板料上的冲压件数；L 为条料长度，mm；L_0 为板料长度，mm；B_0 为板料宽度，mm。

4.2 提高材料利用率的途径

如图4.7-7所示为某一零件冲裁时的排样，阴影部分为废料，其中废料包括工艺废料和结构性废料，提高材料利用率的关键就是减少废料。主要有以下几种方法。

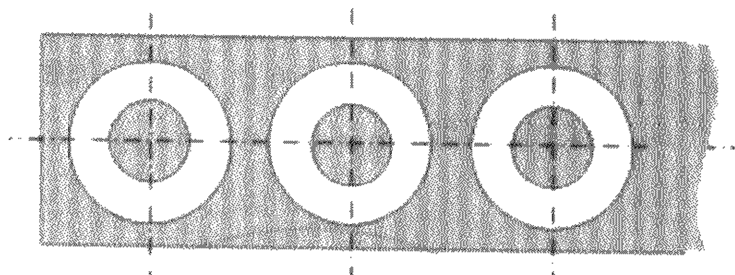


图4.7-7　冲裁件的排样示意图

1）采用少、无废料的排样方式　如图4.7-8所示为少、无废料排样方式，其中包括直排、斜排、直对排和斜对排。这种排样方式兼有生产率高、模具结构简化和冲裁力低的优点，但是工件的尺寸精度和断面质量较差，而且模具寿命低。

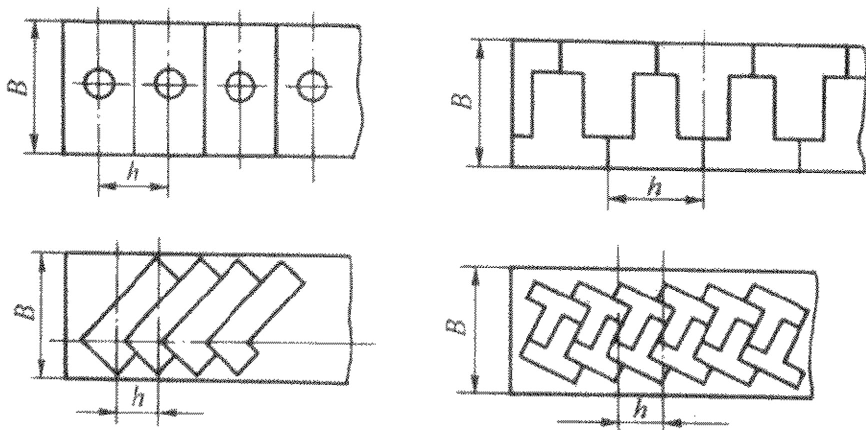


图4.7-8　少、无废料排样方式

2）采用混合排样方式　如图4.7-9所示为混合排样方式，即将不同形状的零件在一副模具中冲裁。

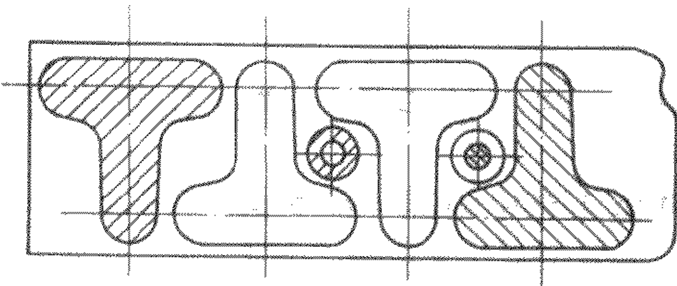


图4.7-9　混合排样方式

3）采用套冲排样方式　如图4.7-10所示为零件套冲排样的方式，充分利用了大零件的工艺废料。

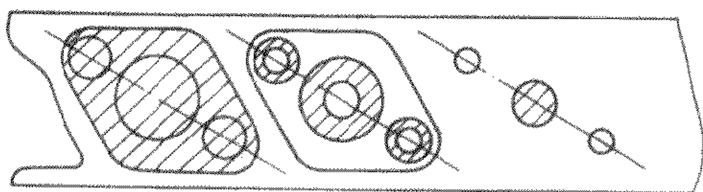


图4.7-10　套冲排样方式

4）修改零件的形状　如果可能，在保证零件主要技术要求的前提下，适当修改零件的形状，以利于合理排样，从而提高材料利用率。图4.7-11a所示为某工件初始形状时的排样，材料利用率只有38%，在保证孔距50 mm和60 mm的条件下，将零件形状稍做修改（如图4.7-11b所示），材料利用率提高到79%。

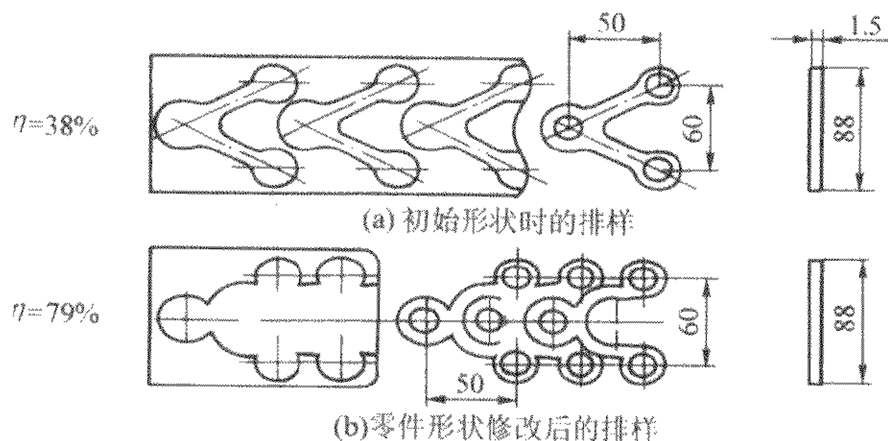


图4.7-11　通过修改零件形状提高材料利用率

4.3 毛坯排样的优化技术应用

由于手工毛坯排样和优化效率低，不可能达到最优的材料利用率，因此基于计算机技术的毛坯排样优化技术自20世纪80年代以来一直是冲模CAD领域研究的热点技术之一。毛坯排样属于多约束条件下的复杂优化问题。由于问题的高复杂度，除了穷举所有解外，没有高效的结构化的方法来搜索解的空间，因此实际上无法用单纯的数学优化方法求得精确解。毛坯排样的计算机优化过程可以按照以下步骤实现。

1）生成毛坯的二维模型。

2）选择不同的毛坯排样模式，目前最常采用的模式有普通单排、对头单排、普通双排和对头双排四种，如图4.7-12所示。

3）根据选定的排样模式、给定的条料尺寸和初始设定的工艺参数值（如图4.7-13中所示 ϕ 和 m），进行毛坯图的等距放大。

4）计算当前排样的材料利用率，对于初始排样，置最佳排样模式为初始排样模式，并置最优材料利用率为初始材料利用率；对于非初始排样，将当前材料利用率与当前最佳排样模式下的材料利用率进行比较，确定新的最优排样模式和相应的材料利用率。

5）按照一定的步长调整工艺参数 ϕ 和 m。

6）重复3）～5），直至完成约束空间的所有离散解的搜索。

毛坯图形等距放大处理时，考虑到工件间搭边，一般将工件图外轮廓沿法线方向等距偏置半个工件间搭边值的距离。在排样设计中，将以这样的等距图代替原工件外轮廓图形，等距图之间若相切即保证了冲件之间必需的最小搭边值。常规的由直线和圆弧构成的闭环的等距放大算法比较成熟，关键是依次求取毛坯轮廓图形环上两相邻（首尾相接共端点）图元偏置后的有效交（切）点。但对于带有内凹的图形，等距图形可能发生自交。

由于从原始毛坯图到等距放大图，其形状可能发生变异（特别是非圆弧过渡处），因而传统算法在生成毛坯排样图时，很难非常准确地由等距图间的位置关系来安排各原始毛坯的位置；并且从变异的等距图计算出的料宽不便修正。通过安排原始毛坯图参与排样过程并记录其位置关系，可以方便地生成排样图，而且料宽计算准确。

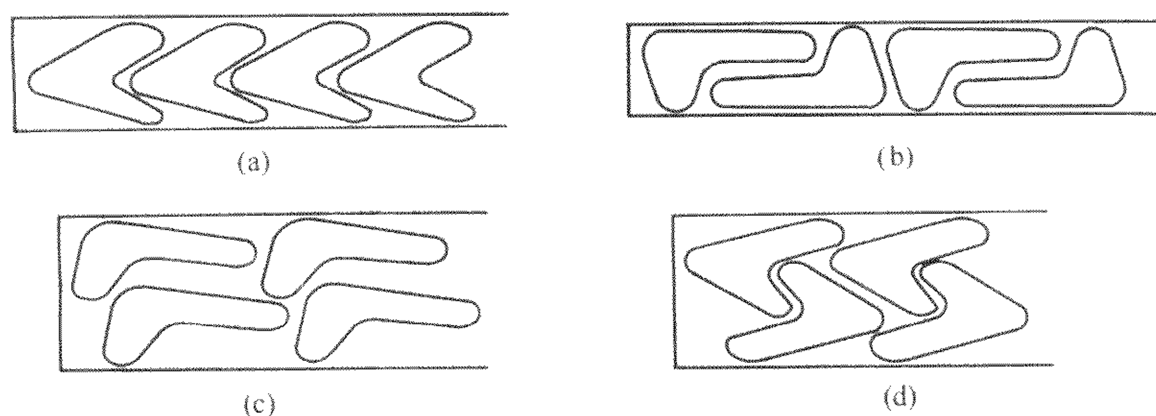


图 4.7-12　目前常用的排样模式

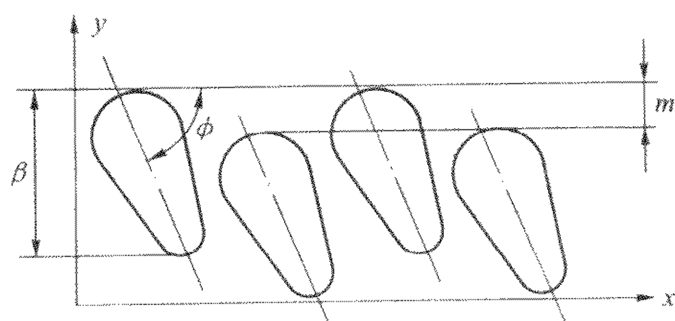
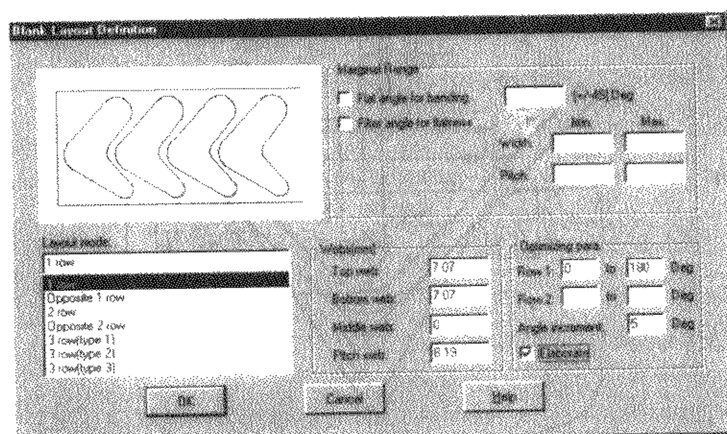
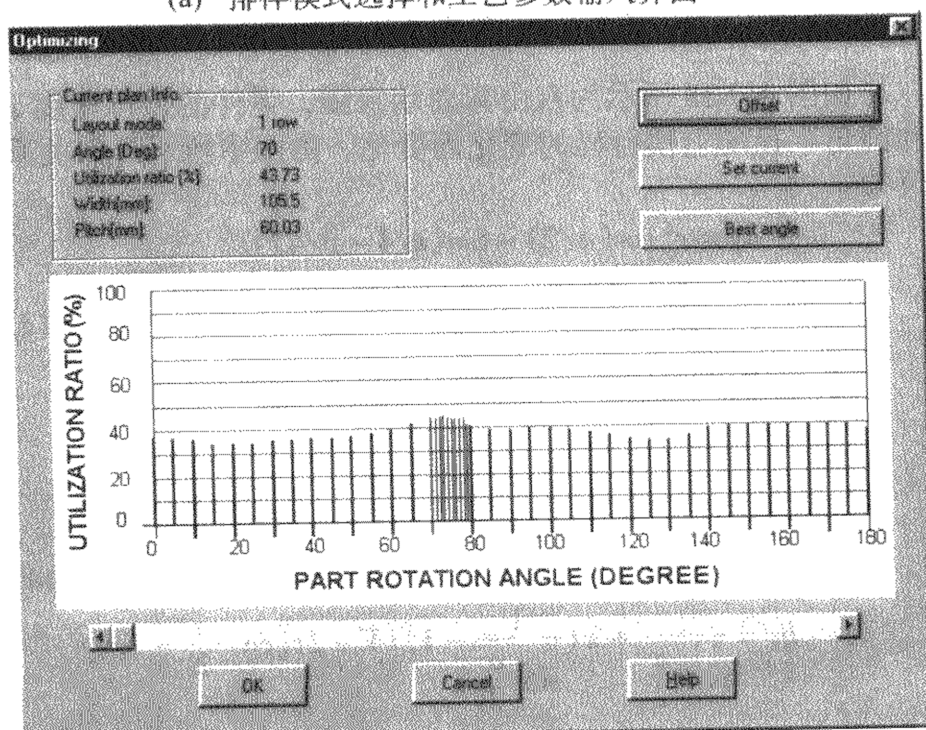


图 4.7-13　排样图

目前许多商业化的 CAD/CAM 软件大都提供二次开发函数，使得 CAD 软件中的成熟算法（如曲线偏置、几何图形间求最短距离和对应点的坐标等）可以被成功应用，大大减少了专业化软件的开发工作量。如图 4.7-14a 所示为在某 CAD 软件上开发的冲裁毛坯排样优化模块的界面，图4.7-14b 所示为材料利用率随零件旋转角度变化的规律，友好的用户界面可以方便地实现毛坯排样最佳参数的选择。



(a) 排样模式选择和工艺参数输入界面



(b) 毛坯排样优化结果

图 4.7-14　冲裁毛坯排样优化软件功能界面

5　弯曲工艺中回弹和下料尺寸控制的优化技术

5.1　弯曲工艺中下料尺寸控制的优化技术

板料弯曲过程中，要保证得到满足尺寸要求的零件，毛坯下料尺寸的计算至关重要，弯曲件的展开长度是根据弯曲前后中性层长度不变的原则进行计算的。在弯曲成形过程中，板料外层受拉，尺寸变长；而内层受压，尺寸变短；内、外层之间必然存在着弯曲前后长度不变的某一中性层。根据弯曲件板料展开长度等于中性层长度的原则，只要确定中性层的位置，就能够得到其展开长度。中性层位移系数与板料厚度、弯曲半径和弯曲角度有着复杂的关系，这种关系很难用公式或其他形式显式表示。可以以企业试验和生产数据为基础，利用人工神经网络模型表示中性层位移系数与板料的厚度、弯曲半径和弯曲角度间的复杂关系，以此计算板料的下料尺寸。以单角弯曲板料展开长度计算为例，如图 4.7-15 所示。单角弯曲板料展开长度采用以下公式计算：

$$
\begin{aligned}
L &= L_1 + L_2 + \frac{\pi\ (180^\circ - \alpha)}{180^\circ}\ (R + xt)\ - 2\ (R + t)\ \cot\frac{\alpha}{2} \\
&= l_1 + l_2 + \frac{\pi\ (180^\circ - \alpha)}{180^\circ}\ (R + xt) \qquad (4.7\text{-}12)
\end{aligned}
$$

式中，L_1、L_2 为板料外形长度；l_1、l_2 为板料各平直部分长度；α 为弯曲成形角度；R 为弯曲成形内半径；t 为板料厚度；x 为板料中性层位移系数。

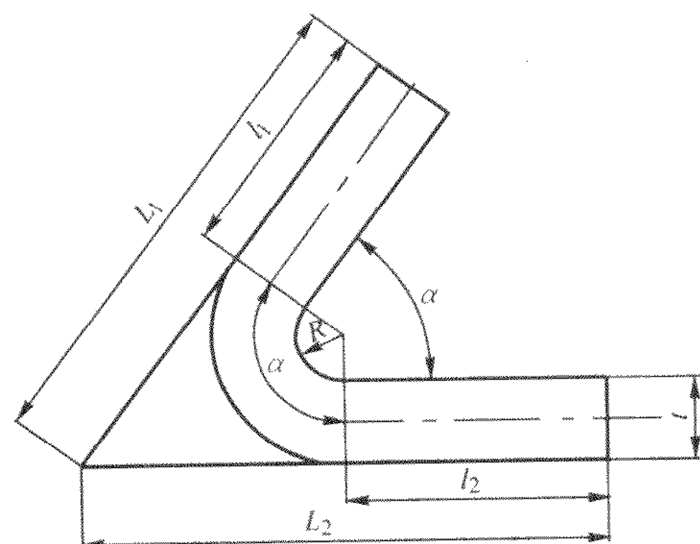


图 4.7-15　单角弯曲的零件示意图

在利用式（4.7-12）计算板料的展开长度时，需要得到中性层位移系数 x。以下采用神经网络来模拟弯曲角度、弯曲半径、板料厚度和中性层位移系数间的隐含关系。

一个 3 层的 BP 人工神经网络模型能够实现任意精度的近似连续的函数。因而采用 BP 网络来获得中性层位移系数，从而计算弯曲板料展开长度，网络模型如图 4.7-4 所示。BP

表 4.7-4　部分样本的原始数据

序号	板料厚度 t/mm	成形内半径 R/mm	成形角度 α/(°)	板料外形长度 L_1/mm	板料外形长度 L_2/mm	展开长度 L/mm
1	1	0.6	147	50.50	50.53	100.50
2	1	0.6	129	50.50	50.80	100.50
3	2	1.5	121	51.00	51.35	100.5
4	3	4	150	51.00	51.12	100.5
5	3	7	135	50.60	52.33	100.5
6	4	9	120	100.80	102.59	200

算法是一种有教师的学习算法，必须利用样本对其进行学习，这里利用的是企业试验和生产中积累的大量数据。样本由板料厚度、弯曲半径、成形角度、弯曲件外形长度和下料尺寸组成，如表 4.7-4 所示。测量不能直接得到成形后板料的中性层位移系数，但是利用式（4.7-12）反向求得中性层位移系数，形成神经网络学习所需的样本。

弯曲成形过程中，中性层位移系数与弯曲角度、弯曲半径和板料厚度相关。因此将这三个参数作为网络模型的输入层变量。中性层位移系数为待识别量，因而作为输出变量。该神经网络模型采用 $3\times5\times5\times5\times1$ 结构。

采用多层神经网络时，隐层单元数目的选择是一个重要的问题。在没有很准确的解析式表示的情况下，隐层单元数目与问题的要求、输入与输出单元的多少都有直接关系。根据经验和识别精度的要求，选择隐层单元数为 $5\times5\times5$。

对神经网络进行训练时，学习步长的选择很重要，学习步长大，收敛快，但过大可能引起解的振荡；学习步长小可避免不稳定，但过小会导致收敛速度慢。此处学习步长取为 0.7。为了提高收敛的质量，在权值调整时，增加了一项动量项，其作用是推动权值随上次权值的变化而变化。动量项为动量系数乘上一次的权值的调整值，动量系数可以置为 0.5。

针对某种钢板，根据企业提供的原始数据，提取大部分作为样本训练数据，其余作为验证数据。通过对网络进行训练，得到板料厚度、成形半径、成形角度和中性层位移系数的关系。神经网络训练结束后，采用验证数据对其进行了检验。发现计算值与验证数据的实测值之间的误差较小，能够满足企业的生产需要。

5.2　弯曲工艺中回弹控制的优化方法

由于板料在弯曲过程中发生弹塑性变形，在弯曲圆角部分，内层材料承受压应力，外层材料则承受方向相反的拉应力。因此，卸载后材料必然产生弹性回复，如图 4.7-16 所示。为了保证弯曲件的形状和尺寸，必须控制弹性回复。

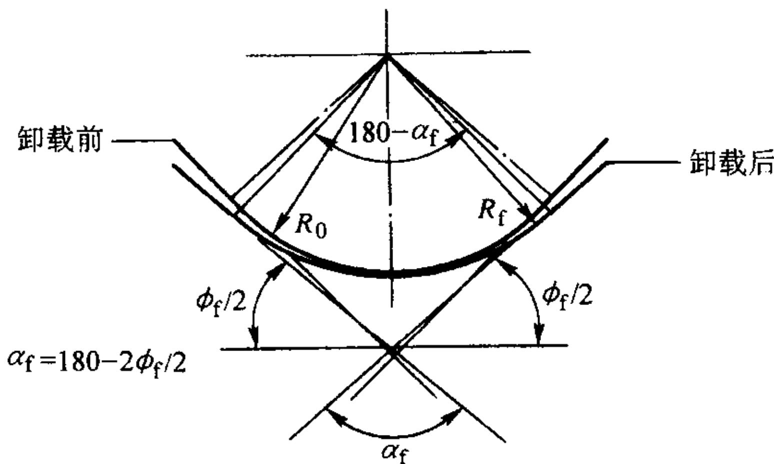


图 4.7-16　弯曲结束后的弹性回复

弯曲工艺可以分为自由弯曲和校正弯曲，如图 4.7-17 所示为弯曲力的变化曲线，校正弯曲前的阶段属于自由弯曲。自由弯曲一般在折弯机上利用通用的冲头和凹模实现，而校正弯曲则常利用专门设计的凸、凹模完成特定零件的弯曲。关于校正弯曲时回弹的控制比较复杂，有时甚至要修改模具的形状。此处重点给出自由弯曲时回弹的控制与优化方法。

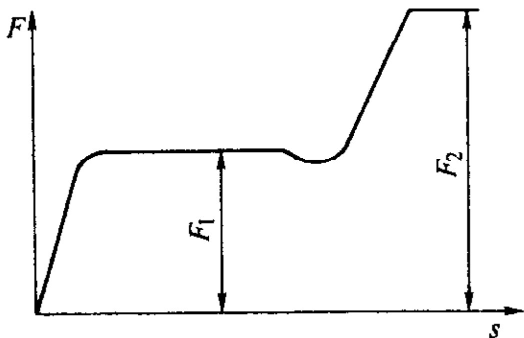


图 4.7-17　弯曲过程中力的变化曲线

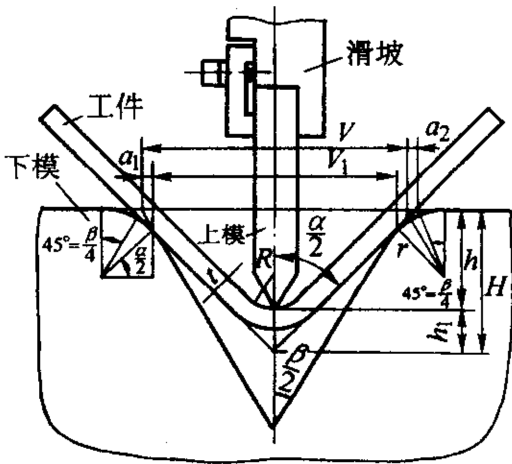


图 4.7-18　自由弯曲示意图

如图 4.7-18 所示为自由弯曲示意图，零件的折弯角度 α 和圆角半径随上模行程的变化而变化。由图中所示几何关系可知：

$$
\begin{aligned}
a_1 &= r[\cos(\alpha/2)-\tan(\pi/4-\beta/4)]\\
a_2 &= r[\tan(\pi/4-\beta/4)-\tan(\pi/4-\alpha/4)]\\
h &= H-h_1\\
R &= kV_1
\end{aligned}
\tag{4.7-13}
$$

而：

$$
H=\left\{\frac{V}{2}+r[\tan(\pi/4-\beta/4)-\tan(\pi/4-\alpha/4)]\right\}/\tan\frac{\alpha}{2} \tag{4.7-14}
$$

$$
h_1=\frac{t}{\sin\frac{\alpha}{2}}+k\left(\frac{1}{\sin\frac{\alpha}{2}}-1\right)\left\{V-2r\left[\cos\frac{\alpha}{2}-\tan\left(\frac{\pi}{4}-\frac{\beta}{4}\right)\right]\right\} \tag{4.7-15}
$$

将式（4.7-14）和式（4.7-15）代入式（4.7-13），简化可得：

$$h = [\frac{1}{2\tan(\alpha/2)} - \frac{k}{\sin(\alpha/2)}]V - \frac{t}{\sin(\alpha/2)} + r\{[1 - \sin\frac{\alpha}{2}] - 2[\cos\frac{\alpha}{2} - \tan(\frac{\pi}{4} - \frac{\beta}{4})][\frac{1}{2\tan(\alpha/2)} - \frac{k}{\sin(\alpha/2)} + k]\} \quad (4.7\text{-}16)$$

式中，α 为工件回弹前的弯曲角度；β 为下模 V 形槽的角度；V 为 V 形槽名义开口宽度，mm；V_1 为弯曲工件时的实际开口宽度，mm；r 为 V 形槽开口处圆弧半径，mm；R 为工件内圆弧半径，mm；k 为内圆弧半径系数；t 为板料厚度，mm；h 为滑块进深，mm。

式（4.7-16）建立了折弯角度 α 和冲头行程 h 的数学关系，即给定 α，即可求出 h；反之，已知 h，则同样可以确定 α。

式（4.7-13）~式（4.7-16）推导的是当 $V > V_1$ 的情况，当 $V < V_1$ 时，推导出的公式与式（4.7-16）相同。

由于工件卸载后存在回弹，因此折弯角度需要进行校正。利用下式：

$$\Delta h \approx \mathrm{d}h = \frac{\mathrm{d}h}{\mathrm{d}\alpha}\mathrm{d}\alpha \approx \frac{\mathrm{d}h}{\mathrm{d}\alpha}\Delta\alpha \quad (4.7\text{-}17)$$

$$\Delta h \approx -\{\frac{V}{2\sin^2\frac{\alpha}{2}}[\frac{1}{2} - \frac{k}{\cos\frac{\alpha}{2}}] - t\frac{\cos\frac{\alpha}{2}}{2\sin^2\frac{\alpha}{2}} + r[k(1 - \sin\frac{\alpha}{2}) - \frac{1}{\sin^2\frac{\alpha}{2}}(\frac{1}{2} - k\cos\frac{\alpha}{2})(\cos\frac{\alpha}{2} - \tan(\frac{\pi}{4} - \frac{\beta}{4}))]\}\Delta\alpha \quad (4.7\text{-}18)$$

式中，Δh 为进深校正值，mm；$\Delta\alpha$ 为工件实测角度与设定角度之差。

关于公式中 k 的取值，各个公司提供的经验值，主要适用于各公司制造的数控折弯机，如比利时 LVD 公司取 5/32，瑞士 Hammerle 公司取 0.16，美国《模具学》一书推荐为 1/8。上述公式应用时必须针对具体的机床。

例如，采用某公司生产的数控机床，在凹模角度 $\beta = 85°$ 的折弯机上，用凹模开口 $V = 16$ mm 的槽宽，把板厚 $t = 2$ mm 的板料折成 $\alpha = 90°$的零件，则根据式（4.7-16）和 $r = V/20$ 的规范，计算得滑块的进深值为：

$$h = 0.453\,9V - 1.433\,8t \approx 4.39 \text{ mm}$$

进深校正值为：

$$\Delta h \approx 5.157\Delta\alpha$$

假定工件首次折弯的实测角度为 92.5°，则校正后的滑块进深为：

$$h' = h - \Delta h = 4.39 \text{ mm} - 5.157\pi(90 - 92.5)/180 \text{ mm} \approx 4.62 \text{ mm}$$

6 板料拉深成形中的缺陷控制与工艺优化技术

板料在拉深成形过程中，各部位因应力状态不同，如果模具设计和工艺参数不合理，会产生不同的缺陷，如拉裂（径向拉应力超过强度极限）和起皱（切向压应力超过临界值）。对于复杂形状的零件的成形，因存在弯曲效应，卸载后还会产生回弹。在冲压工艺设计中，毛坯的尺寸和形状、模具型面、拉深筋的结构参数或拉深肋阻力和压边力等都会对成形质量产生影响。以下将给出上述几个工艺参数优化的方法。

6.1 基于增量有限元数值模拟的毛坯下料形状的优化

复杂形状的零件拉深成形时，初始毛坯的形状和尺寸对于成形质量至关重要。初始毛坯尺寸取得太大，导致压边力过大，易产生拉裂；如果初始毛坯尺寸太小，反而会起皱。采用增量有限元数值模拟的方法可以对初始毛坯尺寸的合理性进行检验，实现最优方案的搜索和选择。

如图 4.7-19 所示为某汽车减振器杯的拉深成形有限元数值模拟结果，变薄现象比较严重，尤其是在冲头底部，变薄量达到 20%。在对初始毛坯进行优化时，毛坯的外轮廓线由 B 样条曲线表示，如图 4.7-20 表示内轮廓线（Lower bounds），外轮廓线（Upper bounds）为目前采用的下料型线。8 个设计变量为图中所示的 8 个控制节点的位置，这 8 个控制点只能沿着各自位置的法向移动。优化的目标是使材料的用量最小和材料内部厚度变化量的不均匀程度最小。初始解为内部的点划线轮廓位置，优化方法选择序列二次规划法（Sequential Quadratic Programming Method），经过 13 次迭代后，材料的体积降低了 16.6%，厚度的不均匀性大大改善，如图 4.7-21 所示。

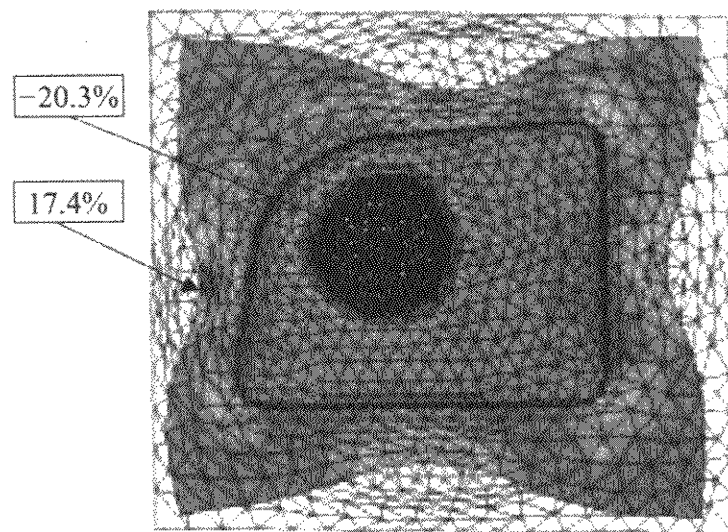


图 4.7-19 某减振器杯拉深成形数值模拟结果
（厚度的相对变化量）

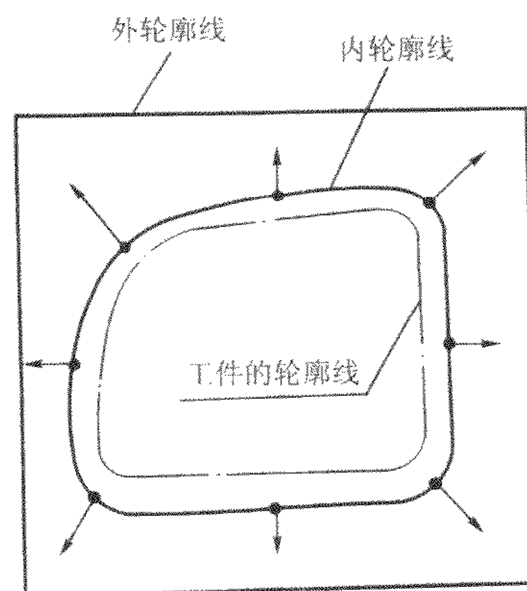


图 4.7-20 毛坯的参数化轮廓

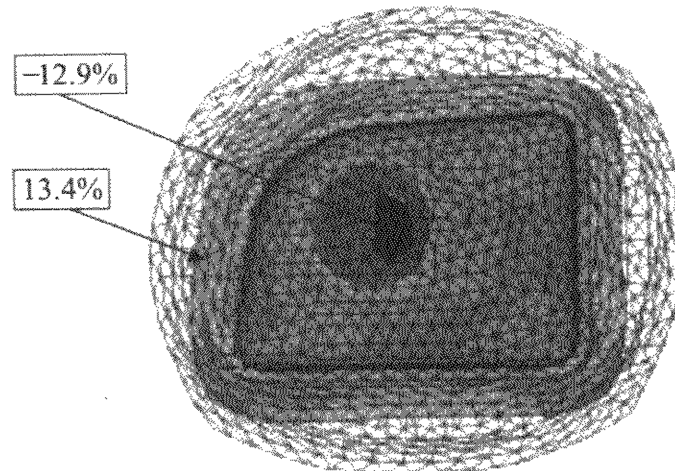
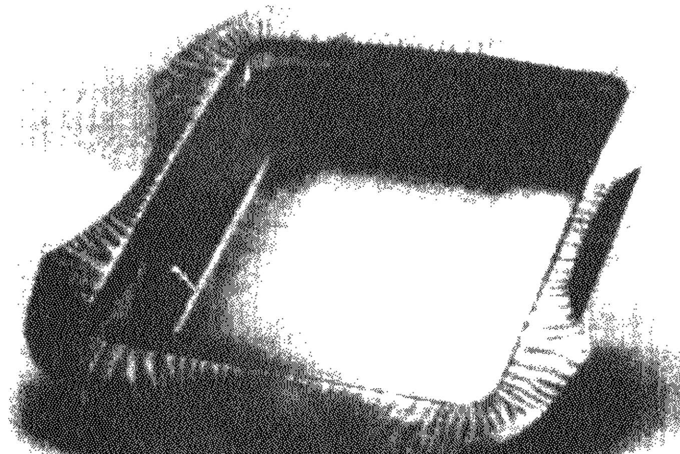


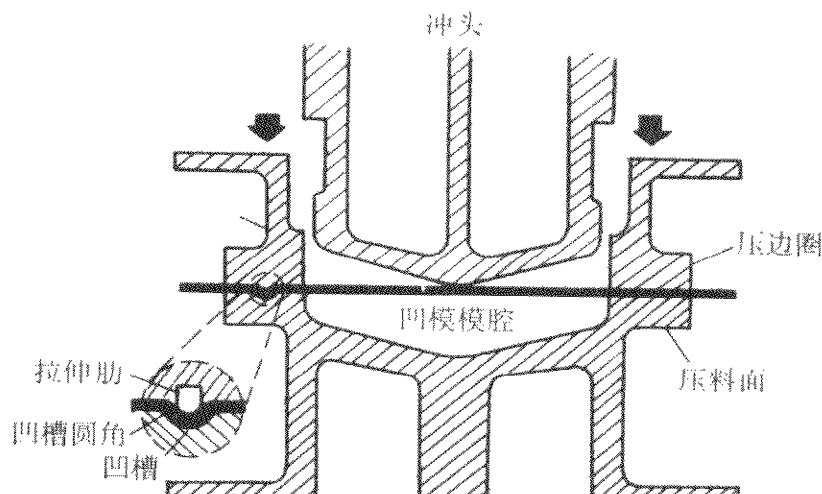
图 4.7-21 毛坯尺寸优化后的数值模拟结果
（厚度相对变化量）

6.2 板料拉深成形中的拉深肋设计优化

板料拉深成形过程中，常常存在变形的不均匀性，如盒形件在拉深成形过程中，直边部分的金属的流动要比圆角部分快，如果不加以控制，常产生起皱缺陷，如图4.7-22a所示。因此常在直边附近设置拉深肋，对金属流动的不均匀性加以限制，如图4.7-22b。拉深肋的作用主要是使得板料在流过拉深肋部位时经历弯曲和反弯曲的作用，从而对板料的流动产生阻力。不同的拉深肋尺寸对应不同的拉深肋阻力。拉深肋的阻力过大容易使得板料产生颈缩和拉裂，阻力过小则起不到作用、无法避免起皱。拉深肋设置属于工程优化问题。此处给出基于一步模拟方法的拉深肋阻力优化的方法。



(a) 带缺陷的盒形件



(b) 拉深肋设置

图4.7-22 盒形件拉深缺陷和拉深筋设置

一步模拟（One Step Simulation）又叫作反向法（Inverse Approach）或反向模拟，是由郭迎乔和J.L.Batoz等人提出的基于形变理论的数值计算方法。其基本思想为：从工件形状C出发，将其作为变形终了时工件的中面，对其离散，通过有限元方法确定在满足一定边界条件下工件中各个节点P在初始平板毛坯C^0中的位置P^0，然后根据位移量可以计算应变和厚度分布，如图4.7-23所示。通过有限元法离散，可以建立下式：

$$\{R(U)\} = \sum_e(\{F^e_{ext}\} - \{F^e_{int}\}) = 0 \qquad (4.7\text{-}19)$$

式中，$\{R(U)\}$为残余力向量；U为节点位移，mm；$\{F^e_{int}\}$为单元内力向量；$\{F^e_{ext}\}$为单元外力向量。

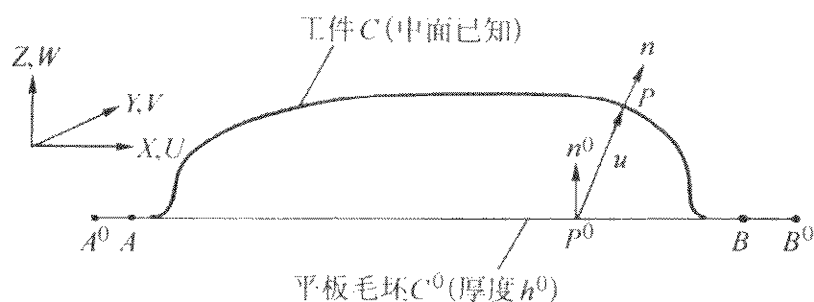


图4.7-23 一步模拟方法示意图

一步模拟因其有限元建模简单、分析速度快而成为设计初期工艺可成形性分析的有效工具。一步模拟主要包括两个计算过程：第一，通过将成形件上划分的网格反推出变形前的平面网格，确定出毛坯的形状；第二，通过将下料毛坯上划分的网格计算出零件上的畸变网格，确定出各质点的变形程度。利用第一个计算过程可以比较准确地确定出下料毛坯的形状和尺寸，而利用第二个计算过程可以快速确定出各点的变形程度，以初步判断是否产生成形缺陷。

由于拉深肋部位的几何尺寸比较小，所以如果将流经拉深肋部位的板料划分成较细的网格，则数值计算需要花费较多时间。因此常将拉深肋的作用等效为拉深肋阻力，拉深肋阻力的优化主要包括拉深肋的位置和拉深肋设计参数的优化。拉深肋的设置必须确保不产生颈缩，优化目标函数可以表示为：

$$J = \sum_e J^e = \sum_e \left|\frac{t^e - t^0}{t^0}\right|^p = \sum_e |\lambda^e - 1|^p \qquad (4.7\text{-}20)$$

式中，t^e为单元e内板料的最终厚度，mm；t^0为板料的初始厚度，mm；λ^e为板料的厚向延伸率；p为强调厚度变化影响的指数。

首先假定沿模具型腔的连续曲线C为拉深肋阻力的作用位置，选择几个关键位置的拉深肋阻力值为设计变量，其他位置的拉深肋阻力采用线性插值的方法。优化的目标首先是确定在哪些需要拉深肋阻力比较大的地方布置拉深肋。优化方法可以采用BFGS法或者序列二次规划法。

采用一步模拟可以进行灵敏度分析，目标函数J关于设计变量的偏导数为：

$$\frac{\mathrm{d}J[q,U(q)]}{\mathrm{d}q_i} = \frac{\partial J}{\partial q_i} + \left\{\frac{\partial J}{\partial U}\right\}\left\{\frac{\partial U}{\partial q_i}\right\} = \{M\}\left\{\frac{\partial R}{\partial q_i}\right\} \qquad (4.7\text{-}21)$$

式中，q_i为节点i的力，N；$\frac{\partial J}{\partial q_i}=0$；$\{M\}$为与力的平衡方程有关的伴随矩阵状态向量。

$\{M\}$可由下式求得：

$$[K]\{M\} = -\left\{\frac{\partial J}{\partial U}\right\} \qquad (4.7\text{-}22)$$

式中，$[K]$为切向矩阵。

如果有约束$g_i(q,U)$，则其关于q的导数的表达式类似于式（4.7-20）。

为了减少设计变量，拉深肋阻力作用线上21个节点的作用力定义为设计变量，其他点的力采用插值的方法求得。21个节点的作用力的初始值取100 N，经过8次迭代后收敛。各节点拉深肋阻力如图4.7-24所示，板料的变厚量已经得到大大的降低。根据初步优化的结果，可以在需要拉深肋阻力最大的两个位置设置拉深肋。

6.3 板料拉深成形压边力优化

复杂形状的零件在拉深成形过程中，不同位置的受力状态可能是不一样的，例如，为了避免产生起皱缺陷的出现，常通过压边圈向尚未进入凹模的板料法兰部分施加一定的压边力。如果施加恒定的压边力，虽然有些位置的起皱缺陷可以避免，但有可能在其他拉应力比较大的部位产生过度变薄

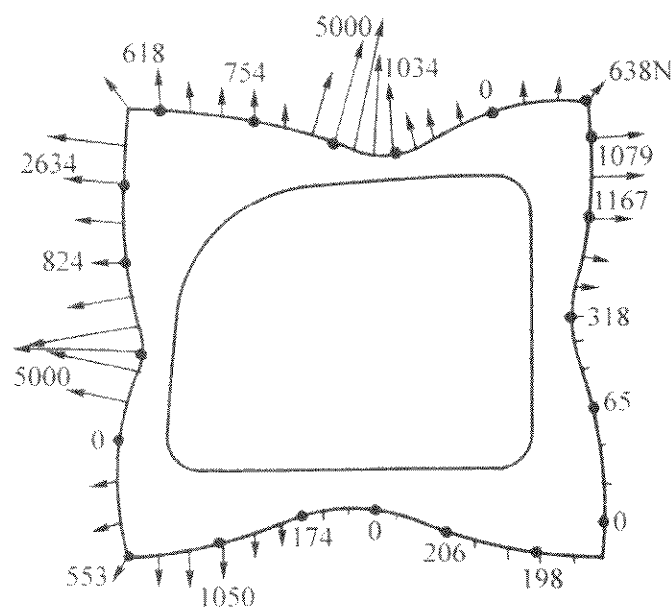


图4.7-24　优化后沿拉深肋线的拉深肋阻力分布

或拉裂。因此在不必要的位置施加过大的压边力反而导致无法充分发挥板料的拉深性能。

因此，压边力优化的最终目标就是获得随行程变化的压边力。

采用自适应或者优化的压边力控制策略，可使得板料的成形性能得到充分发挥，如图4.7-25所示，与采用刚性压边圈相比，采用柔性压边圈时板料安全成形的设计空间大大增加。

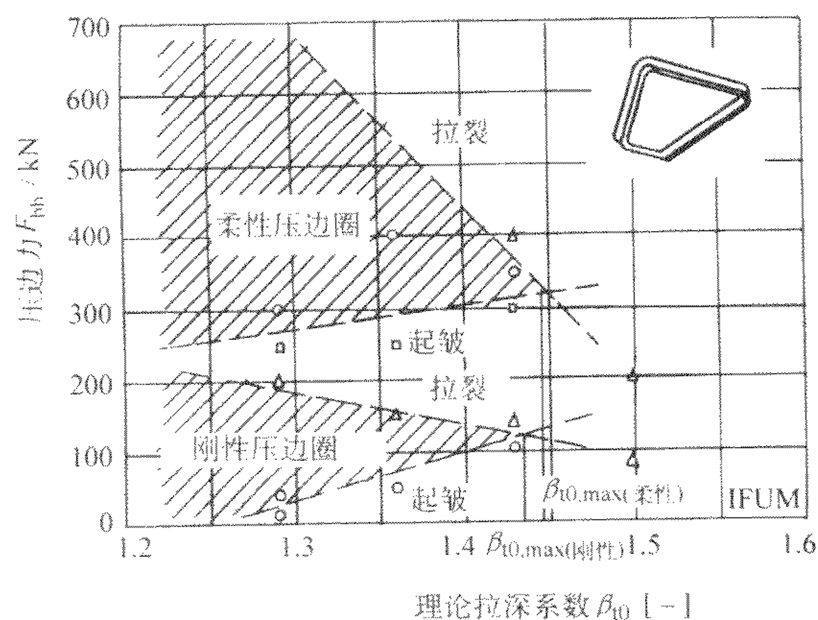


图4.7-25　材料H340的可加工区间

6.4　基于正交试验和增量有限元数值模拟的工艺方案优化

金属板料拉深成形过程中，板料的成形质量受多种因素影响，如板材的性能参数和模具工艺参数等。传统的设计过程要使得工艺和模具设计保证生产出满足质量要求的冲压件，除了依靠专家的经验，还要进行反复的试验（如正交试验）。有限元数值模拟技术的成熟，使得设计人员可以在计算机上通过改变设计参数，考察金属的变形规律，获得最佳的解决方案。这里给出基于正交试验和有限元数值模拟的冲压成形设计参数优化的方法。

研究对象为一盒形件，几何形状如图4.7-26所示。盒形件深25 mm，初始料厚t_0为0.8 mm，长度方向斜度为30°，宽度方向斜度为15°，采用一次拉深成形，宽边两侧的拉深肋依据设计经验确定。采用正交试验，采用L_{16}（4^4），即屈服强度σ_s、各向异性指数r、应变硬化指数n和摩擦系数μ，不考虑各因子的交互作用，各个因子及其水平如表4.7-5所示，冲压成形过程数值模拟采用Autoform软件。考核目标为没有开裂和壁部起皱情况下的冲压件的变薄程度，即研究最薄处的厚向应变：

$$\varepsilon_t = -\ln(t_{min}/t_0) \tag{4.7-23}$$

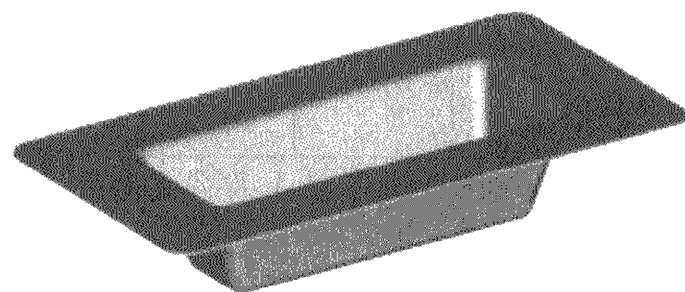


图4.7-26　盒形冲压件示意图

表4.7-5　各因子及其水平数

水平	σ_s/MPa	r	n	μ
1	110	2.1	0.19	0.05
2	120	2.3	0.22	0.08
3	130	2.5	0.25	0.11
4	140	2.7	0.28	0.14

如表4.7-6所示为数值模拟试验计划安排、试验结果和数据处理结果。图4.7-27所示为各因子和水平对成形质量的影响，即$R_D > R_B > R_A > R_C$，各因子的最优组合为：A1B4C1D1。但是，在生产实践中，如果σ_s、μ过低，r过大，常导致板料拉深成形起皱严重。因此，具体的优化方案还要进一步根据工程约束来确定。

表4.7-6　数值模拟试验方案、试验结果和数据处理结果

数值模拟试验方案					试验结果	
试验号	σ_s/MPa	r	n	μ	t_{min}/mm	ε_t/%
1	1	1	1	1	0.72	10.536
2	1	2	2	2	0.719	10.675
3	1	3	3	3	0.72	10.536
4	1	4	4	4	0.715	11.233
5	2	1	2	3	0.703	12.925
6	2	2	1	4	0.7	13.353
7	2	3	4	1	0.727	9.569
8	2	4	3	2	0.726	9.706
9	3	1	3	4	0.682	15.958
10	3	2	4	3	0.705	12.641
11	3	3	1	2	0.721	10.397
12	3	4	2	1	0.724	9.982
13	4	1	4	2	0.689	14.937
14	4	2	3	1	0.719	10.675
15	4	3	2	4	0.694	14.214
16	4	4	1	3	0.713	11.513
试验数据处理分析						
K_1	42.98	54.35	45.799	40.762		
K_2	45.553	47.344	47.796	45.715		
K_3	48.798	44.716	46.875	47.615		
K_4	51.339	42.434	48.38	54.758		
R	2.09	2.98	0.646	3.499		
排序	3	2	4	1		

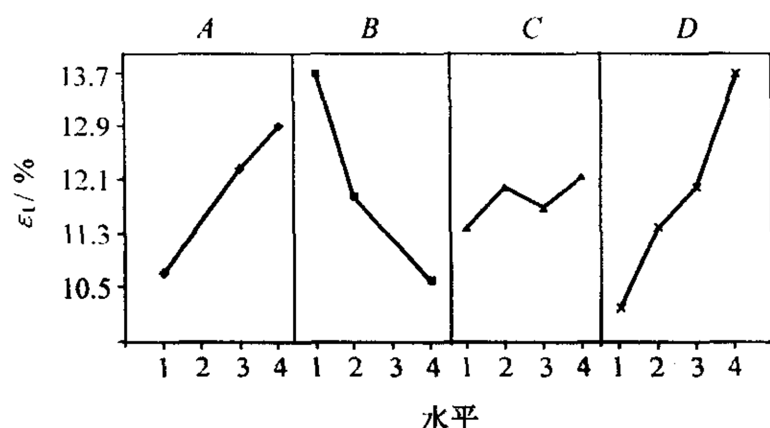


图 4.7-27 各因子和水平对成形质量的影响

7 激光板料成形过程模拟与优化设计

7.1 激光板料成形过程的数值模拟

板料激光弯曲成形的根源在于不均匀加热导致的温度梯度，即温度的变化使板料产生变形。实际上，板料的变形也会反过来改变热边界条件，进而影响温度的变化。也就是说，变形与温度两种不同场变量之间存在很强的相互作用，因此板料激光弯曲成形过程属于复杂的热力耦合问题。变形对温度的反作用往往表现在以下两个方面。

1) 经历大变形后物体几何形状发生变化，单元体积或边界面积也随之改变，施加在这些单元上的热边界条件也因此变化。

2) 非弹性功耗散转换成热。如绝大部分的塑性变形功率都会转化成体积热流，几乎全部摩擦力的功率也不可逆地转化成热的情形。

对于上述温度与位移存在强耦合作用的问题，可以采用先分析温度，后分析热应力的解耦方法分析，但显然会带来较大的分析误差。鉴于此，在板料激光弯曲成形的仿真分析中采用三维的热力耦合分析模型，同时处理热传导和力平衡两类不同场方程。

同时，板料的激光弯曲成形过程还有其自身的诸多特点，在进行有限元仿真时必须针对性地采取相应的措施，以与实际过程吻合，提高仿真结果的精度与可靠性。

7.1.1 板料激光弯曲成形仿真中模型处理

(1) 单元选取与处理

由于在激光弯曲成形中，板料局部的温度及变形发生剧变，存在强烈的温度与应力梯度，八节点六面体单元在解决这类问题时可达到较高的精度，因此在仿真时均采用八节点六面体单元。

八节点六面体单元按数值积分方案又分为全积分单元和减缩积分单元。在薄板的弯曲问题中，如果选择全积分有可能因精度损失过大而得出不合理的解。此时，采用减缩积分单元，不仅有助于降低单元集成的计算量，提高计算效率，而且可以去除刚度过分刚硬的不完全高解项的影响，提高计算精度。

同时加热区强烈的温度与应力梯度也要求在对板料进行离散时，加热区及其邻近区域应尽量细分。尤其是在板料的厚度方向应具有足够的积分点，实践证明，在板料厚度方向分为四层单元比较合适。为了提高求解效率，远离加热区的区域在单元剖分时可分得较为粗大。实际计算表明，远离加热区，材料的温度及变形变化不大，因此粗大的单元网格并不会显著影响所在区域，甚至整个求解域的精度。

(2) 材料性能参数的取值

温度对材料应力和变形的影响不仅仅在于热应变和热应力，材料的性能参数（热物理参数及力学参数）由于随温度的改变而剧烈变化，也会对应力分析的结果产生较大的影响，所以材料性能的正确取值在板料激光弯曲成形分析中至关重要。在仿真时必须充分考虑各参数随温度的真实变化情况，否则将会带来很大的分析误差。本节以08F钢板为研究对象，图4.7-28～图4.7-32分别给出了其热导率、比热容、弹性模量、屈服强度、线胀系数、泊松比随温度的变化规律，除此之外还考虑了密度与温度的相关性，密度ρ(g/cm³)随温度的变化由下式决定：

$$\rho = 7.85 - 2.625 \times 0.0001T \qquad (4.7\text{-}24)$$

(3) 热传导边界条件与初始条件的处理

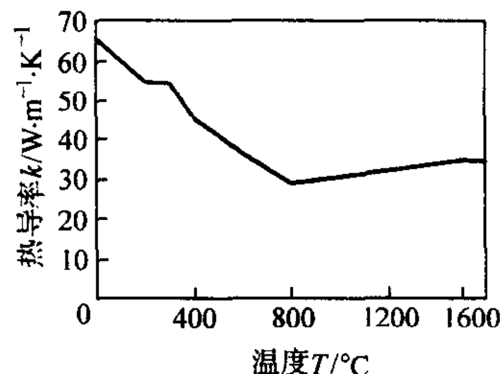


图 4.7-28 热导率与温度的关系

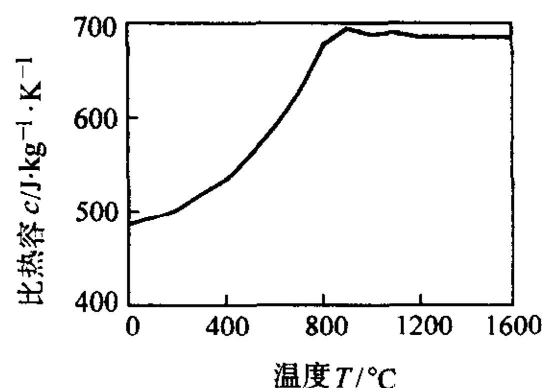


图 4.7-29 比热容与温度的关系图

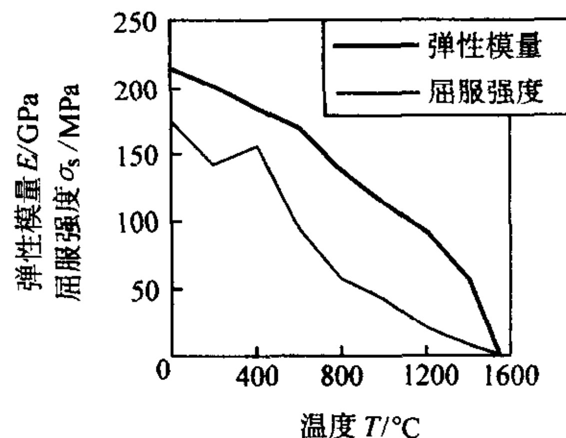


图 4.7-30 弹性模量、屈服强度与温度的关系

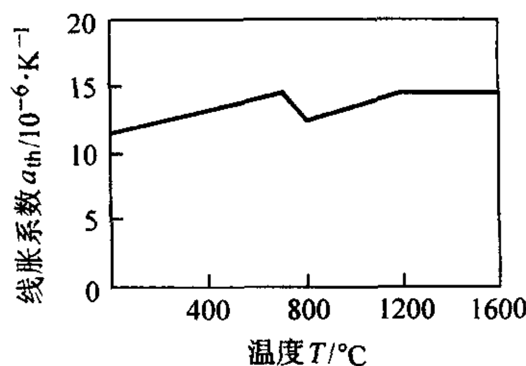


图 4.7-31 线胀系数与温度的关系

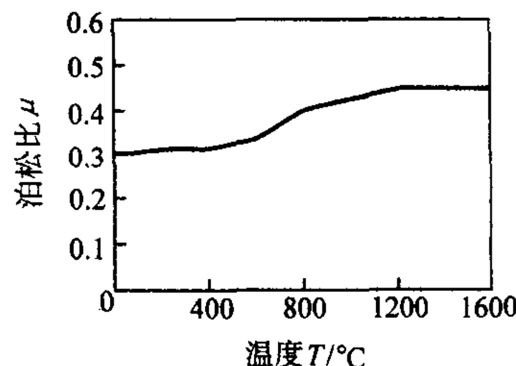


图 4.7-32 泊松比与温度的关系

在结构外表面 Γ_T 上通常可建立四类不同的边界条件：

1）表面 Γ_T 上给定温度：

$$T = T(x_i, t) \quad (4.7\text{-}25)$$

这一温度边界可能随空间和时间的不同而变化。

2）在 Γ_q 边界上给定热流强度的边界条件

$$q(x_i, t) = \lambda_n \frac{\partial T}{\partial n} \quad (4.7\text{-}26)$$

式中，n 表示 Γ_q 的外法线方向；λ 为不同方向的导热系数；$q(x_i, t)$ 是随空间位置和时间变化的给定热流强度值。

3）在 Γ 边界上给定对流的边界条件

$$q = h(T_\Gamma - T_\infty) \quad (4.7\text{-}27)$$

式中，h 为表面对流放热系数；T_Γ 为表面温度；T_∞ 为外界环境温度。对流放热系数 h 通常是环境介质的温度及其他材料特性和结构尺寸的复杂函数。

4）在 Γ 边界上给定热辐射边界

$$q = \sigma\varepsilon(T_\Gamma^4 - T_\infty^4) \quad (4.7\text{-}28)$$

式中，σ 为玻耳兹曼常数，$\sigma = 5.67\times10^{-8}$ W/（m^2·K）；ε 为材料表面热辐射率。

除了上述边界条件外，还需给定初始时刻的结构温度分布，作为初始条件：

$$T(x_i, 0) = T_0 \quad (4.7\text{-}29)$$

在板料激光弯曲成形中，板料局部加热后有两种冷却方式——强制冷却和自然冷却，这里只考虑板料经激光束扫描后在空气中自然冷却的情况。这时板料与周围环境存在对流和辐射换热。此两类边界条件在传热学中称为第三类边界条件，可将式（4.7-27）与式（4.7-28）统一写为下式

$$q = -\lambda\left(\frac{\partial T}{\partial n}\right) = a\ (T_\Gamma - T_\infty) \quad (4.7\text{-}30)$$

式中，a 为传热系数。传热系数 a 可写成表面传热系数 h 与等效辐射传热系数 h_r 之和。与环境的表面传热系数，结合实验现场条件，取为 30 W/（m^2·K）。与环境之间的辐射传热系数，根据辐射定律，并转换成式（4.7-27）的标准形式，可得等效辐射传热系数为

$$h_r = \sigma\varepsilon\ (T_\Gamma + T_\infty)\ (T_\Gamma^2 + T_\infty^2) \quad (4.7\text{-}31)$$

材料表面热辐射率 ε 取为0.8。显然，等效辐射传热系数与温度相关，因而辐射边界条件是非线性的。根据式（4.7-31），即可建立等效辐射传热系数随温度变化的曲线（如图4.7-33所示），作为仿真时等效辐射传热系数的取值。

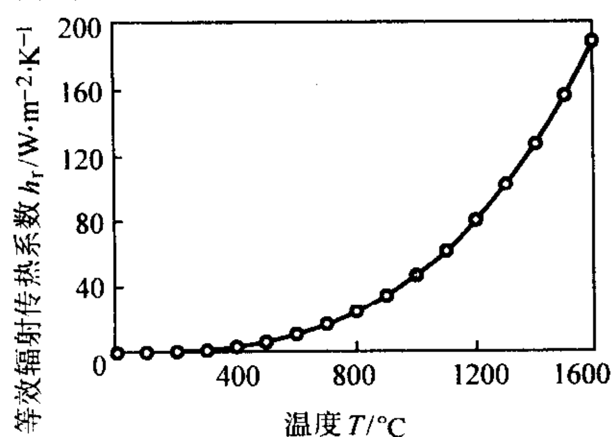


图4.7-33 等效辐射传热系数与温度的关系

此外，仿真过程中认为板料的初始温度与环境温度相同，均为常温，即 $T_\Gamma = T_0 = 20$℃。

（4）板料吸收系数的处理

板料吸收系数的大小取决于材料对激光的吸收率和板料的表面状况等。

激光束照射金属表面时，金属通过光子、自由电子、晶格三者之间的相互作用实现对光能的吸收，材料的吸收系数不仅与材料的物理性质、化学成分及表面状况有关，还与激光波长、光强度、作用时间以及材料的电阻率有密切关系。而且材料对激光的吸收系数还随温度升高而增大。同时由于温度的升高，材料表面氧化也会增大吸收率。材料的表面状况主要指板料表面的处理状况，为提高吸收系数，一般要对材料表面进行黑化处理，所采用的处理方法和涂层材料的不同对吸收系数也会产生显著的影响。可见，材料对激光的吸收系数在加工过程中的变化异常复杂，很难全面而准确地加以反映，因此在仿真中可不考虑材料吸收系数的变化，认为吸收系数为一常数。

根据板料表面的处理状况，以及综合考虑各因素的影响，仿真时板料对激光能量的吸收系数 A 取为0.5～0.8。

（5）激光热源的处理

激光束的模式很多，光斑形状有圆形、方形、矩形等，光束的能量密度分布也存在均匀型、高斯分布型等。这里认为光束能量密度为高斯分布的圆形光斑。为方便边界条件的施加，同时又不改变光束单位时间内注入板料的热量的前提下，将圆形光斑按照面积等效的原则转化为方形光斑来处理。

在板料激光弯曲成形中，激光束呈运动态，即边界热流的作用位置随时间的变化而改变，不仅如此，就板料表面的光斑而言，其热流密度的分布并不均匀，但由于激光光斑的尺寸很小（圆形光斑的直径一般在10 mm以下），同时板料的长度与宽度远远大于光斑尺寸，因此可将光斑作为一匀强面热源来处理，其热流密度为光束的平均热流密度 I_m。

对于高斯光束，其能量密度分布 I 为：

$$I = \frac{2AP}{\pi R^2}\exp\left(-\frac{2r^2}{R^2}\right) \quad (4.7\text{-}32)$$

于是，在光束直径范围内的平均热流密度 I_m 为：

$$I_m = \frac{1}{\pi R^2}\int_0^R 2\pi r I \mathrm{d}r = \frac{0.865AP}{\pi R^2} \quad (4.7\text{-}33)$$

式中，R 为照射到板料表面的光束半径；r 为考察点至光束中心的距离。

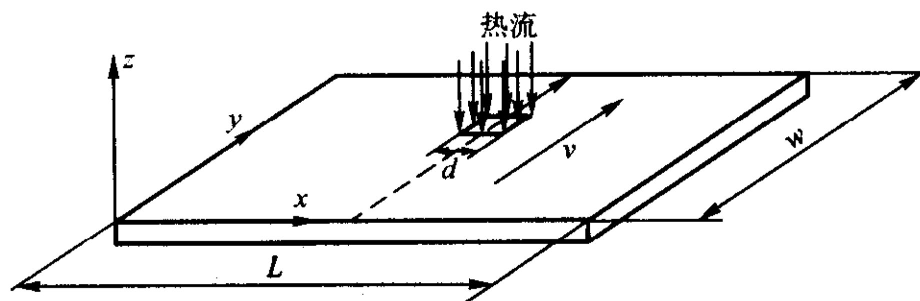


图4.7-34 激光热源处理模型

因此，为准确模拟处于运动态的光束热源，可将其简化为一平面尺寸为 $d\times d$、热流密度为 I_m、沿板料表面以扫描速度 v 移动的方形热源，简化模型如图4.7-34所示，简化后的激光束可作为一个移动的热载荷，以热流的形式施加到相应的单元表面。

（6）其他处理

在激光加热过程中，光斑周围材料的温度急剧升高，而在已扫描过的区域，材料温度急剧降低，因此分析时时间步长应取得足够小。但当光斑移出板料表面后，板料完全进入冷却过程，板料只是通过热传导方式与周围环境进行辐射和对流散热。与加热过程相比，冷却过程中温度与应力的变化较为平缓，此时则可以采用较大的时间步长，以提高分析效率。此外，时间步长的确定还应考虑光束的输出功率、板料对激光的吸收系数、光斑直径以及扫描速度等工艺参数。譬如在功率很高、光斑直径小、扫描速度慢的情况下，由于温度与应力的变化更加剧烈，时间步长应相应变小。

同时在时间步长控制时，根据成形参数的不同选用两种不同的控制方式：①固定的等时间步长。步长太小，可能使整个加载历程的完成需要不必要的过多增量步分析；太大的加载步长，造成一些增量步内非线性程度过高，出现收敛困难甚至发散。②自适应时间步长。时间步长的大小由每个增

量步内满足误差的非线性迭代次数与设定的目标收敛次数的偏离程度决定，对于与设定的目标收敛次数相比收敛较快的步长，可以按一定比例增加下一个步长的大小，反之则需减小下一个步长。这样由程序自动地确定出能够保证足够快的收敛性的较大步长。

位移边界条件的确定与施加根据不同的仿真过程而定，应针对具体问题，在模型建立时具体对待。

7.1.2　板料激光弯曲成形工艺仿真

(1) 模型处理及仿真工艺参数的选取

建立如图4.7-35所示的有限元计算模型，单元采用8节点六面体单元，共划分3 402个节点，2 600个单元。将板料一端（长度方向）所有节点的平动自由度予以约束，相关工艺参数如下：

光束功率：1 000 W

扫描速度：25 mm/s

光斑直径：4 mm

吸收系数：0.5

板料尺寸：40 mm × 20 mm × 2 mm

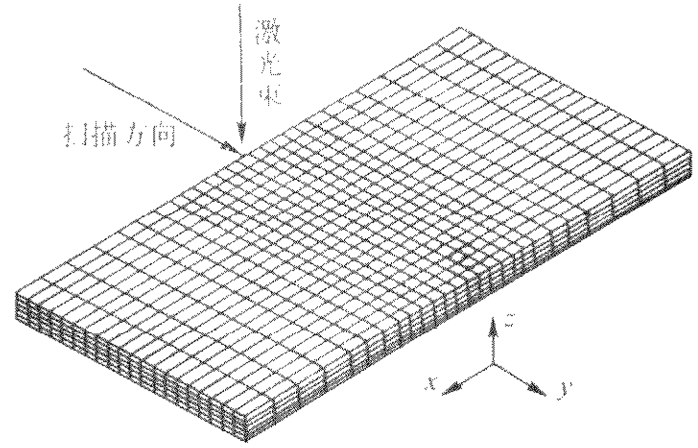
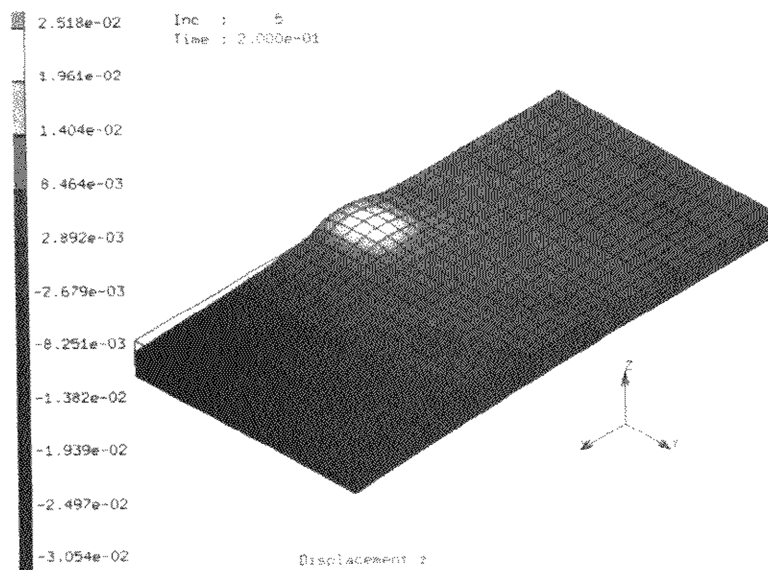


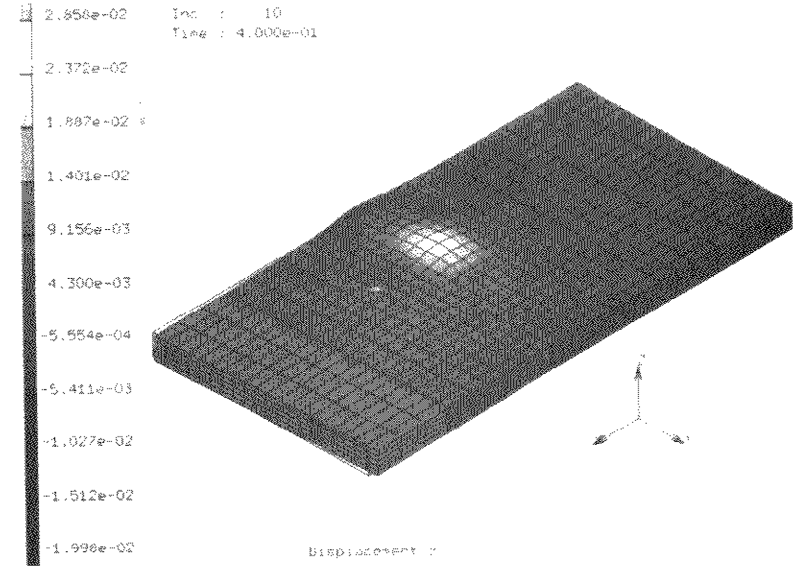
图4.7-35　有限元计算模型

(2) 成形过程仿真结果

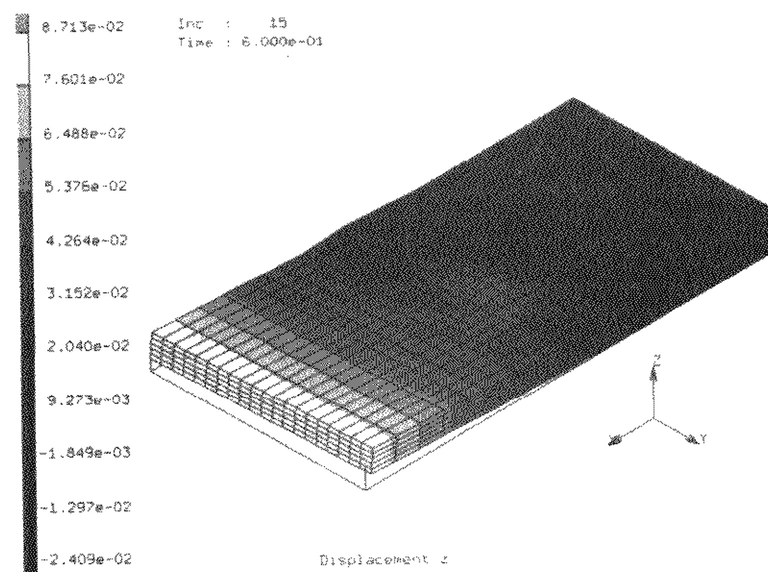
利用所建立的有限元模型，实现了特定工艺参数下的激光弯曲成形过程的有限元仿真。得到了成形过程中板料的应力应变场、变形场、温度场、热流场等结果。由于篇幅所限，这里只给出部分计算结果。图4.7-36为激光弯曲成形过程中的z向变形变化情况。图4.7-37为成形过程中的温度



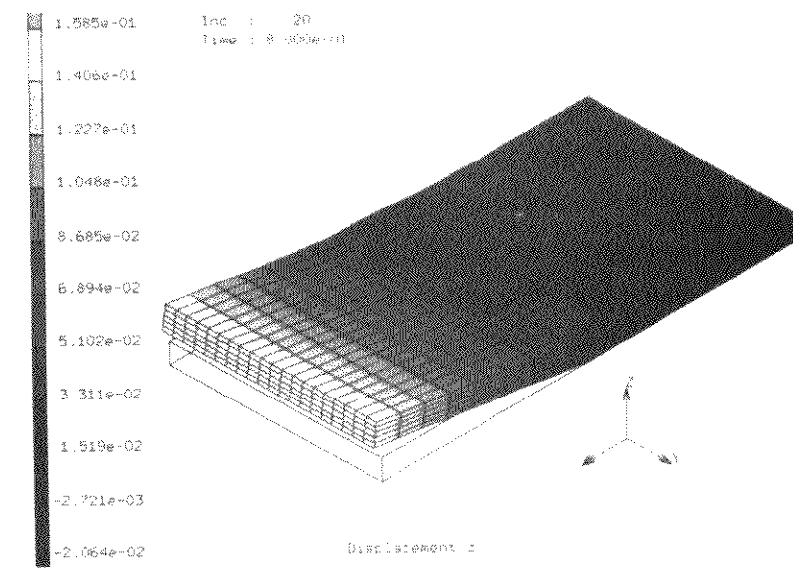
(a) t=0.2s



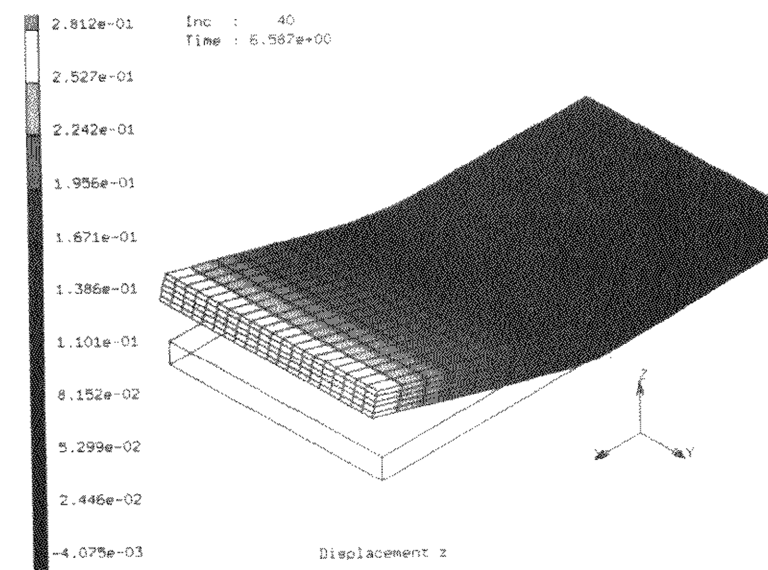
(b) t=0.4s



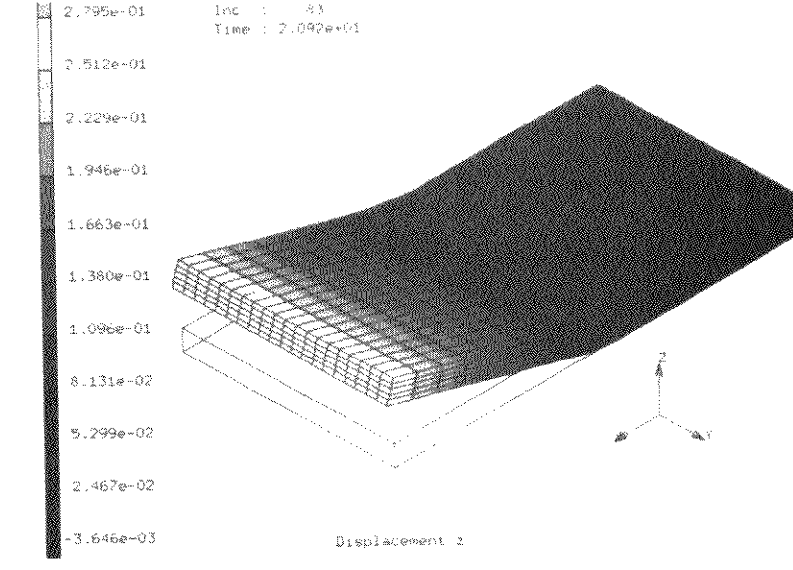
(c) t=0.6s



(d) t=0.8s



(e) t=6.587s



(f) t=20.92s

图4.7-36　z向变形（mm）过程（正向弯曲变形过程）

及横向（x 方向）的塑性应变的变化情况。可以看出，在激光弯曲成形过程中，激光束以一定的速度从板料的一侧向另一侧移动，对板料沿扫描线实施扫描，直至光斑完全移出板料表面，完成一次扫描过程。与扫描过程相对应，板料依次经历了加热和冷却两个过程，伴随着加热与冷却过程，板料表现出了相应的变形特征。

(a) t=0.4s　(d) t=0.4s

(b) t=0.8s　(e) t=0.8s

(c) t=20.92s　(f) t=20.92s

图 4.7-37　成形过程中的温度（℃）及横向（x 向）塑性应变 ε_x 变化

（a）～（c）为不同时刻的温度场变化；（d）～（f）为不同时刻的应变场变化

高能激光束沿板料表面扫描，当光束的能量密度较大而扫描速度又相对快时，在板料的厚度方向产生了非常强烈的温度梯度，致使加热区材料的塑性应变沿板厚方向也差别很大，结果当冷却阶段结束后，板料上表面的材料发生了较大的横向收缩变形，而下表面的横向收缩较小，甚至产生了伸长，从而导致了正向弯曲角度的产生。图 4.7-38 表示出了板料边缘的 z 向变形随时间的变化过程，可以看出，在变形的开始阶段（大约 t = 0.2 s 时），板料在 z 方向产生的最大位移为 − 0.03 mm，说明这时板料产生了反向弯曲变形，但随着冷却过程的继续，板料产生较大的正向弯曲变形。

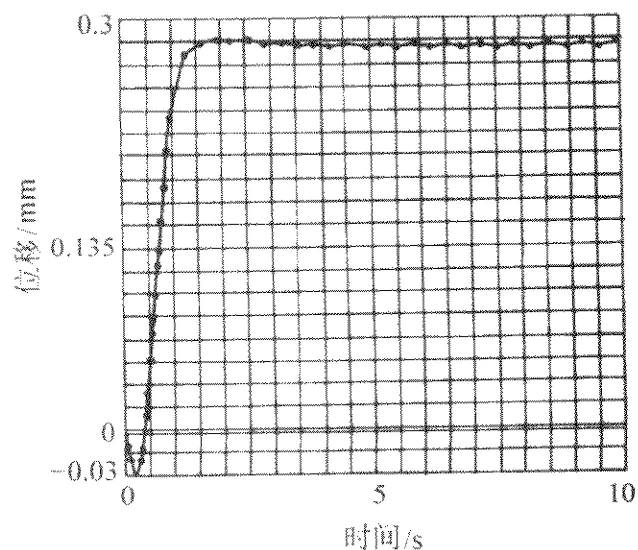


图 4.7-38　z 向变形随时间的变化过程

7.2　基于遗传算法的激光板料成形优化

遗传算法是一种全局优化方法，能够处理大规模优化问题，而且不会陷于局部最优，理论上可适用于任何优化目标。

而且遗传算法只需要目标函数值，不需要导数信息，工艺仿真模块可以采用不同的有限元软件来完成。

7.2.1 板料激光弯曲成形工艺优化模型的建立

(1) 设计变量的确定

影响激光成形的因素很多，但是对于某一具体成形工件来讲，板料的几何参数一定，可变的因素主要是激光扫描参数，包括激光功率 p、扫描速度 v、光斑直径 d 等，这些参数对激光弯曲成形的最终成形起决定性的作用，因此把它们作为激光成形工艺优化的设计变量。

(2) 约束条件的确定

激光成形工艺的约束条件并不多，主要是限制激光扫描过程中板料的峰值温度，不至于使板料熔化，或使组织和性能发生严重破坏，即

$$T_{\max}(p,d,v) \leqslant [T] \tag{4.7-34}$$

式中，$T_{\max}$为成形过程中板料加热区域所达到的温度峰值；$[T]$为加工后仍能使材料保持良好组织与性能所允许的最高温度。

(3) 设计变量的范围

设计变量的范围一般由具体的加工条件决定，即由进行加工的激光器及加工机床的技术指标决定。

(4) 目标函数的确定

在板料激光弯曲成形中，最首要的工艺要求是保证成形获得所要求的弯曲角度。在满足弯曲角度的成形精度的基础上，工艺要尽量合理，力求实现高效率、低能耗的加工。因此对应不同的工艺要求，目标函数的确定也就依具体情况而定。

1) 最大角度成形　由于激光一次扫描成形角度很小，所以激光弯曲成形一般需要多次扫描。为提高激光加工的效率，希望一次扫描能获得尽可能大的弯曲角度。即通过优化成形工艺参数组合，使一次成形的弯曲角度达到最大。该优化目标对应的目标函数为：

$$\max f = \alpha(p,d,v) \tag{4.7-35}$$

式中 α 为一组设计变量所对应的激光弯曲成形角度。

2) 定角度成形　这是工艺优化的主要目标，因为只有满足一定的精度要求，才能成形合格的工件，工艺才是有效的。要控制板料的弯曲角度在一次扫描过程中达到某一确定的定值，就要求定角度成形，该优化目标对应的目标函数为：

$$\min f = |\alpha(p,d,v) - \beta| \tag{4.7-36}$$

式中，β 为工艺要求的成形角度。随着优化计算的进行，目标函数逐渐逼近其理论最小值 0，在目标函数值小于预定精度时，$\alpha(p,d,v)$和 β 可以认为相等。因此，此时的设计变量对应的工艺参数就可以认为是能够在一定的精度要求下成形弯曲角度 β 的工艺参数。

3) 低能耗成形　节能是在任何时候都能被期望的，在激光成形中，能量的消耗由激光的输出功率 p 和扫描时间 t 决定，即：

$$W = pt \tag{4.7-37}$$

而扫描时间由扫描速度 v 和扫描路径的长度 l 决定，即：

$$t = l/v \tag{4.7-38}$$

于是输出功率为 p 的激光束以扫描速度 v 沿长度为 l 的扫描路径对金属板料进行扫描，其能量消耗为：

$$W = pl/v \tag{4.7-39}$$

式中，p/v 为激光的线能量密度，它表征了板料对激光的吸收率为 1 时，激光在单位长度上注入金属板料的能量大小。如果在相同的弯曲角度下，p/v 值越小，表明激光成形的能量消耗越低。综合激光成形的弯曲角度和能量消耗等因素，对于低能耗的目标函数为：

$$\min f = \alpha(p,d,v)\frac{v}{p} \tag{4.7-40}$$

(5) 优化系统的建立

激光成形工艺优化模型的目标函数中，都包含激光弯曲角度$\alpha(p,d,v)$项，由于激光加工的复杂性，目前激光弯曲角度的理论解析模型还很不完善，所提出的弯曲角度的解析表达式的求解精度无法满足优化求解的要求。

由于遗传算法只需要目标函数值，不需要导数信息，因此理论上可以适用于任何优化目标，只要此目标函数可以求出。所以在基于遗传算法的优化过程中，可以选择专门的有限元分析系统，如 MSC/MARC、ANSYS、LS－DYNA3D 等作为目标函数的解算器。图 4.7-39 为基于遗传算法的板料激光弯曲成形工艺优化系统的流程框图，系统由遗传算法和有限元成形过程模拟两个相对独立的模块组成。根据确定的初始工艺参数（激光功率 p、扫描速度 v、光斑直径 d），修改有限元前处理数据，进行成形过程仿真，通过对结果的后处理文件的相关数据进行处理，得到激光弯曲成形的角度，遗传算法模块根据该角度求得目标函数值，并进行遗传操作，返回优化后的工艺参数，如果不满足优化设计条件，则优化后的工艺参数→成形过程模拟→遗传优化→更优的结果…，如此循环往复，直到得到一组满足优化设计条件的工艺参数。

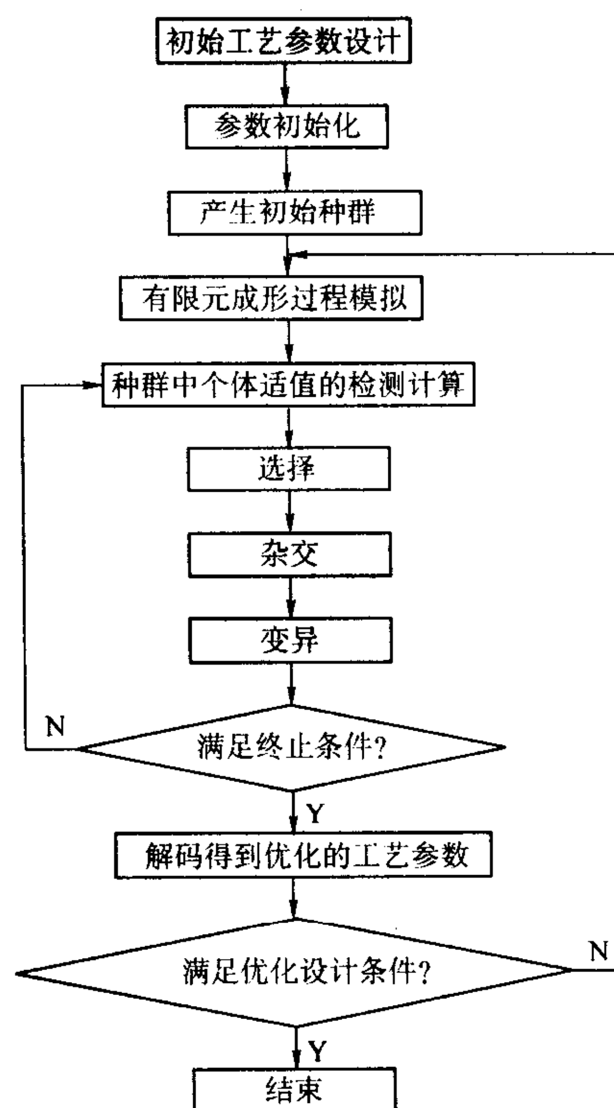


图 4.7-39　优化系统框图

需要注意的是在优化中适值函数的确定，适值函数是衡量染色体优劣的度量工具，直接影响到是否收敛和收敛速度，是遗传优化的重要因素。可以直接取目标函数为适值函数，但为提高收敛速度，一般要进行加工，使其在优化具体的操作中有适当的敏感度，在有约束的情况下应能作可行性调整，应尽可能满足以下三个原则：①与目标函数值一一对应；②适应值不论目标函数符号如何均应保持正值，且最好是单调的；③二者具有相同的极值点。根据激光成形的特点和所建立目标函数的变化情况，取目标函数的绝对值为适值函数比较理想。

7.2.2　板料激光弯曲成形工艺优化实例

应用遗传算法，对薄板激光弯曲成形的工艺参数进行优化。材料为08F钢板，钢板尺寸为长40 mm×宽20 mm×厚2 mm，将板料沿长度方向的一端约束，激光束的扫描路径在板料的中间位置（如图4.7-35所示）。

以定角度成形说明板料激光弯曲成形工艺优化的过程。假定工艺要求一次成形角度为0.8°，于是相应的优化模型为：

目标函数：$\min f = |\alpha(p,d,v) - 0.8|$

约束条件：$T_{max}(p,d,v) \leq 1\,000℃$

其中优化中材料承受的最高温度 T 由材料的性能要求和工艺要求决定，在此取1 000℃。根据激光成形的具体情况，各设计变量的取值范围如下：

激光功率 p：800～3 000 W

扫描速度 v：10～100 mm/s

光斑直径 d：1～10 mm

优化过程中的收敛过程如图4.7-40所示，可以看出，随着遗传代数的增加，弯曲角度逐渐逼近0.8°，说明优化正向着逐步逼近的方向搜索最优解。经过32代优化后得到的工艺参数为激光功率948.4 W、扫描速度30.5 mm/s、光斑直径3.33 mm，此时对应的板料成形角度为0.801 80°，误差小于工艺精度要求。

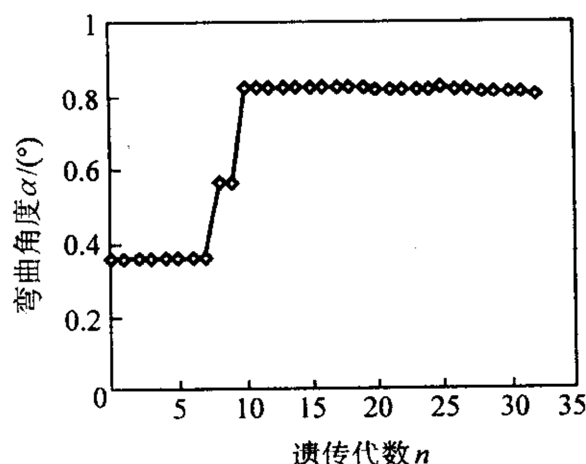


图4.7-40　优化收敛过程

8　多点成形技术工艺优化

8.1　多点成形技术

多点成形是利用一系列规则排列的、高度可调的基本体，通过对各基本体运动的实时控制，自由地构造出成形面，从而实现板材三维曲面成形的一种柔性加工技术。

多点成形技术涵盖的领域很广，包含了从多点成形原理、工艺、成形方法到多点成形设备的设计与制造等诸多方面。多点成形方法与传统模具成形方法的一个主要区别就是它具有“柔性”加工特点，即可以实时控制各基本体的位移和速度。利用这个特点，既可以在成形前也可在成形过程中改变基本体的相对位移状态，从而改变被成形件的变形路径及受力状态，以达到不同的成形结果。多点成形设备的这种柔性加工的特点与传统模具成形相比能为工件提供更多的变形路径的选择，从而能够实现如反复成形、分段成形等很多特色加工方法。

多点成形有二种典型的方式，一种称为多点模具成形，其特点是在板材成形前就调整好各基本体的位置，而在成形过程中基本体间没有相对运动。其实质与模具成形基本相同，只是把模具分成离散点。另一种称为多点压机成形，其特点是在成形前不对基本体进行调整，而在成形过程中调整基本体位置，此时的每对基本体都相当于一台小型压机，故称为多点压机成形法。多点压机成形更能体现多点成形的优点。目前，为适应不同的需求，多点成形已经形成了系列化的实用技术。

（1）反复成形技术

反复成形法就是利用多点成形柔性化的特点，在多点成形中成形件围绕着目标形状连续不间断地反复成形，逐渐靠近目标形状，减小工件的回弹及材料内部的残余应力，实现板材小回弹或无回弹成形。反复成形时，首先使材料变形到比目标形状加上应有的回弹值还大一点的程度，在此状态下，再使材料往回变形。如果此时的变形量等于回弹值，就相当于卸载过程，继续加载使材料沿其回弹方向继续变形到超过目标形状。这样，以目标形状为中心，重复上述成形过程，使板料逐渐地靠近到目标形状，最后在目标处结束成形。在多次反复成形过程中，可使残余应力的峰值逐渐变小，最终可实现无回弹、无残余应力的变形。

（2）分段成形技术

多点分段成形充分利用了多点成形设备的柔性特点，把工件在不分离的情况下分成若干个成形区域分别成形，从而能够实现利用小设备对大型板材的成形，如图4.7-41所示。

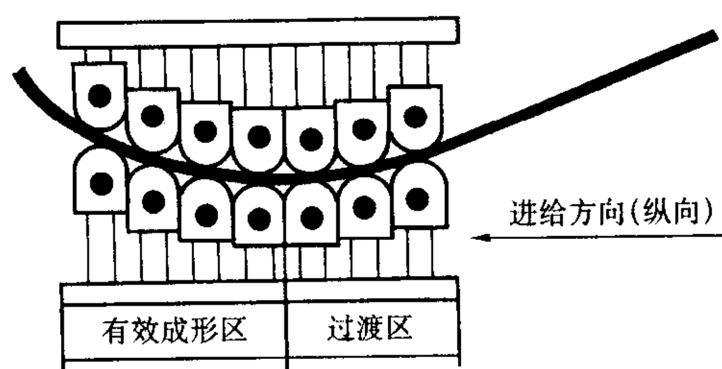


图4.7-41　分段成形示意图

这种成形方法可以减小设备尺寸，实现以往只能利用手工完成的大型板材的压制，从而大大降低产品的成本。同时，多点分段成形方法也可以提高板材的成形权限，但由于分段成形时，板材在每次成形时，都要受到未成形区及已成形区的影响，成形区的受力及变形情况比整体成形时要复杂得多，控制较难。

（3）多道成形技术

对于变形量很大的制品，选取最佳路径多道成形，使成形过程中板材各部分变形尽量均匀，以消除起皱等成形缺陷，提高板材的成形能力。

（4）闭环成形技术

即将自动控制技术与CAT、CAD结合起来，对成形后的工件进行三维测量，将测量的数据反馈到CAD系统，经过控制算法运算后，计算出基本体群形状的修正量，传递给控制系统再次成形，这样反复几次，可以达到精确的目标形状。

在上述几种成形技术中，分段成形时工件在不分离的情况下分成若干个成形区域分别成形，从而能够实现利用小设备对大型板材的成形。这种成形方法大大降低了设备的成形面积和吨位。甚至能够实现以往只能利用手工完成的大型和超大型板材（如船板、飞机外壳等）的压制，极大地提高了劳动生产率，降低了产品的成本。但是，分段成形时的受力及变形情况比整体成形时要复杂得多，控制较难。以往的研究表明，分段成形时的缺陷往往集中表现为出现在压制区与自由变形区交界上的波浪形曲面，这种缺陷是由于压制区受到自由变形区的抑制作用而产生的局部剧烈塑性变形而引起的。在此交界线附近设置一定的过渡区来分散这种局部剧烈塑性变形是解决此类问题的基本方法。因此，怎样设计过渡区的大小、形状才能既保证避免出现波浪形曲面的缺陷，又

保证获得最高的成形效率是分段成形技术中亟待解决的关键问题。实际上过渡区设计也可视为一种优化问题，优化目标在于寻找最佳过渡区的大小和形状，保证在不产生缺陷的同时获得最高的成形效率。李明哲教授提出了一种分段成形过渡区最优化设计的方法。

8.2 分段成形过渡区最优化问题描述

利用过渡区变形协调技术进行分段成形时，可以把成形区分成两个部分，一个是有效成形区，另一个是过渡区，如图4.7-41。在有效成形区内的冲头群要压制工件达到目标形状，在过渡区内的冲头群则要求能保证板材不发生局部剧烈塑变。可见有效成形区的大小决定了成形效率，有效成形区的排数越多，成形效率就越高。过渡区的排数越多，工件的变形协调性就越好，成形件的最终成形效果也就越好，但成形效率随之下降。过渡区的形状则是保证板材不发生局部剧烈塑变的关键因素，其形状的设定要协调目标形状和坯料形状，尽量使过渡区内变形均匀，避免出现局部剧烈塑变。

通过上述分析可以明确，优化目标在于合理确定过渡区的排数以及过渡区内的各冲头行程，使过渡区排数最小的同时，过渡区内的各冲头排布仍然满足不发生局部剧烈塑变的要求。

根据以往的研究，过渡区形状设计的基本思想是使处于过渡区变形的板材的曲率从已变形区到未变形区的变化最均匀，这时板材上将不会因过度局部变形产生缺陷，变形区间的衔接也会最平滑。即目标泛函

$$\min: \Phi[k(F)] = \iint_{R_t} \left\{ w_{x_1} \sum_{i=1}^{2}\sum_{j=1}^{2}\left[\frac{\partial^2 k_{x_1}(F)}{\partial x_i \partial x_j}\right]^2 + w_{x_2} \sum_{i=1}^{2}\sum_{j=1}^{2}\left[\frac{\partial^2 k_{x_2}(F)}{\partial x_i \partial x_j}\right]^2 + w_{x_1x_2} \sum_{i=1}^{2}\sum_{j=1}^{2}\left[\frac{\partial^2 k_{x_1x_2}(F)}{\partial x_i \partial x_j}\right]^2 \right\} \mathrm{d}x\mathrm{d}y \tag{4.7-41}$$

式中，k_{x_1}、k_{x_2}、$k_{x_1x_2}$分别为过渡区成形曲面的 x_1、x_2 方向的曲率及扭率；w_{x_1}、w_{x_2}、$w_{x_1x_2}$为权因子；F、R_t 分别为过渡区成形曲面及其面积。

求解上述泛函的极值便能得到过渡区成形曲面方程 F。

如果在过渡区内划分 $m\times n$ 的网格，设过渡区曲面对应每个网格点的高度为 d_{ij}，则过渡区成形曲面可以表示为 $F(s)$，其中

$$s=(d_{00},d_{01},\cdots,d_{11},d_{12},\cdots,d_{mn}) \tag{4.7-42}$$

所有可能的 s 构成的解空间可计为 S。则对于任一 $s\in S$，s 都代表了可能的过渡区形状。所要求的问题是找到一点 $s_{\max}\in S$，使$f(s_{\max})$是 S 上的全局极大值，即求一点 $s_{\max}\in S$，满足

$$\forall s\in S: f(s_{\max})\geqslant f(s) \tag{4.7-43}$$

函数$f(s)$是 s 的适应值函数，显然在此应取

$$f=K-\Phi \tag{4.7-44}$$

式中 K 为一常数，代表了函数 Φ 的上界。上式的写法保证了$f(s)\geqslant 0$，且较差的过渡区表面具有较低的适应值。

需要注意的是，在解空间 S 中，没有把过渡区的排数考虑进去，主要是因为过渡区排数的引入会引起解空间维数的改变，而遗传算法要求解空间维数是确定的。在确定过渡区排数的时候，可以从过渡区排数的较小值开始逐次递增测试，直到找到符合要求的形状为止。

8.3 算法描述

应用遗传算法进行过渡区优化设计的主要过程如下。

1）编码表示　采用浮点编码方案，直接把所求的变量 s 作为一个染色体，d_{ij}为基因。如式（4.7-42）所示。

2）确定适应值函数　适应值函数的选取直接影响最终的设计结果，是利用遗传算法进行过渡区设计的关键。确定适应值函数一方面要正确反映过渡区形状设计的意图，另一方面要尽量简化适应值函数，以便提高遗传算法的效率。

式（4.7-44）给出的是适应值函数$f(s)$理论上的表达式。但由于泛函 Φ 的复杂性，直接利用式（4.7-44）来求适应值几乎是不可能的，必须依据实际情况作必要的简化。

分段成形法按其分段方法的不同，可分为单向分段成形和双向分段成形。其中，单向分段成形在目前的研究中应用较多，是研究双向分段成形的基础。在此只探讨利用遗传算法进行单向分段成形的设计，对双向分段成形可参照处理。

试验表明，对于单向分段成形，进给方向的曲率调节起着主要作用。此时在式（4.7-41）中可取 $w_{x_1}=0$，$w_{x_2}=w_{x_1x_2}=1$，并可忽略混合导的影响。对式（4.7-41）进行离散化，用求和近似代替积分，用差分近似代替导数，均匀划分过渡区内的网格，此时的式（4.7-41）可以简化为

$$f(s)=K-\sum_{j=0}^{n}\sum_{i=0}^{m} w_1(k_{i+1,j}^{x_1}-k_{ij}^{x_1})^2-\sum_{j=0}^{n}\sum_{i=0}^{m} w_2(k_{i,j+1}^{x_1}-k_{ij}^{x_1})^2 \tag{4.7-45}$$

式中，$k_{ij}^{x_1}$ 为网格点（i，j）处过渡区曲面在进给方向的曲率。

3）确定遗传算子　遗传算子主要包括选择、交叉、变异3个主要算子。群体通过遗传算子的作用后，将会进化出下一代群体。

选择算子：采用轮盘赌方法进行选择，把当前群体中的个体按与适应值成比例的概率复制到新的群体中。对选择产生的新群体要经过交叉、变异操作后才能最终产生进化后的群体。

在上述几步完成后，即可反复求取每代群体各个个体的适应值，并利用遗传算子进化出下一代。这样重复进行进化，即可不断地逼近于理想的过渡区形状。

8.4 应用实例

下面给出利用上述遗传算法对球形及双曲抛物形（马鞍形）工件设计过渡区的两个实例。

算例所使用的多点成形设备主要工艺参数如下：冲头数为10×10，冲头间距为14 mm，冲头球头半径为10 mm。两个算例所用的坯料均为铝板，尺寸为280 mm×140 mm。板厚为3 mm。

马鞍形工件的双向目标曲率半径均为120 mm，过渡区排数为6排。遗传算法中主要参数为群体规模 $N=100$，杂交概率 $p_c=0.6$，变异概率 $p_m=0.005$。在进化100代后，优化结果趋向于收敛，共进化150代，耗时283 s。

图4.7-42为利用马鞍面分段成形的计算结果反算出的工件形状，由图4.7-42可知，利用遗传算法设计的过渡区形状基本符合预想的设计意图，过渡区部分的横纵方向曲率从原始的曲率半径 $R=120$ mm 的圆弧较均匀地向坯料形状过渡。

球形工件的目标曲率半径为180 mm。遗传算法中主要参数除群体规模 $N=200$ 外，其他参数与上例一致。在进化80代后，优化结果趋向于收敛，共进化150代，耗时356 s。图4.7-43所示为球形工件的分段成形设计结果。从图4.7-43中可以看出，利用遗传算法设计出的过渡区形状基本符合设

计要求。

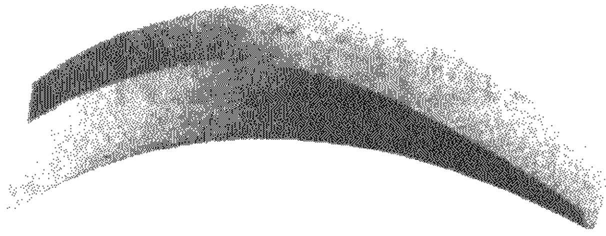


图4.7-42 马鞍面分段成形设计结果



图4.7-43 球面分段成形设计结果

经实际压制证明，采用上述过渡区设计方案，不但成功地抑制了局部剧烈塑变的产生，还可以提高成形效率。

编写：陈军（上海交通大学）
管延锦（山东大学）

第8章　塑料注射成形过程的优化设计

如何以更低的价格、更好的质量，及时地将产品推向市场、赢得竞争，是注射成形企业的主要目标。要使产品具有市场竞争力，产品质量起着关键作用。注射成形零件质量可以用力学性能、尺寸精度及外观等来评价。影响注射成形零件质量的因素很多，包括：成形方案（与成形有关的模具结构，主要包括浇注系统和冷却系统）、材料性能（主要包括聚合物材料的流变性能、物理性能、力学性能等）、成形过程中的工艺参数（主要包括注射温度、注射速度、保压压力、保压时间、冷却介质温度、冷却介质流速等众多变量）等。

1　成形方案的优化设计

1.1　浇注系统优化设计

设计合理的浇注系统，对注射成形而言十分重要，但很多时候设计人员并没有认识到这一点，因而也就没有给予足够的重视。大的流道中塑料熔体流动所需的压力较低，但需要较长的冷却时间，消耗和废弃更多的材料，需要更大的锁模力。设计适当的较小的流道系统将最大程度地在成形中提高原材料和能源的利用率。当然，流道尺寸的减小受注射机注射压力的限制。

浇注系统优化设计的目标之一就是要减小注射压力。通常，多个浇口可以减小流动的长度，因此就可以减小所需的注射压力，图4.8-1所示的实例显示出了这种关系。如果为了减小注射压力而选择多浇口系统，则可能会导致出现可见的熔接缝，并且需要复杂的流道系统。借助于注射成形模拟软件（CAE）可以分析浇口的位置与个数对注射压力的影响。模拟软件也能预测熔接缝和气穴位置，这些信息有利于选择最佳的浇口位置。

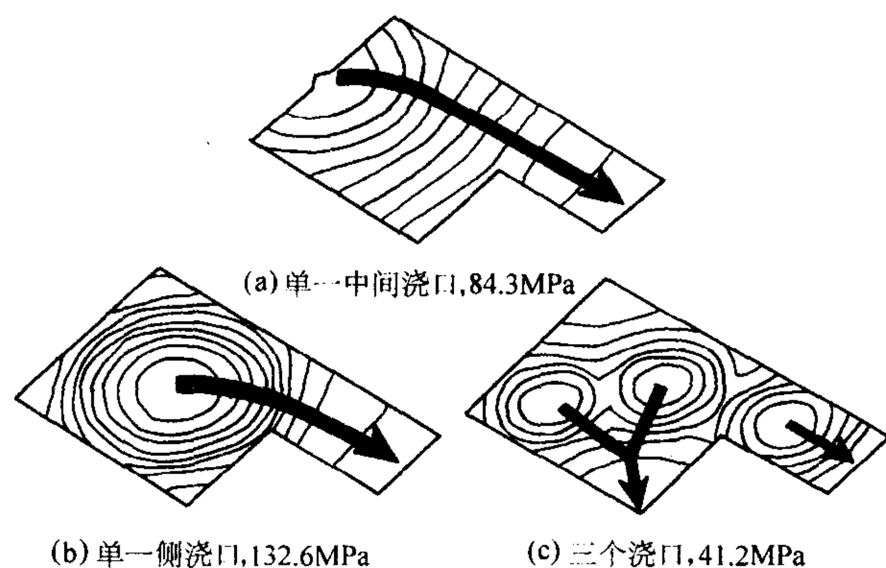


图4.8-1　注射压力与浇口位置、个数关系的实例

浇注系统优化设计的另一个目标就是要实现流动平衡。常用的多型腔模具的流道系统有三种：独立式（鱼骨形）流道系统、H桥（分支）式流道系统和辐射状（星形）流道系统，如图4.8-2所示。H桥式流道系统和辐射状流道系统是自然平衡的。所谓自然平衡流道是指对每个型腔而言，都有完全相同的流道路径与尺寸，所以每个型腔都在相同的条件下充填。虽然鱼骨形流道是非自然平衡的，但它可以适应更多的型腔，并且具有最小的流道体积和较低的制造成本。这一点要优于相应的自然平衡流道。非平衡流道系统可以通过改变流道直径和长度来达到人工平衡。

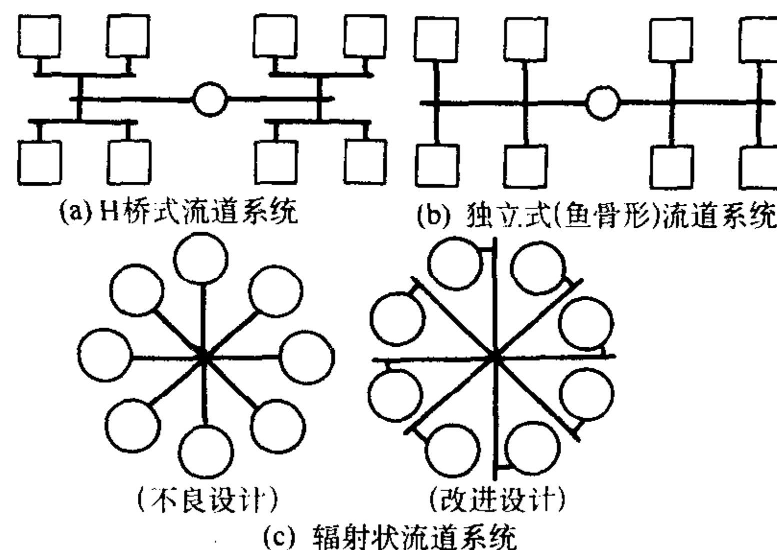


图4.8-2　基本流道系统布局

对高质量的塑件成形来说，型腔中的平衡流动是必要的。改变浇口直径能得到形式上的平衡充填，但却对浇口的凝固时间影响很大，从而导致塑件的不一致，同时浇口由于容易磨损，其尺寸不易得到保证。如果可能，应尽量采用自然平衡式流道系统，如果不能采用自然平衡式流道，那么就得进行人工平衡。用CAE来设计人工平衡流道系统，不仅省时省力，也大大提高了塑件充填的一致性。

人工平衡流道系统如果为某一种材料设计，当换用另一种材料时就不一定能平衡。另外，人工平衡流道系统需要更严格的工艺过程控制，工艺控制中的小变化也会影响模具的充填模式，导致不平衡充填。例如，如果使用独立式（鱼骨形）流道系统而使用不同的注射速率，结果会产生不同的充填模式。一般来讲，慢的注射速率将首先充填较远的型腔，而较快的注射速率将首先充填靠近浇口的型腔。因为在低速注射下，熔体在第一次碰到浇口时往往会迟滞，流向流道系统的其他部分。当所有的流道分支填满之后，上游的浇口会变得比下游浇口的阻力更大（因为该浇口处的熔体温度有所下降，甚至已开始凝固）。不同的注射速率会使充填模式介于这两种极端情况之间，如图4.8-3所示。在流道设计的最后阶段，CAE分析能帮助确定流道对于流动速率的敏感程度，并提供合适的加工窗口。

如前面所述，CAE分析在浇注系统设计中的主要作用可总结为：①计算最优的流道尺寸，以便使熔体以适当的速度输送并得到合理的压力降；②实现流道人工平衡；③预测不同工艺条件（如注射速率）对流道平衡的影响。

浇注系统的设计规则可归纳如下。

① 流道必须设计成有利于迅速充填型腔。

② 容易脱模且容易把浇口从塑件上取下。

③ 对多腔模系统，流道平衡对于保证塑件质量一致至关重要，流道系统可以自然或人工平衡。

④ 人工平衡流道可以通过改变流道尺寸与长度达到，改变浇口直径能达到表面上的平衡充填，但是会对浇口凝固时间有显著影响，不利于塑件质量的均匀一致。

⑤ 尽量采用小的流道，以使废料最少并产生摩擦热，在流道中产生摩擦热是在不提高料筒温度的情况下提高熔体温度的有效途径，因为过高的料筒温度会使材料降解。

⑥ 当流道有分支时，分支的直径应当小于主流道，假设 N 为流道分支数，则主流道与分支流道的直径有如下关系：

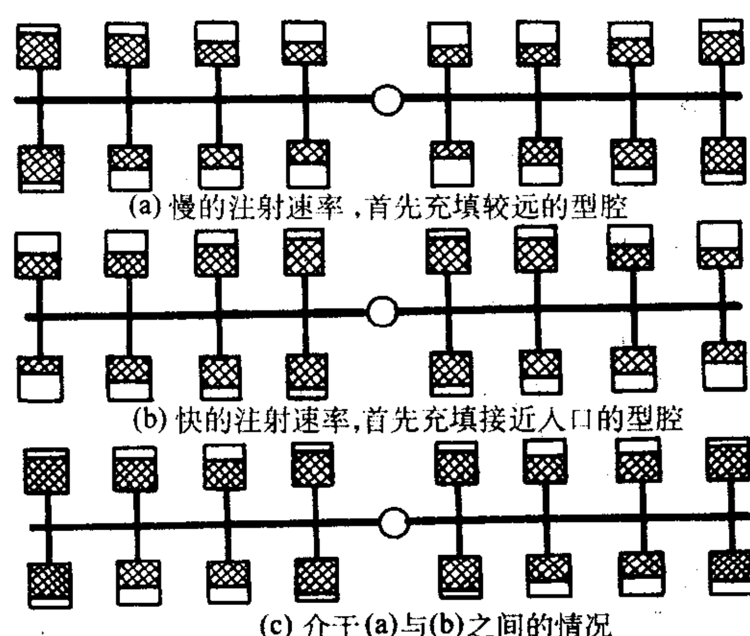


图 4.8-3　在非自然平衡的流道系统中不同的注射速率导致不同的充填模式

$$d_{主} = d_{分} N^{1/3} \tag{4.8-1}$$

⑦ 梯形流道深度应等于它的宽，每边应有 5°～15°的斜度。

⑧ 对大多数材料，推荐的最小流道直径为 1～5 mm。

⑨ 所有流道交叉点应有冷料井，以存储前锋冷料。冷料井的长度一般等于流道的直径。

1.2　冷却系统优化设计

如图 4.8-4 所示，注射模具冷却系统由以下几部分构成：温度控制单元、泵、供水歧管、软管、模具中的冷却管道以及一个集水歧管。模具本身可认为是一个热交换器，高温塑料熔体的热被循环冷却液带走。

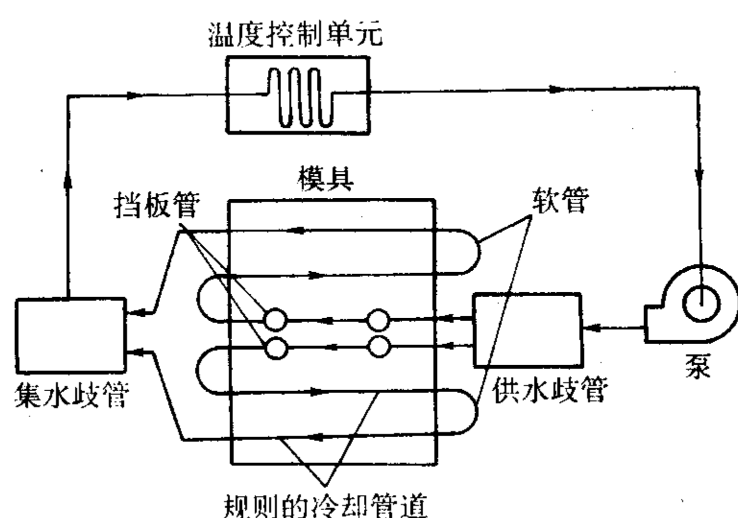


图 4.8-4　冷却系统的组成

冷却管道的配置可以是并联的，也可以是串联的，如图 4.8-5 所示。并联冷却管道从供水歧管直通到集水歧管，由于并联设计的流动特点，不同的冷却管道的流动速率是不同的，取决于每个冷却管道的流动阻力。这种变化的流动速率反过来导致了不同的冷却管道传热效率。因此，使用并联冷却管道配置的模具冷却效果并不一致。一般情况下，型腔和型芯部分都设置各自的并联冷却管道系统，每个系统冷却管道的数目根据模具的尺寸与复杂性而有所变化。

冷却管道从冷却介质入口到出口通过一个单回路连接起来，这称为串联冷却管道。这种类型的冷却管道配置是最常用和推荐使用的。如果各段冷却管道在尺寸上是一致的，冷却介质就可以在它的整个长度上维持较好的紊流速率（紊流能使传热传效果更好）。另外，还需注意应使冷却介质的温升最小（温度升高是因为冷却介质吸收了整个冷却管道中的所有热量），对于一般模具而言，出入口冷却介质的温差通

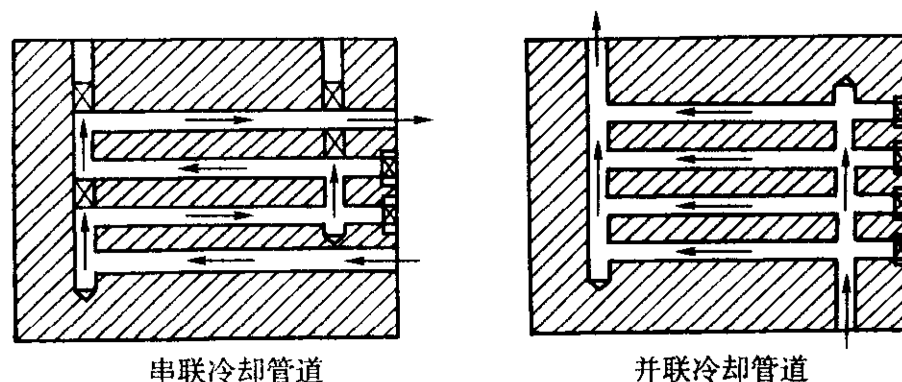


图 4.8-5　冷却管道配置

常应在 5℃之内。对于精密模具应在 3℃以内。对于大型的模具，可以考虑多个串行的冷却通道，以保证均匀的冷却介质温度和模具冷却的均匀。

冷却通道一般是在模具型腔和型芯部分钻孔形成的。但对于一些较深的型芯，需要进行足够的冷却模具而无法设置常规的冷却管道，这时应该考虑使用隔板、喷流管、热管等特殊冷却方式，如图 4.8-6 所示。

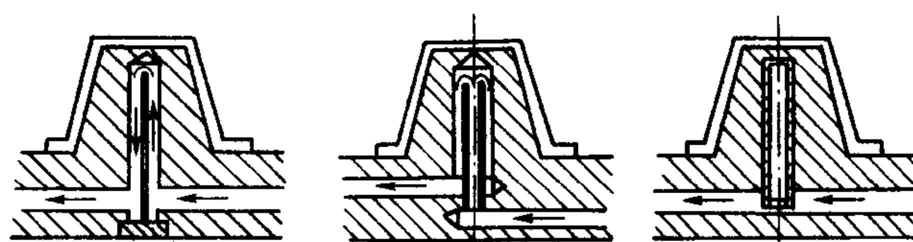


图 4.8-6　隔板、喷流管、热管

隔板实际上是在垂直于主冷却管路的方向上开一条冷却通道（其形状可以根据型芯的形状设计成圆形、矩形等），然后用一个隔板将该冷却通道分成两半，冷却介质从隔板一侧的主冷却管路流入，流过隔板顶部之后，又流回主冷却管路。

喷流管与隔板类似，将隔板用一个小的管道代替，冷却介质流入管的底部并从顶部“喷出”，就像喷泉一样，然后再从管外流下来继续在冷却管路中流动。

热管与前二者不同，它是一个充满液体的密封的圆柱，液体在接近型芯区域的一端从其外壳中（一般是钢材料）吸热并蒸发，蒸气到达接近冷却管路的一端后再液化，向冷却介质放热，如图 4.8-6 所示热管的传热效率几乎是铜导管的 10 倍。为了保证良好的导热，热管和模具之间应避免间隙，并应充填高导热的密封剂。

在进行模具的冷却系统设计时，需要确定以下设计参数：①冷却管道位置；②冷却管道尺寸；③冷却管道类型；④冷却管道的布局与连接；⑤冷却管道回路长度；⑥冷却介质的流动速率。应强调指出的是，冷却管道应使用标准尺寸，以便使用标准的加工设备、标准接头，实现快速拆卸。

注射模冷却系统的设计原则可总结如下。

1）为了缩短冷却时间，应避免使制品壁过厚，否则所需冷却时间会随壁厚增加而迅速增加。

2）制品壁厚应尽可能均匀一致（图 4.8-7）。

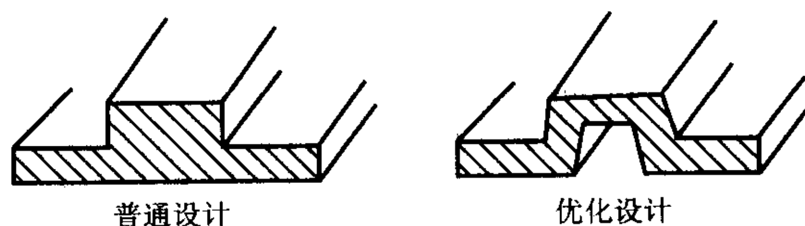


图 4.8-7　为获得一致的制品壁厚而改进设计

3）冷却管道的最好位置是在模具型腔与型芯板上，将

冷却管道置于型腔或型芯板之外会使模具冷却不充分。

4）制品对应的型腔或型芯两侧的温差应尽可能小，对于要求较严的制品温差不应超过10℃。

5）对于金属模具，最好使冷却管道（布置深度）离型腔或型芯表面1到2个直径，具体而言：对于钢模具布置深度应为1个直径，对于铜模具布置深度应为1.5个直径，对于铝模具布置深度应为2个直径。间距（冷却管道之间距离）应为3到5个直径，典型的冷却管道直径范围为10～14 mm，如图4.8-8所示。

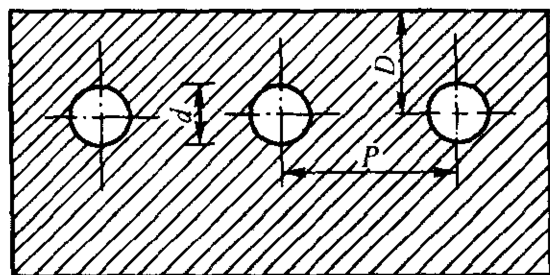


图4.8-8　典型的冷却管道直径（d）、**深度**（D）、**间距**（P）
$d=10\sim14$ mm；$D=(1\sim2)\ d$；$P=(3\sim5)\ d$

6）当冷却介质由层流转变为紊流时，热传输效率将显著增大。对于层流，热能是依靠层与层之间的热传导来传输；但在紊流中，径向的质量交换使得热量通过传导和对流两种方式进行交换，传热效率明显提高。

7）一旦冷却介质的流动变成紊流，提高冷却介质流速对于提高热流量和降低冷却时间就不很显著了，而压力损失和泵输送的费用却迅速增加。因此，当Re数超过10 000之后，就再也没有必要增加冷却介质的流速。

8）应注意使冷却介质在冷却系统的所有地方均达到紊流，利用CAE软件的冷却管道分析功能可以帮助用户确定和解决诸如淤塞冷却管道、短路冷却管道以及在某些冷却回路中的压力损失过大等问题。

9）空气层会降低传输效率，因此，应清除模具嵌件和模板之间的气隙和冷却管道中的气泡，以提高传热效率。

2　注射成形工艺参数的优化

成形工艺参数是影响注射成形零件质量非常重要的一个方面，注射成形的工艺参数通常有十多个，所有参数都可以影响到零件的质量。针对工艺参数对成形缺陷及产品力学性能的影响，已经进行了许多研究工作。实验表明成形参数对产品质量有很大的影响，如：增加保压压力和保压时间可以减少凹陷、对无定型材料降低注射速度有利于消除流痕等。

传统上注射成形的工艺参数由有经验的试模人员来确定。在实践中，试模师根据以前工作中类似的零件来设定当前零件的初始工艺方案，然后不断的调整、试验，直到获得满意的零件，调整过程的长短依赖于工艺师个人的经验。然而，由于工业企业对有经验工艺师的需求不断增长，导致有经验的工艺师远远供不应求，通常一个工艺师需要10～20年的经验才能成为试模专家。

面对注射成形企业的全球化竞争，采用传统的试错法（trial-and-error）来确定成形工艺参数已不可行，大量的研究工作都在尝试采用不同的方法来优化注射成形工艺参数，以便缩短市场响应时间和获得满意的产品质量。本节将简单地综述注射成形工艺参数优化的相关技术，主要包括数学模型、数值模拟、工艺窗口、实验设计方法、专家系统、人工神经网络、基于实例的推理、遗传算法等。

（1）数学模型（Mathematical Model）

基于流体力学的基本原理已经开发出一些注射成形过程的数学模型，如中心浇口圆盘的注射成形、圆管和矩形型腔的注射成形流动模型，这些简单的充填模型主要局限于规则的一维模型，多型腔的流动模型可以用分支流和网络流方法来描述。这些充填数学模型可以用来估算充填时间等。关于模具冷却，有的研究采用理论与统计相结合的方法来估算冷却时间。

数学模型也可以通过与其他技术相结合来确定工艺参数，如通过模糊推理来确定初始的工艺条件（包括熔体温度、压力，模具温度，充填时间），然后根据初始工艺条件通过数学模型来计算相应的机器配置，如喷嘴温度、注射速度等。显而易见，通过数学模型可以很快地得到注射成形参数，但是，注射成形过程非常复杂，对复杂零件不可能获得一个通用、可靠的数学模型。

（2）数值模拟（Numerical Simulation）

典型的数值模拟系统包括数学建模、数值方法和用户界面开发。针对注射成形充填、保压和冷却过程，已经建立了较多的数值模拟模型，如Hiber和Wang采用一个有限元/有限差分混合方法来模拟型腔的充填过程，该模型实现了非等温条件下无弹性、非牛顿流体的Hele-Shaw流动；Chiang等发展了一个统一的流动和保压阶段的集成模拟模型；Himasekhar等建立了冷却阶段三维模具传热模拟模型。数值模拟方面的一些研究成果直接转化成了商品化的模拟软件，如C-MOLD和Moldflow。

数值模拟可以为零件设计、模具设计和工艺设计提供有用的信息，数值模拟所预测的成形过程可以帮助工程师们克服经验的不足，并获得合理的工艺参数。然而，数值模拟方法要创建有限元模型并需要运行一系列的模拟后才能获得优化的成形参数。由于进行一个中等复杂零件的模拟目前还需要约一个多小时，因此在生产车间的环境中运行一系列的数值模拟是不切实际的。另外，数值模拟的结果需要使用人员正确的理解并自行进行相应的工艺调整，因此，用户需要具备很强的注射成形工艺背景知识。

（3）工艺窗口（Process Window）

成形条件对模塑制品的质量有重要影响，因此可以尝试建立一个可行工艺条件的区域，这种工艺区域通常称为工艺窗口。如图4.8-9所示，工艺窗口由一组边界组成，类似于一个窗口的形状。工艺窗口表明了注射压力与熔体温度的相互影响，如果熔体温度过低，就需要高的注射压力来驱动熔体充填型腔，如果熔体温度过高，材料就可能发生降解，另一方面，如果注射压力太低，可能会出现短射，如果注射压力过大，就会产生飞边。初始的工艺参数一般都选取在工艺窗口的中间位置。

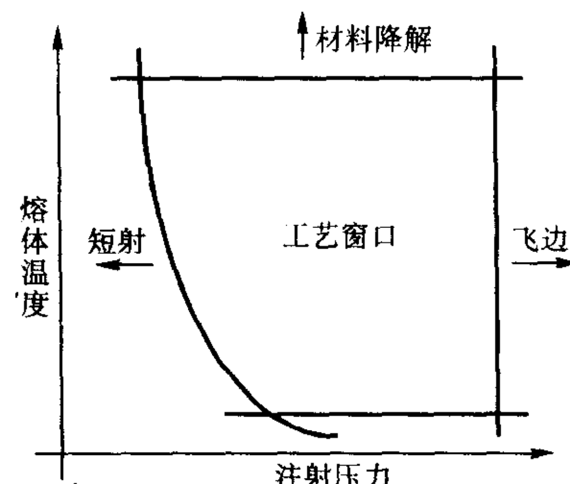


图4.8-9　典型的工艺窗口（注射压力对熔体温度）

有多种技术可以用来开发注射成形的工艺窗口，如对给定的材料、注射机和零件大概形状，可以通过组合不同的工艺条件进行模拟来确定工艺窗口，也可以采用约束处理方法和模拟结果来完成工艺窗口。为了从工艺窗口内获得最优的

或较优的参数设置，可以采用线形方法、函数优化方法等。

通过工艺窗口可以很容易地获得工艺参数设置，但是，要形成一个全系列的工艺窗口十分困难，因为涉及的工艺参数过多，并且参数之间相互关联。同时，确定工艺窗口需要进行大量的模拟测试，这需要相当的时间和费用。

(4) 实验设计方法（Design of Experiment）

在注射成形研究中，传统的实验方法只是反复改变某一个工艺参数直到获得满意的零件质量。这种方法忽略了工艺参数相互关联的影响，事实上，注射成形中的工艺参数是相互关联的，实验设计方法（Design of Experiment，DOE）有助于全面理解注射成形工艺参数对零件质量的影响。在 DOE 技术中，Taguchi 方法广泛应用于注射成形工艺参数优化，图 4.8-10 所示为采用 DOE 技术来确定最优工艺参数的典型流程，研究工作表明田口方法可以通过最少的实验次数来发现工艺变量之间的隐含关系。同时，对实验结果的分析也有助于发展注射成形模型，建立工艺条件与产品特性间的数学关系。

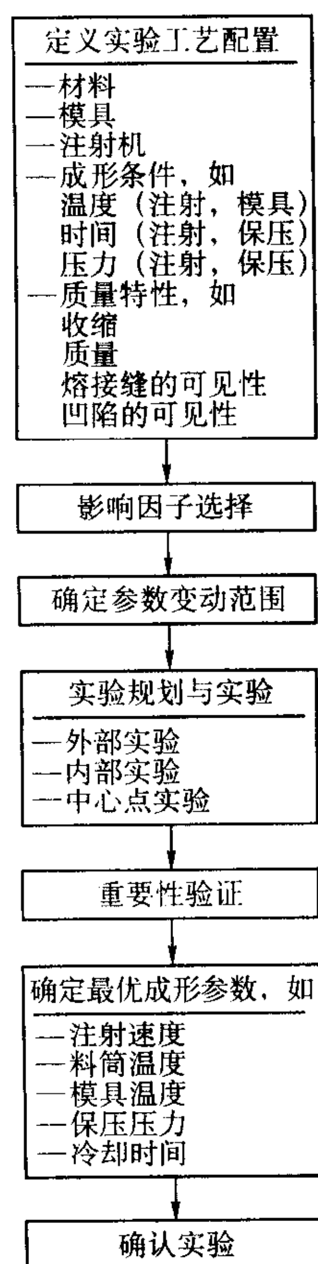


图 4.8-10 基于 DOE 技术的注射成形工艺参数优化过程

为了提高 DOE 的有效性，可以将 Taguchi 方法与其他的一些技术集成，如将质量功能配置（Quality Function Deployment，QFD）与 DOE 方法联系起来以区分工艺参数调整的优先次序，实现用户的满意度、控制参数的重要性、质量控制响应之间的最佳反馈。另外，将 Taguchi 方法与 Shainin 方法结合也可以提高注射成形过程的鲁棒性。

在 DOE 方法的实验规划中，需要一定的统计学和工艺过程知识。如果对实验规划的特性缺乏理解，两个参数间的相互作用就可能会给其他参数带来伪效应，从而导致错误的解释。因此，在 DOE 方法中，需要成形专家来确定工艺参数规划，而且，实验过程中也通常需要不断地中断生产以便设置实验参数和确定目标值。

(5) 专家系统（Expert System）

专家系统是人工智能技术的一个分支，它通过特定的知识和推理来模拟人的思维过程。一个典型的专家系统包括两个主要部分：知识库和推理机。知识库中包括事实和领域知识，知识的表达形式比较多，包括产生式规则、决策表、框架等。推理机可以看作是知识激活的组织和控制机构，其功能是根据问题的有关数据调用和操作知识库。注射成形专家系统可以根据输入的成形缺陷定量地调整成形参数，图 4.8-11 所示为注射成形缺陷修复专家系统的典型框图，系统要求输入所用的材料、当前的成形参数及所观察到的成型缺陷，知识库通过推理提供调整意见。

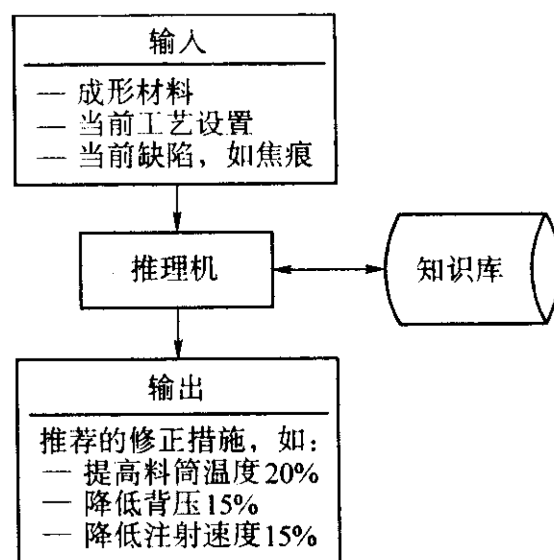


图 4.8-11 典型的注射成形缺陷修复专家系统

开发注射成形缺陷修复专家系统可以采用多种不同方法，如计算决策指数的方法，通过决策指数表明导致缺陷的各个变量的可能性。命题逻辑和模糊逻辑也可以用来推荐成形参数的定量修改，而多维矩阵方法则可用来将有经验的操作人员的思维过程结合到专家系统中。

在注射成形中比较适合采用专家系统，因为它可以囊括大量的工作中获得的经验和经验知识。专家系统可以给操作人员提供设置成形参数方面的帮助，然而要成功开发一个注射成形专家系统却并不容易，这主要是因为人类专家并不是一个规则系统，相反，他们是一个经验库，而经验本身都是零散的、无结构的，并不能很容易地转化成简单的规则形式。另外，专家的经验也有可能会超出他们对技术的理解，有时候他们自己也无法解释所做决定的原因。因此，要构造一个专家系统，必须要通过知识工程师与模塑人员的交流来引出适当的知识，知识获取的瓶颈是构造专家系统的主要困难。

(6) 基于实例的推理（Case Based Reasoning）

在 1980 年代中期，Riesbeck 和 Schank 首先提出了基于实例推理（CBR）的思想。基于实例推理的基本思路是采用解决老问题的方案来解决新问题，首先在实例库中对新问题进行匹配，找到一个或多个相似实例，通过修改最相似的实例来解决新问题。CBR 在很多工程领域中都获得了广泛的应用，如装配规划、建筑设计、系统诊断及注射成形中的注射模与注射零件成本估算系统等。关于工艺设计，CBR 方法已用于铸造及磨削工艺参数的确定，最近也被用来确定注射成形的工艺参数，CBR 系统推理的基本算法如图 4.8-12 所示。

CBR 系统中的实例保存在实例库中，可以通过索引来检索每一个实例，检索关键词可以是材料类型、零件复杂度、壁厚等。输入新问题后就在实例库中进行匹配，检索出一些可能的实例并进行相似度计算，找出最相似的实例来进行适当的修改以适用于新问题。

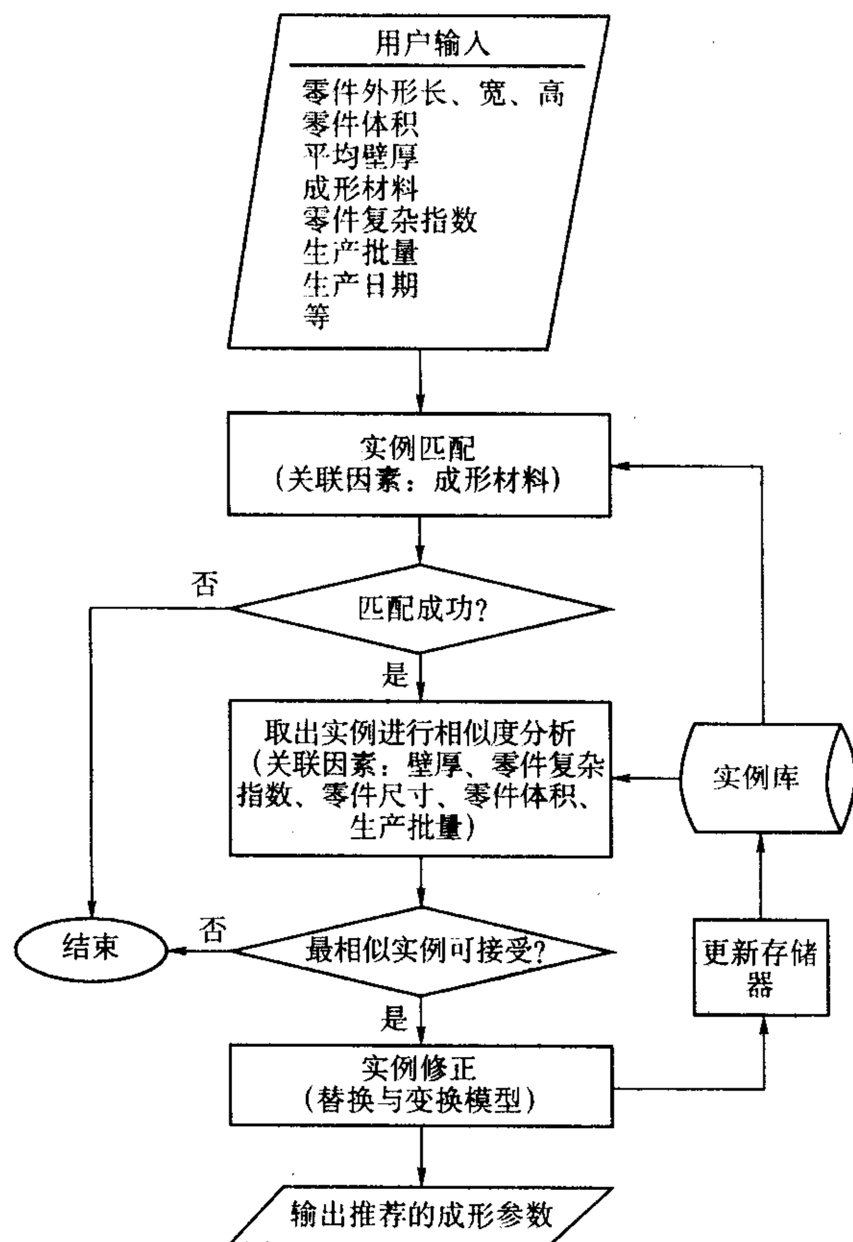


图 4.8-12　CBR 系统推理的基本算法

采用 CBR 方法可以迅速而方便地确定注射成形参数，该方法也不依赖于模塑专家，不存在知识获取的瓶颈问题，CBR 方法还通过实例库的更新积累而具有自学习的能力。CBR 系统的效果与效率依赖于实例库的大小、实例的检索、相似度计算及设计的修改模型。

（7）人工神经网络（Artifical Neural Network）

可以采用人工神经网络（ANN）来建立注射成形质量控制中的工艺模型，网络的输入为工艺参数，输出为制品的质量特征，网络训练成功后，就可以将在线质量预测或控制中的成形参数映射到质量特性。图 4.8-13 所示为一质量预测神经网络的结构，为了获得优化的注射成形参数设置，还应将零件的质量定量化为网络输出的函数，如采用性能指数并定义为线性收缩变化量和凹陷指数的函数，通过使性能指数最小化来获得优化的工艺参数。

神经网络可以与模糊逻辑结合起来构造注射成形参数设置的智能系统。模糊逻辑为参数设置中的不确定性提供了数学工具，在这样的系统中，输入为成形缺陷和模糊化的尺寸参数，输出为推荐的成形参数调整方法，所训练的神经网络充当一个模糊推理系统。

神经网络是复杂非线性系统建模的有效技术，由于神经网络是并行工作方式，处理过程比较快。不同于非线性拟合等方法，神经网络不需要预先假定任何的函数形式，如线性、一阶指数、对数等。另外，神经网络从训练样本中自动获取其“知识”，因此比较适合处理注射成形参数设置这种具有大量零散的、不易组织知识的问题。

神经网络从一定数量的训练样本中产生自身的规则，这个训练过程非常费时，如果新的输入超出了网络的预测范围，还必须获取新样本并重新训练。另外，对于一个给定的问题并没有构造神经网络的现成方法，开发人员设计网络结构的自由度很大，需要确定节点数目、激活函数、网络层数等。设计过程本质上是一个试错过程，非常费时。而且神经网络在工作过程中并不与使用人员进行任何通讯，因此使用人员不可能通过神经网络来进一步了解注射成形过程。

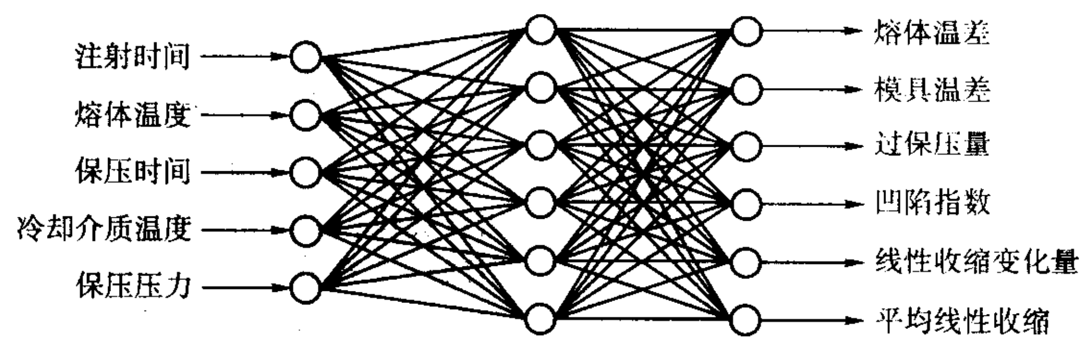


图 4.8-13　制品质量预测的神经网络结构

（8）遗传算法（Genetic Algorithms）

从对自然界的观察中发展了越来越多的人工智能技术，遗传算法（GA）就是其中的一种。基于流动模拟的结果，可以采用 GA 方法来建立注射成形工艺参数的优化系统，图 4.8-14 所示是基于 GA 方法的注射成形优化典型流程。在该系统中，初始的成形参数在取值空间内随机产生，并通过流动模拟来评价，制品的质量通过适合度函数来定量化，进化过程不断地循环迭代直到获得最优的或较优的成形设置。

通过遗传算法，即使没有注射成形的工艺知识也可以获得优化的参数设置，如果采用动态优化，在生产中的任何扰动也都可以得到及时的补偿。遗传算法的效率在很大程度上依赖于工艺参数的选择空间和定量化成形零件质量的适合度函数的设计，如果采用不恰当的适合度函数或取值空间，计算过程就可能永远也不能收敛到一个较优的解，或者变成了一个低效率的解空间的自由行走问题，因此，在系统开发中需要有经验的模塑工艺师。同时，遗传算法的有关控制参数，如群体大小、杂交比率、突变比率等也需要合理地进行设置。

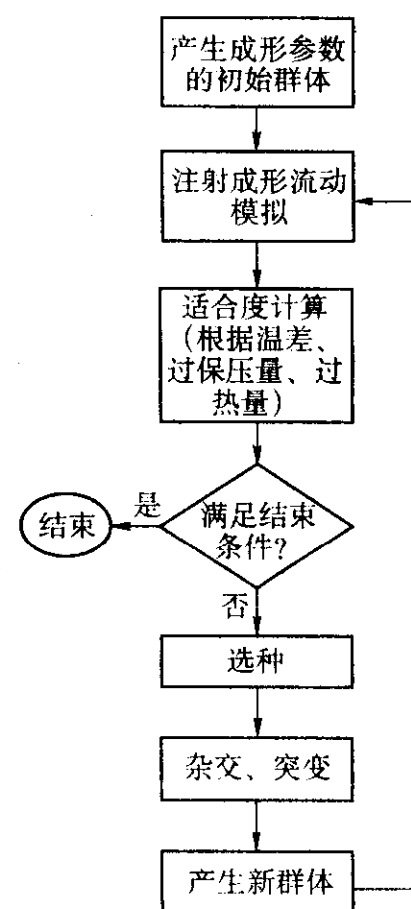


图 4.8-14　基于 GA 方法的注射成形工艺参数优化流程

(9) 总结

本节综述了注射成形工艺参数确定方面不同的技术方法，包括数学模型、数值模拟、工艺窗口、实验设计、专家系统、人工神经网络、基于实例推理、遗传算法等。

由于缺少有经验的模塑工艺人员，传统的试错法已不很适用于注射成形参数设置，虽然也有一些参数设置的基本原则和数学公式，但是由于注射成形过程非常复杂，仅凭这些简单的原则和公式无法实现对成形过程的充分理解和最优控制。通过引入数值模拟技术可以在一定程度上避免这个问题，模拟结果可以用来帮助确定成形参数，也可以用来定义无缺陷的成形区域或工作窗口。同样，实验设计技术也可以起到数值模拟类似的效果。然而，上述这些方法都需要一定的试注射，由于注射成形涉及参数较多，且参数之间相互关联，因此也需要有经验的模塑人员来设计工艺窗口或实验。

一些人工智能技术也可以用来确定注射成形工艺参数，如专家系统、人工神经网络、遗传算法、基于实例的推理等。基于专家系统的方法可以为工艺设置提供有用信息，存在的问题是模塑工艺师的一些经验很难转化为规则。ANN、GA和CBR是确定注射成形工艺参数中较新的方法。基于其所预测的产品质量，一个训练好的ANN系统可以迅速地优化一组成形参数，但是，网络训练的时间会比较长。采用GA方法，即便没有注射成形的工艺知识也可以优化成形参数，但在实际应用中，某些情况下优化过程的收敛速度会相当慢。基于相似的实例，CBR系统可以获得一组初始的工艺参数，系统的效率依赖于实例库的大小、检索方法和修改算法。

注射成形工艺参数确定技术的发展趋势是采用人工智能技术，应该看到采用单一的一种技术很难成功解决这个复杂问题，不同智能技术的结合是建立注射成形参数设置系统的有效途径。

3 充填模式优化

(1) 充填模式的基本用途

传统上进行充填模式分析时，必须先制作原型模具，并在注射机上进行缺料注射（短射）。三维流动模拟软件能直观地显示熔体从进料口逐渐充满型腔的动态过程，不需要实际成形就可以确定熔体前端在不同时刻的位置，并由此可判断熔体的流动是否为较理想的单向流形式、各个流动分支是否在同时充满型腔的各个角落、熔接缝是否产生在预定的位置、熔体在何处最后充满型腔（可能产生气穴的位置）。通过流动前沿的动态显示，可以预测成形中的跑道效应、滞流现象、流动的平衡性等。若熔体的填充过程不理想，可以改变塑件与模具设计（特别是改变进料口的尺寸、数量和位置）、工艺参数，反复运行流动分析软件，一直到获得理想的流动形式为止。

(2) 熔体前锋面速度和熔体前锋面面积

充填模式的另一个重要作用是分析充填过程中的熔体前锋面速度和熔体前锋面面积。所谓熔体前锋面面积（Melt Front Area，MFA）是指充填中熔体前端锋面的面积，熔体前锋面速度（Melt Front Velocity，MFV）就是熔体前端锋面的推进速度，MFV与MFA的乘积等于熔体体积流率。

应当注意的是，对于任何复杂形状的型腔，当螺杆的推进速度或熔体体积流率不变时，并不能保证熔体前锋面速度不变，熔体前锋面速度会随着熔体前锋面面积变化而变化。图4.8-15显示在嵌件周围尽管体积流率恒定，但MFA减小了，反之MFV增加了，在嵌件周围产生了高的应力和取向，从而导致不同的收缩和翘曲。

熔体前锋面速度与分子及纤维的取向程度关系极大。在注射成形的充填阶段，塑料材料分子和增强纤维（如果有的话）的取向受剪切流的影响，由于熔体是被注射到一冷模具中，塑件表面的分子取向在熔体到达的瞬间就被冻结（称为表面定向），因此分子与纤维的取向状态取决于熔体前锋面的流动形态。由于“喷泉效应”的影响，在充填过程中，熔体前锋面的形态较复杂，但总的来说，熔体前锋面流动速度越高，表面应力、分子和纤维取向越高，变化的前锋面流动速度会导致不同的取向，从而产生不同的收缩，使塑件产生翘曲。所以，在充填时要求维持熔体前锋面有稳定的充填速度，以产生一致的分子和纤维取向。

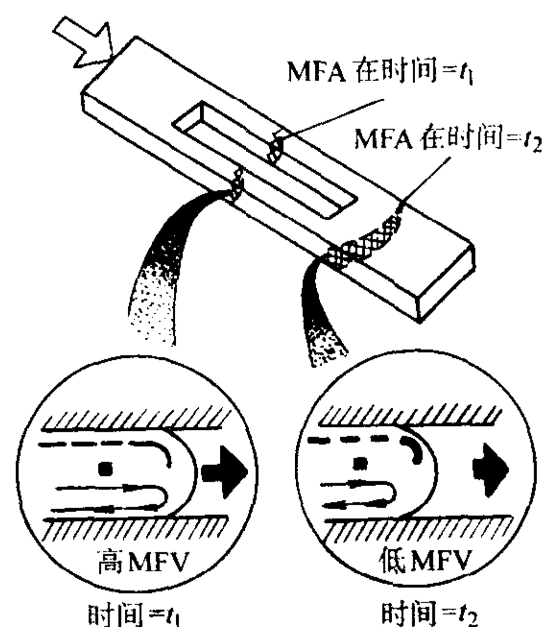


图4.8-15 熔体前锋面速度与熔体前锋面面积

(3) MFV、MFA和流动平衡

MFV与MFA是重要的设计参数，尤其是在型腔充填的流动平衡中，MFA可作为比较流动平衡程度的量化指标。确切地说，流动不平衡是指当熔体前锋面的一部分到达型腔的末端时，另外一些熔体仍在流动。当不平衡状态一出现，熔体前锋面面积就会突然变化。另一方面，平衡的流动通常是型腔中的熔体前锋面面积有极小的变化。对于给定的复杂的塑件几何形状，可以通过使MFA变化最小来决定最优的浇口位置。图4.8-16显示了对于平衡与不平衡两种不同充填模式的MFA曲线。

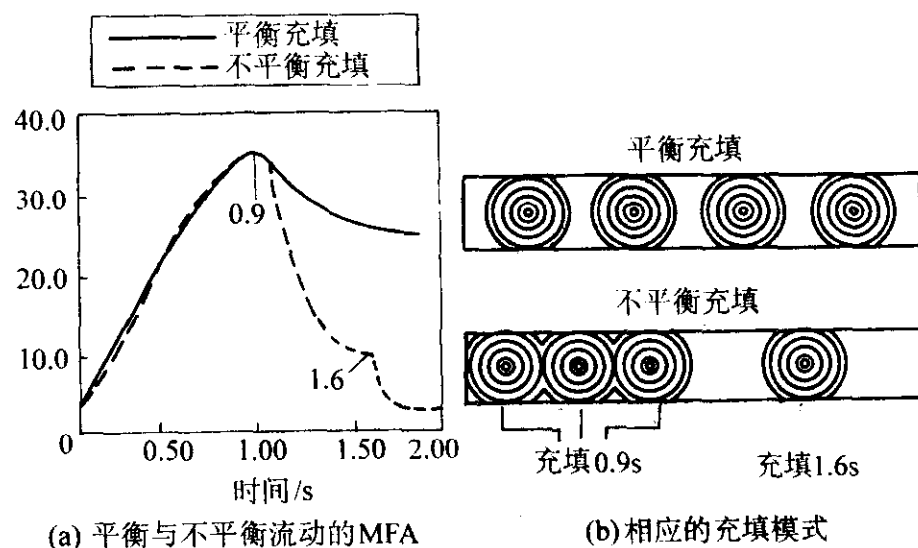


图4.8-16 不同充填模式的MFA曲线

(4) 分级注射优化

在充填过程中维持恒定的MFV的另一措施就是调节螺杆推进速度（即采用多级注射），几乎所有的现代注射机都具有多级注射的功能。下面通过一个简单的例子来说明如何根据CAE的结果来调整多级注射方案。

如图4.8-17所示，塑料熔体开始沿中间浇口半径扩展，起初随着充填区域的扩大MFA随之增大，但在型腔充填的后期，由于制品结构的原因MFA又降低，如图4.8-18所示。在恒定的螺杆推进速度下（恒定的熔体体积流率），MFA的

大幅度变化导致 MFV 也产生较大变化，图 4.8-17 中等值线间隔不均匀就是反映了这一现象。考虑到对分子与纤维取向的影响，这种变化的 MFV 是不理想的。目前有些 CAE 软件（如华中科技大学模具技术国家重点实验室开发的 HSCAE）可以计算出 MFA 随螺杆冲程（充填体积）的变化曲线，并根据该曲线拟合出最佳的分级注射方案，即优化的螺杆推进速度曲线。若采用该分级注射方案，在同样的型腔充填中 MFV 将变得更一致，这可以从预测的熔体前锋面推进等值线间隔更均匀中观察到，如图 4.8-19 所示。

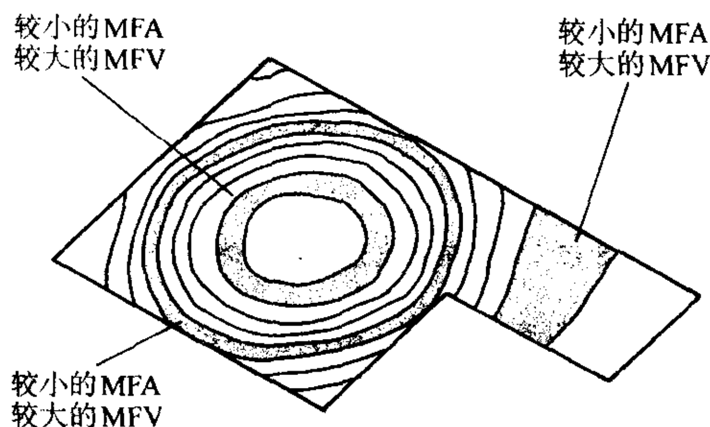


图 4.8-17　恒定的螺杆推进速度导致 MFV 发生变化

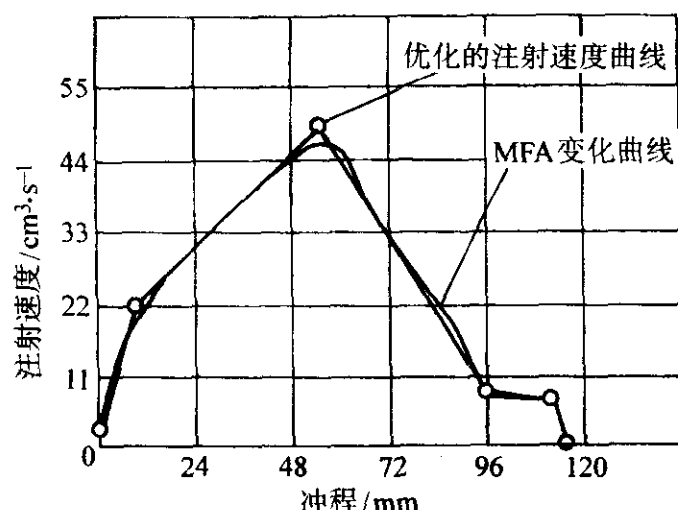


图 4.8-18　MFA 随螺杆冲程的变化及分级注射优化

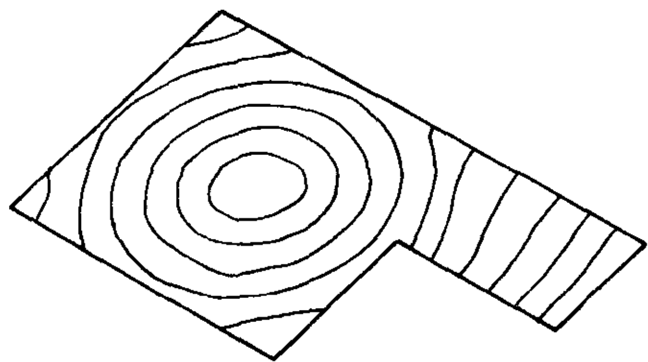


图 4.8-19　优化的分级注射方案产生稳定的 MFV

4　注射成形模拟软件及其智能化

4.1　注射成形模拟软件

注射成形模拟（CAE）是指注射过程的计算机分析，即模拟注射成形中熔体充模、保压与冷却过程以及预测塑料制品在脱模后的翘曲变形。CAE 技术借助于有限元法、有限差分法和边界元法等数值计算方法，分析型腔中塑料的流动、保压和冷却过程，计算制品和模具的应力分布，预测制品的翘曲变形，并由此分析工艺条件、材料参数及模具结构对制品质量的影响，达到优化制品和模具结构、优选成形工艺参数的目的。塑料注射成形 CAE 软件主要包括流动保压模拟、流道平衡设计、冷却模拟、翘曲预测等功能。其中流动保压模拟软件能提供不同时刻型腔内塑料熔体的温度、压力、剪切应力等的分布，其预测结果能直接指导工艺参数的选定及流道系统的设计；流道平衡设计软件能帮助用户对一模多腔模具的流道系统进行平衡设计，计算各个流道和浇口的尺寸，以保证塑料熔体能同时充满各个模腔；冷却模拟软件能计算冷却时间、制品及模腔的温度分布，其分析结果可以用来优化冷却系统的设计；应力计算和翘曲预测软件则能计算出制品的收缩情况和内应力的分布，预测制品出模后的变形。

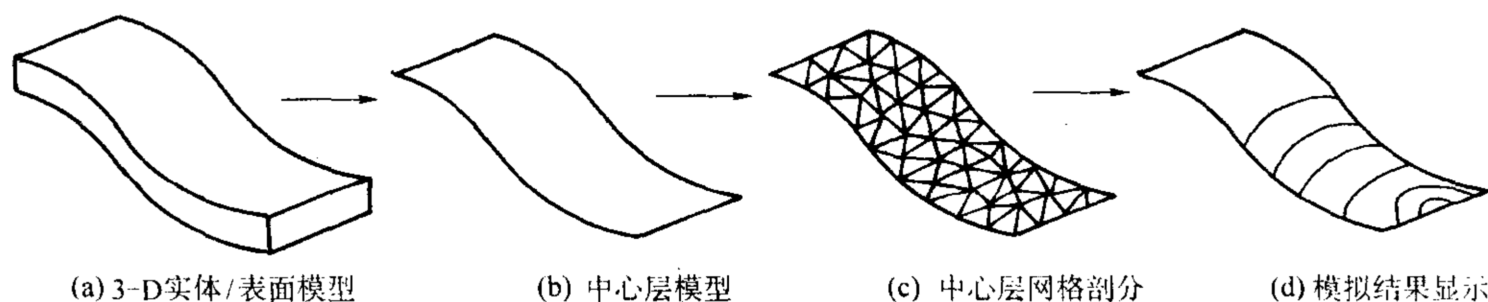
CAE 技术在当今的塑料工业领域得到了广泛应用。工程师们借助 CAE 技术可以全面地把握塑料注射成形过程，寻求一种增加产量、提高质量、节省时间和费用的最佳方案。CAE 分析对使用者来说是非常实用的。CAE 分析可以输出重要的设计数据，如压力分布、温度、剪切速率、切应力、速度等，设计者可由 CAE 获取诸如充填模式、熔接缝与气穴的位置、注射压力和锁模力大小、纤维取向、冷却时间、最终成形情况等信息。作为一种设计工具，CAE 能够辅助模具设计师优化流道系统与模具结构，协助产品设计师从工艺的角度改进产品形状、选择最佳成形性能的塑料，帮助模具制造者选择合适的注射机，指导模塑工程师设置合理的工艺条件。使用 CAE 分析可以对不同的成形方案进行反复的评测对比，寻求最优设计。同时，CAE 软件又是一种教学工具，通过对注射成形过程各阶段的定性与定量描述，CAE 能够帮助设计者熟悉熔体在型腔内的流动行为，把握熔体流动、保压、凝固的基本原则，帮助设计新手克服经验的不足，帮助有经验的工程师注意那些也许会被忽视的细节。

虽然 CAE 技术具有节省时间和原材料、降低废品率、提高产品质量、缩短产品开发周期等优势，但它并不是解决所有成形问题的万能方法。实际上 CAE 只是一种分析工具，用来帮助或完善工程师的设计而不是取代他们。因此，和其他工具一样，CAE 技术作用的大小取决于使用者的经验水平和熟练程度。确切地讲，CAE 分析结果的精度在很大程度上取决于设计人员提供的输入数据，以及对输出结果的正确判断与解释。

由于算法的局限性，一直以来，传统的注射成形 CAE 软件模拟软件都是基于制品的“中心层”模型（Middle Plane）。所谓中心层是假想的位于模具型腔和型芯中间的层面，其模拟过程如图 4.8-20 所示。基于这种型腔模型的 CAE 软件在应用中具有很大的局限性，主要表现在：①CAE 软件的使用人员必须理解中心层的概念，用户直接由产品模型构造中心层感到困难；②独立开发的 CAE 系统造型功能往往很差，依据模腔的 CAD 模型自动生成中心层模型又十分困难，这极大地妨碍了 CAE 软件的推广和普及；③由于 CAD、CAE 软件的模型不统一，二次建模不可避免。设计效率因此大打折扣，CAD、CAE 的集成也不可能实现。

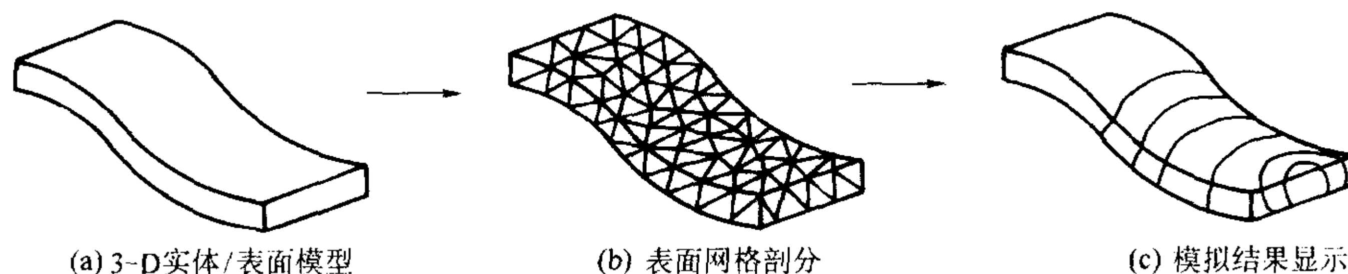
新近发展起来的基于实体/表面模型的模拟分析引入型腔的表面模型来取代其中心层模型，这一创新使得由现行的模具设计 CAD 系统产生的型腔模型可以直接应用于 CAE 软件，突破了传统软件的建模瓶颈。华中科技大学模具技术国家重点实验室和澳大利亚 MOLDFLOW 公司这方面的研究已走在世界前列，分别推出了成熟的商品化软件。基于型腔实体/表面模型的模拟过程如图 4.8-21 所示。

基于型腔实体/表面模型的模拟软件无需根据零件的 CAD 模型再构造产品的中心层模型，而是直接利用三维实体模型进行分析和显示。系统首先利用实体的表面信息生成模型表面的三角网格。对于薄壁塑料制品，可对所有三角网格的节点进行厚度方向的配对，产生节点配对信息及网格厚度信息。然后再利用网格数据、节点配对信息进行流动保压模拟，产生的分析结果直接在三维实体表面上显示。软件的执行的流程如图 4.8-22 所示。首先读取塑料制品的 STL 文件，显示制品图形，构造邻接拓扑关系，搜索制品边界（用于线



(a) 3-D实体/表面模型　(b) 中心层模型　(c) 中心层网格剖分　(d) 模拟结果显示

图 4.8-20　基于中心层模型的模拟过程



(a) 3-D实体/表面模型　(b) 表面网格剖分　(c) 模拟结果显示

图 4.8-21　基于实体/表面模型的模拟过程

框模型的显示)。然后进行表面重整、网格划分、网格优化以及节点配对。在此基础上用户定义浇口、设置工艺条件后即可进行流动保压分析。流动和保压分析结束后用户可以直观地浏览各项分析结果，包括流动前沿、气穴、熔接缝、温度场、压力场、剪切力场、剪切速率场、表面纤维定向、收缩率场、密度场等数据。

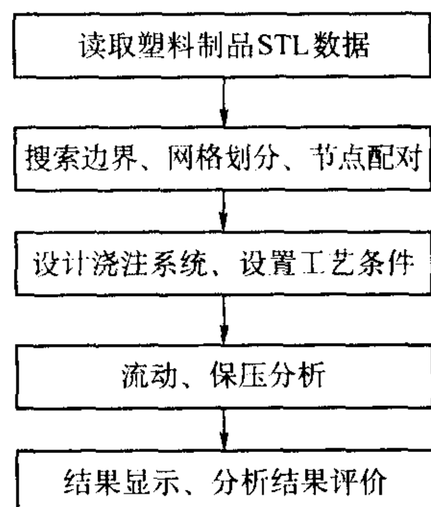


图 4.8-22　软件执行流程图

4.2　模拟软件的智能化

近年来，塑料注射成形 CAE 技术不论是在理论上还是在应用上都取得了长足进展，数学模型经历了一维、二维、中心层和目前的实体/表面模型，数值模拟的真实感和准确度显著提高，软件覆盖面不断扩大，能成功分析以往无法处理的大型复杂制品（如电视机外壳、洗衣机筒等）。但是，现有的注射成形 CAE 软件还只是一个“被动式”的模拟计算工具，即由用户给定设计方案后进行模拟计算，提供指定方案的分析结果。这样，要获得理想的成形方案，就需要设计者反复地分析与改进，并需要结合设计者的经验和技巧。基于 CAE 分析的自动优化技术可以大大减少用户分析的反复次数，显著提高系统的使用效率，这已成为注射成形 CAE 发展的重要方向。在现有的研究中，基于 CAE 的注射工艺优化设计主要是采用数值优化的方法（如灵敏度分析），但在注射成形中，由于成形工艺本身的特点，很难给出连续的、规则的优化求解空间，往往只是在有限个候选方案之间进行优选，同时，数值优化方法还存在计算量大、效率低、不能保证收敛、优化目标单一、优化过程不直接等缺点。因此，数值优化方法在注射成形 CAE 中的应用并不成功。另一方面，CAE 系统在分析能力显著提高的同时，对软件使用者的要求也在提高。在现有的注射成形 CAE 系统前置处理中，需要用户自行设置成形方案和工艺条件，但如何选择合理的初始成形方案和工艺条件常常要依赖于设计人员的经验；而在后置处理中，分析软件发挥了优异的图形功能，以等值线、矢量、阴影、动画、多模型、多窗口、图表、表格等方式充分表现分析结果，但都不具备对 CAE 分析结果进行解释和评价的功能，需要设计者自己正确分析和理解 CAE 模拟结果，同时，对设计的修改也要由设计者经过分析和判断后自行给出，对用户的要求过高。因此，目前的注射成形 CAE 分析软件尚属“专家型”的分析工具，这在很大程度上影响了 CAE 应用的普及。

知识工程的方法和技术是解决注射成形 CAE 现有不足的有效途径。20 世纪 90 年代以来，知识工程的研究异常活跃，相关的方法与技术层出不穷。考虑到现有 CAE 系统的不足，结合知识工程的发展，有必要把知识工程技术系统地引入注射成形 CAE 领域，通过建立包括产生式规则、实例、样本等集成多知识表达方式的知识库和规则推理（RBR）、实例推理（CBR）、人工神经网络（ANN）等集成多处理方式的多智能主体，建立支持 CAE 应用全过程的智能化系统，应用人工智能与专家系统方法在数值模拟结果的基础上自动对成形方案进行分析和评价，并提出方案的优化建议，通过 CAE 与知识工程技术共同实现“设计 - 评价 - 再设计”任务的自动化，提高设计的精确程度和效率，将仿真软件由传统的“被动式”计算工具提升为“主动式”的优化系统；与此同时，通过“机器专家”来指导完成分析模型的建立、提供工程化的分析结论，缩短了用户需求与分析结果之间存在的距离，将仿真软件由传统的“专家型”计算工具转变为“普及型”应用系统。

知识工程技术在注射成形 CAE 中主要应用于初始工艺方案设计、CAE 分析结果的解释与评价、分析方案的改进与优化三个方面，基于知识的智能型 CAE 系统的结构如图 4.8-23 所示，其中，工艺设计指导模块是根据一定的优化目标，自动为用户提供较优的初始参数，分析结果评价模块是利用专家系统处理 CAE 分析产生的大量数据，依据专家知识进行推理，将工程化的分析结论提供给用户，成形方案优化模块则是在分析结果评价的基础上，为用户提供参考的改进方案。结合人工智能和专家系统技术，注射成形 CAE 软件形成了以最优成形质量为目标的封闭回路，用户可在系统的帮助和提示下完成工艺方案的优化设计。

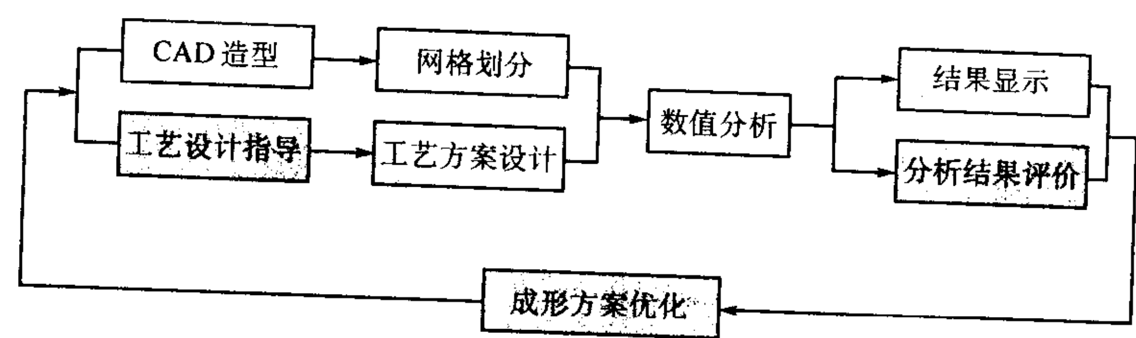


图 4.8-23　基于知识的注射成形 CAE 系统结构框图

编写：周华民（华中科技大学）

第 9 章　铸造工艺优化设计

铸造过程是一个庞大而复杂的系统，在长期的生产实际中积累了许多经验，这些经验对铸造过程的工艺优化起到了积极的促进作用。随着科学技术的发展，一些优化的数学方法在铸造过程中得到了初步的应用。计算机技术的发展，使铸件凝固模拟技术成功地在铸造生产中得到了应用，为铸造工艺的优化提供了可靠的依据。

1　浇注位置

铸件的浇注位置是指在浇注过程中铸件在铸型中所处的位置。正确的浇注位置能够保证铸件的形状完整，利于清理，并且有利于造型和造芯。

为了优化工艺以便能够生产出合格的铸件，对于铸件在浇注过程中浇注位置的确定，有以下几种比较重要的原则：

1）铸件的大平面和重要面应该放在底部或者侧面，以防止在这些位置出现气孔、砂眼等大的缺陷。如图 4.9-1 所示的零件。

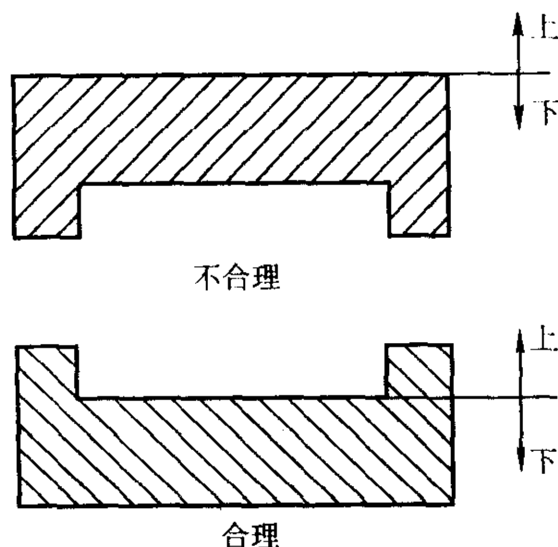


图 4.9-1　铸件的大平面放在底部

2）浇注位置应该有利于所确定的凝固顺序。对于体收缩较大的合金，应该尽量满足顺序凝固的原则，厚大部分一般位于上方，有利于设置冒口进行补缩。如图 4.9-2 所示。

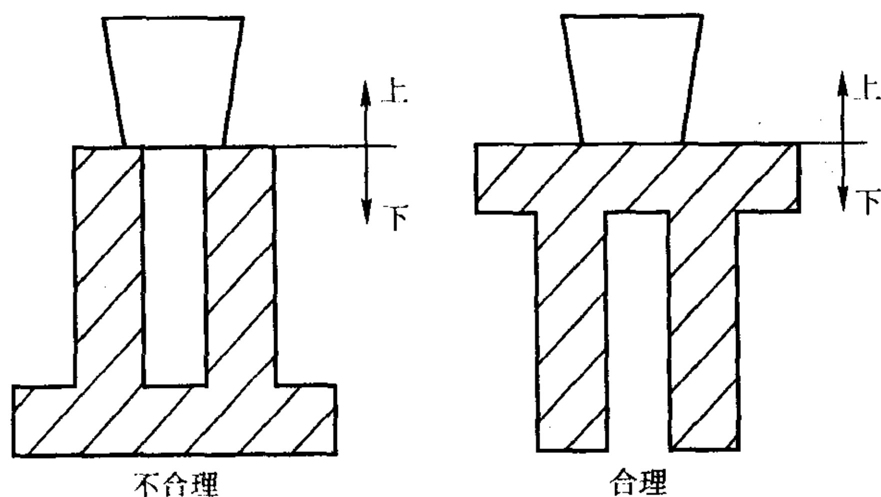


图 4.9-2　厚大部分位于上方

3）铸件的薄壁位置应该尽量放置在侧面或者是下部，以防止浇不足、冷隔等缺陷的发生，保证铸件的完整性。

4）浇注位置应该有利于砂芯的安放和排气通畅，尽量避免用到吊芯、悬臂砂芯，以保证铸件的形状，防止由于液体冲击而引起的破坏。

5）在大批量铸造过程中，应该使铸件的毛刺、飞边利于清理。

6）要防止主要工作面由于冒口而引起偏析，如图 4.9-3 的工作台。

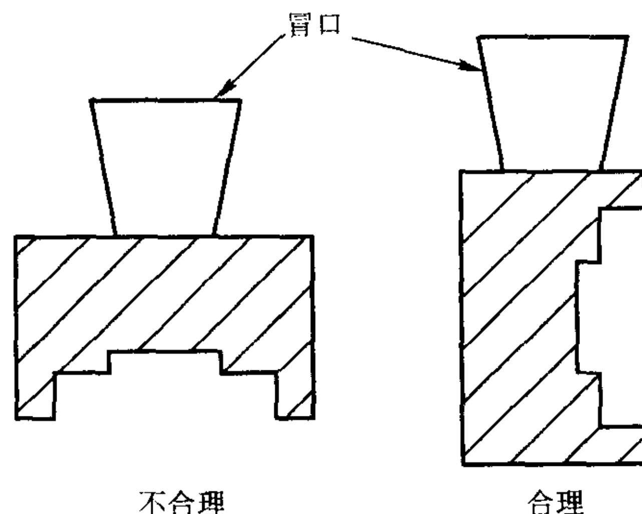


图 4.9-3　避免冒口在工作面上

2　分型面

分型面是指铸型组元间的结合面。合理地选择分型面，对于简化铸造工艺、提高生产率、降低成本、提高铸件质量等都有直接关系。分型面的选择应尽量与浇注位置一致，以避免合型后翻转。在铸造过程中，浇注位置与分型面的设置对于铸造过程中铸件质量，生产条件和效益有重大影响。在分型面的控制上，大体可以有以下几种方式：整体型设计、水平分型设计、垂直分型设计、综合分型设计。

分型面方式的选择应该着重考虑如下两点：①要力求简化铸型的结构，少用或不用活块，以减少加工工作量，降低铸型成本；②要便于浇冒口系统设置，型芯的安放和稳固以及在操作过程容易实现机械化，减少工人劳动强度，改善劳动条件。

在分型面的设计过程中，应该注意以下几个方面，以达到优化的目的。

1）尽量将整个铸件或者大部分放在同一箱内，以减少错箱带来的尺寸误差。如图 4.9-4 所示，增加了两个砂芯，却可以保证加工面的精度。

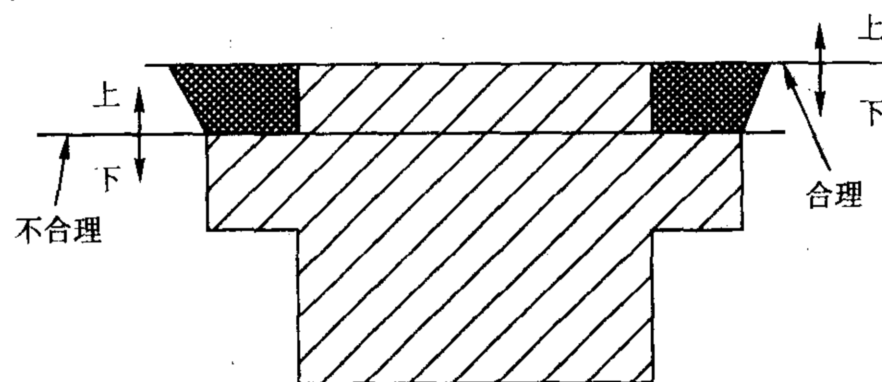


图 4.9-4　铸件整体放在同一箱内

2）尽量减少分型面的数目，在机器造型过程中，一般选用一个分型面。如图 4.9-5 所示，可以将含有三个分型面的铸型简化为只含有一个分型面。

在机器造型过程中，选择分型面时，应该避免使用活块，必要时应该用砂芯代替活块，如图 4.9-6 所示壳体的造型用砂芯代替活块。

3）尽量减少砂芯数量。

4）为了方便起模，分型面尽量选在铸型的最大截面处，对于较高的铸件，尽量不使铸件在一个箱内，以免模样在一箱内过高。

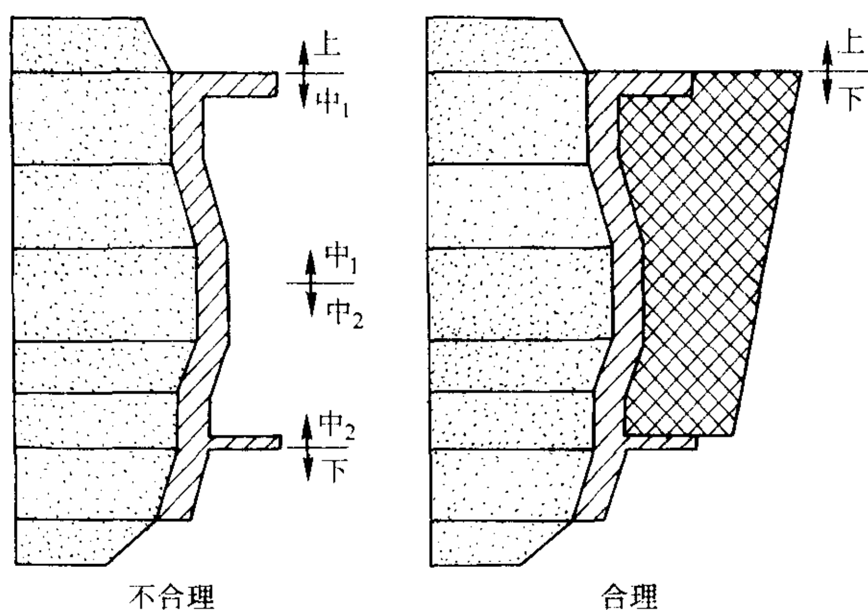


图 4.9-5 将三个分型面简化为一个

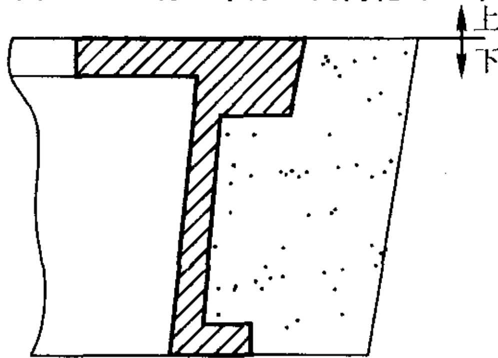


图 4.9-6 用砂芯代替活块

5）选择分型面的时候应该考虑到造型方法，高压造型与震击造型和射压造型相比，砂型紧实度较高，狭小吊砂处易坏型，所以在高压造型中应该避免狭小吊砂。

6）分型面的选择还应该有利于清理，图 4.9-7 中，方案（a）有较大的毛刺产生而不易清理，方案（b）则有利于清理。

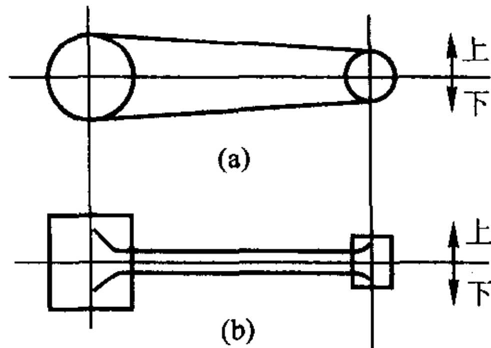


图 4.9-7 易于清理的分型面

以上介绍了几种常见的分型面的设置情况，在具体操作中，应根据具体的工艺要求，尽可能考虑以上几大因素来选择分型面。

3 浇注系统

浇注系统是砂型中引导液态合金流入型腔的通道，对浇注系统的要求是：①充型过程平稳；②有挡渣作用；③控制充型时的流量；④调整铸件的凝固顺序。由于浇注系统设计不合理，造成浇不足、孔偏、冷隔、缩孔、缩松等缺陷，由此引起的废品率占 40%。所以浇注系统与获得优质铸件，提高生产效率和降低生产成本的关系是密切相关的。

3.1 浇注系统的设计原则

通常，在普通铸造中浇注系统应满足下列要求。

1）确保进入型腔内的金属液流平稳，没有涡流现象。

2）选择合理的浇口位置，以利于铸件所需的凝固顺序。

3）浇注系统不引起或少引起铸件热应力。

4）保证在适当的浇注时间内，金属液在铸型内有合适的上升速度。

5）应使浇注系统消耗的金属液最少。

6）浇注系统的结构应具备排气、挡渣以及撇除其他杂质的作用。

为提高浇注系统的挡渣能力可采取如下措施。

1）采用闸门式或带浇口塞的浇口杯，浇口杯的容积应比常规的大些，最好大出 1/5 ~ 1/3。

2）选用结构合理、有挡渣能力的横浇道；加大横浇道的高度或在横浇道上设置集渣包，使金属液先经集渣包再流入铸型。

3.2 浇注系统的形式及适用情况

按金属液注入铸型（或内浇口）位置的不同，浇注系统可分以下几种（如图 4.9-8 所示）：(a) 顶注（包括顶雨淋浇口）；(b) 上（侧）注；(c) 分型面浇注；(d) 阶梯浇注；(e) 底（下侧）浇注；(f) 下雨淋浇注。对于不同的情况可以选择不同形式的浇注系统，达到优化的效果，保证铸件的质量。

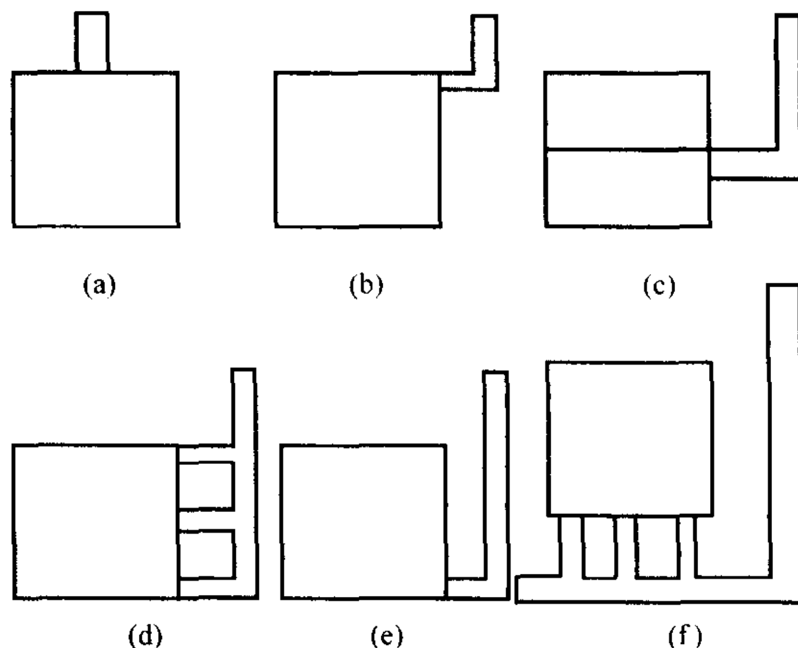


图 4.9-8 按浇口位置分的几种浇注系统

下雨淋浇口的特点是能使金属液均匀地从铸型底部注入型腔，液流平稳有利于排气、排渣，是一种适合实型铸造用的浇口。但这种浇注系统的结构较复杂，设置和制造也较麻烦，所以在生产上除特殊要求的铸件外一般是不多用的。

阶梯浇口多用于高度较大的铸件，使用时必须设置合理，浇口比例和结构也要恰当；否则金属液可能会由上下内浇口同时注入铸型；破坏了液流的平稳，干扰了模样的汽化条件，为此，这种浇注系统一般不主张采用。

底注式或下侧注浇口：合金液自下而上平稳地充填铸型，充填效果好，但不利于自下而上地顺序凝固，当铸件较高时，可在冒口部分设置补充浇道或直接在冒口补浇，以提高冒口补缩作用。因它的结构简单，造型操作方便，而且使用效果也较理想。所以它在大中型铸件的铸造生产中得到广泛应用。

上注式：有利于铸件自下而上地顺序凝固，但充型不平稳，易引起冲击和飞溅，使铸件产生气孔和氧化夹渣等缺陷，适用于外形简单，尺寸较小的铸件。

中注式：由于合金液从铸件高度的中间分型面处引入，因而降低液流下落高度，改善了上注式浇注系统充型不平稳的缺点，适用于高度不大的、容易在中间分型（分型后液落高度小于 100 mm）的大、中型铝、镁合金铸件，也是机械化造型的首选形式。

简单缝隙式：将扁平内浇道竖放，即形成简单缝隙浇道，合金液流动平稳，热分布合理，但去除浇道比较麻烦，金属液消耗多，缝隙厚度设计不良时容易在缝隙附近或对面产生缩松和裂纹，多应用于支架类铸件。

带立筒的缝隙式：具有顶注式和底注式的浇注系统的优点，但结构复杂，浇道去除困难。金属液消耗多，缝隙厚度

和立筒设计不良时，容易在缝隙附近或对面引起缩松和裂纹，大多应用于镁合金筒形铸件和金属液下落高度较大的铸件。

复合式：根据铸件结构复杂程度，为了保证铸件冶金质量，此种形式综合了上述形势的特点，一般使用特殊复杂的铸件。

压边浇道：用于水平截面不大的短小铸件，如小盖板等。

单边浇道：用于形状简单的铸件。

牛角浇道：用于质量要求较高的中、小型铸件，此种浇道具有稳流和建立合理热分布作用。

单向横浇道的连接式结构、搭接式结构用于一般的中、小型铸件（搭接式结构带有缓冲槽，其缓冲、稳流、撇渣效果较好）。其带过滤网的搭接式结构用于重要的中、小型铸件。

双向横浇道的连接式结构、搭接式结构：用于箱体、框架、大平板、管、套等一般铸件（搭接式结构带有缓冲槽，其缓冲、稳流、撇渣效果较好）。其带过滤网的搭接式结构用于重要的复杂的大型铸件。

4 冒口

冒口设计的工作主要是确定冒口的种类、形状、尺寸、数量和安放位置。

冒口的形状直接影响到它的补缩效果。如果冒口在相同体积下其表面积越小，则散热就慢，补缩效果就越好，一般说来，球形冒口补缩效果最好，常用的圆柱形冒口次之，而方形、长方形较差。

冒口在铸件上的位置正确与否，对获得健全铸件有着重要的意义。因此，冒口应尽量放在铸件最后凝固的热节点的上边或侧面；冒口还应尽量放在铸件最高最厚的地方，另外，冒口的安放应尽量不阻碍铸件的收缩，而且冒口最好不放在加工面上。在确定了冒口的种类、形状、位置后，就可以对冒口的尺寸进行设计。

4.1 冒口设计优化原则

（1）冒口位置的确定

铸造合金的液体收缩和凝固收缩，会使铸件肥厚部分产生缩孔或缩松。为了避免缩孔或缩松的形成，必须正确地运用冒口。冒口的作用除补缩铸件以获得致密的组织外，同时还能起排渣、排气的作用。

正确运用冒口的第一步是确定冒口的位置，即找出铸件上可能产生缩孔或缩松的部位。通常运用以下三种方法。

1）结晶等温线法　是在铸件断面上从冷却表面开始，相应于结晶温度的状况划分很多金属层，划到最狭小的断面上金属层相接触为止（见图4.9-9a），此时金属层不相毗连的地方即可能产生缩孔的位置。

2）内切圆法　是在铸件断面上画出一系列的内切圆（见图4.9-9b），在内切圆直径较大的地方（即热节），即可能产生缩孔的位置。

3）铸件凝固过程数值模拟法　利用计算机对铸件的温度场进行数值计算，进而对铸件的凝固过程进行模拟，预测铸件的缩孔、缩松等缺陷，并以图形方式显示其模拟的结果，可以准确地确定热节的位置。

（2）冒口设计要点

1）垂直分型的金属型，其冒口可做成上大下小的形状。水平分型的金属型，为了方便脱型，冒口必须做成上小下大的形状，或设计成活动块。

2）反锥形冒口中缩孔很深，有时会侵入铸件。为防止这种缺陷，冒口底部应设计砂套，使它具有向上扩大的标准形式。

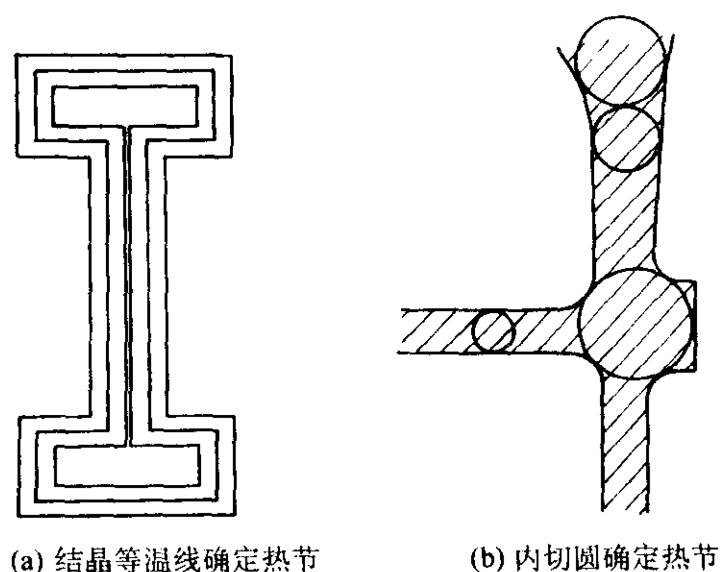

(a) 结晶等温线确定热节　(b) 内切圆确定热节

图4.9-9 金属液凝固顺序的确定方法

3）为提高冒口补缩效果，可运用电加热冒口、石棉纸保温及喷涂保温涂料等方法。

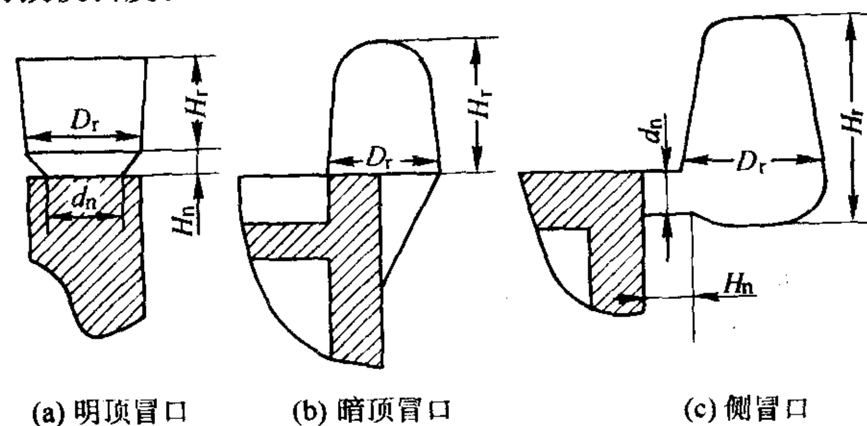
4）轻合金铸件明冒口高度不应小于60 mm，过低时补缩效果不好。但也不应大于200 mm，因为过高的冒口并不能完全保证补缩到铸件的较深处，同时还浪费金属，甚至引起其他缺陷。同样原因，冒口直径一般不超过100 mm为宜。

5）因液柱高100 mm的金属液静压力，对铝合金只等于0.026 6大气压，对镁合金仅0.017大气压，所以铝、镁合金冒口所能产生的静压力与大气压力相比是很小的。因此建议设计金属型时，能采用明冒口的地方，尽量采用明冒口而不用暗冒口。

4.2 冒口优化的数学模型

（1）设计变量

如图4.9-10所示，冒口的结构参数可以归纳为冒口直径 D_r，冒口高度 H_r，冒口颈直径 d_n 和冒口颈高（长）度 H_n。对于顶冒口，如果冒口和冒口颈采用45°过渡，冒口颈高度则为1/2（D_r-d_n）。对于侧冒口，冒口颈长度可取 $0.1d_n$，但不应小于10 mm。具体造型时冒口部分可取适当的拔模斜度。



(a) 明顶冒口　(b) 暗顶冒口　(c) 侧冒口

图4.9-10 冒口的种类

根据以上结构参数，设计变量为：

$$x=[x_1,\ x_2,\ x_3]^T=[D_r,\ d_n,\ H_r]^T$$

（2）目标函数

在保证铸件质量的前提下，应尽量减小冒口体积，这样不仅可以节约金属材料，而且可以减少做木模和造型的工时，从而获得较好的经济效益。对于圆柱形明冒口，其体积 V_r 可以这样计算（其他冒口也可用类似方法计算）：

$$V_r=\frac{1}{4}\pi D_r^2H_r+\frac{1}{12}\pi(D_r^2+d_n^2+D_rd_n)\frac{1}{2}(D_r-d_n)$$
$$=\frac{\pi}{24}[6D_r^2H_r+(D_r^2+d_n^2+D_rd_n)(D_r-d_n)] \quad (4.9\text{-}1)$$

故目标函数 $f(x)$ 可以写为：

$$\min f(x)=V_r=\frac{\pi}{24}[6D_r^2H_r+(D_r^2+d_n^2+D_rd_n)(D_r-d_n)] \quad (4.9\text{-}2)$$

(3) 约束条件

要使冒口达到补缩的目的，应建立如下约束条件。

1) 冒口的凝固时间应大于或等于铸件（或铸件被补缩部分）的凝固时间

为了保证冒口比铸件晚凝固，一般要求 $M_r \geqslant kM_c$（$k \geqslant 1$）

即 $$g_1(x) = kM_c - M_r \leqslant 0 \quad (4.9\text{-}3)$$

式中，M_r 为冒口的模数；M_c 为铸件的模数。

2) 冒口必须有足够的金属液补充铸件（或铸件被补缩部分）的体积收缩

考虑到冒口本身的凝固以及外界各种因素的影响，冒口不可能全部用于补缩，即要乘上补缩效率，因此

$$V_r\eta \geqslant \xi(V_r + V_c)$$

即 $$g_2(x) = \xi V_c - (\eta - \xi)V_r \leqslant 0 \quad (4.9\text{-}4)$$

式中，V_r 为冒口的体积，cm^3；V_c 为铸件被补缩部分的体积，cm^3；ξ 为金属在液态和凝固期总的体积收缩率，%；η 为冒口的补缩效率，%。

3) 必须保证冒口和铸件被补缩部位之间存在补缩通道

据此，冒口颈直径应该比铸件的热节点直径更大，以便冒口颈比热节点晚凝固，从而使冒口和铸件被补缩部位之间存在补缩通道，因此

$$d_n \geqslant d$$

即 $$g_3(x) = d - d_n \leqslant 0 \quad (4.9\text{-}5)$$

式中，d 为热节点直径，cm。

4) 与此同时，冒口颈还需保持一定高度，一般应大于15 mm，因此

$$\frac{1}{2}(D_r - d_n) \geqslant 1.5$$

即 $$g_4(x) = 3 - D_r + d_n \leqslant 0 \quad (4.9\text{-}6)$$

5) 必须保证缩孔产生在冒口内 要获得健全的铸件，必须保证缩孔完全在冒口内。因此，冒口的高度就不能太低，通常可取：$H_r \geqslant 1.4D_r$

即 $$g_5(x) = 1.4D_r - H_r \leqslant 0 \quad (4.9\text{-}7)$$

因此，冒口优化设计的数学模型可以简单记为

$$\min f(x) = \frac{\pi}{24}\left[6D_r^2H_r + (D_r^2 + d_n^2 + D_rd_n)(D_r - d_n)\right] \quad (4.9\text{-}8)$$

$$(x = [x_1, x_2, x_3]^T \quad x \in R^3)$$

约束函数：$g_i(x) \leqslant 0 \quad (i = 1, 2 \cdots\cdots 5) \quad (4.9\text{-}9)$

(4) 模型的应用方法

冒口优化设计数学模型有3个设计变量，5个约束条件，属于小型优化设计问题。由于随机方向法对目标函数的性态无特殊要求，而且其可行搜索方向是从许多随机方向中选择的使目标函数下降最快的方向，加之步长还可以灵活变动，收敛速度比较快。因此，本优化模型的求解采用随机方向法。其步骤如下。

1) 将变量离散化为规格化点，确定初始参数。本模型属于三维问题，所以只比较一层离散点。一般可取8~20个点为宜。比较确定层的离散点的函数值，以函数值最小的点确定出搜索方向。

2) 搜索方向确定后，可从初始点出发，沿着各个搜索方向利用离散变量的一维搜索法——步进法进行搜索，求得下一个极小点。然后重复上一步骤，确定新的搜索方向，求得新的搜索方向上的极小点。依次迭代下去，直到求得某一点，其周围离散点均不符合可行和适用条件，再不能确定出新的搜索方向求极小为止，这一点即为目标函数的最优解。这个迭代求解过程一般通过计算机程序来实现。

5 铸件结构的优化

在铸造过程中，铸件结构对铸件的质量有很大的影响，一方面造成铸件的铸造应力，导致热裂、冷裂和变形等缺陷；另一方面会影响凝固过程，导致缩孔、缩松、偏析和晶粒粗大等缺陷。

5.1 铸件壁厚及其结构形式的优化

(1) 铸件的临界壁厚

在铸件结构设计时，须考虑铸造合金的力学性能对铸件壁厚的敏感性。铸件过于厚大容易产生缩孔、缩松、晶粒粗大、偏析和松软等缺陷，铸件的力学性能也随之下降。从这方面考虑，各种铸造合金都存在一个临界壁厚，铸件的壁厚超过临界壁厚以后，铸件的力学性能并不按比例地随铸件厚度的增加而增加，而是显著地下降。因此，铸件的结构设计应科学地选择壁厚，以节约金属和减小铸件质量。对于可锻铸铁件，为保证获得白口坯件，其壁厚不能过大。球墨铸铁件在相同的球化工艺条件下，厚大铸件易出现球化衰退现象，造成球化不良，使铸件的力学性能显著恶化，因此球墨铸铁件的壁厚亦不宜过大。砂型铸造各种铸造合金铸件的临界壁厚可按其最小壁厚的三倍来考虑，也可按表4.9-1确定。

表4.9-1 砂型铸造各种铸造合金的临界壁厚 mm

合金种类	牌号	当铸件质量(kg)为下列值时		
		0.1~2.5	2.5~10	>10
灰铸铁	HT100，HT150	8~10	10~15	20~25
	HT200，HT250	12~15	12~15	12~18
	HT300	12~18	15~18	25
	HT350	15~20	15~20	25
可锻铸铁	KTH300-06，KTH330-8	6~10	10~12	—
	KTH350-10，KTH370-12	6~10	10~12	—
球墨铸铁	QT400-15，QT450-10	10	15~20	50
	QT500-7，QT600-3	14~18	18~20	60
碳素钢	ZG200-400，ZG230-450	18	25	—
	ZG270-500，ZG310-570，ZG346-640	15	20	—
铝合金		6~10	6~12	10~14
镁合金		10~14	12~18	—
锡青铜		—	6~8	—

上述推荐的临界壁厚数值虽不相同，但可作为对铸件的铸造工艺进行优化的参考。

(2) 铸件的内壁厚度

砂型铸造时，铸件内壁的散热条件差，即使内壁厚度与外壁厚度相等，但由于它比外壁的凝固速度慢，力学性能往往比外壁的要低，同时在铸造过程中易在内、外壁交接处产生热应力致使铸件产生裂纹，对于凝固收缩大的铸造合金还易产生缩孔和缩松。因此，将铸件的内壁厚度设计得比外壁薄一些是合理的。

(3) 铸件壁的过渡和连接

一般情况下，铸件壁的断面尺寸不可能完全相同，同时铸件壁之间有类型各异的接头。壁的过渡形式按接头断面的形状大致可分为L形、T形、V形、Y形和十字形五种形式。在接头处，凝固速度慢，容易产生应力集中、裂纹、变形、缩孔、缩松等缺陷。应注意选择合理的接头，以减小与分散

热节点以及避免交叉连接。如图 4.9-11 ~ 图 4.9-14 是各种接头连接形式及其优化实例。

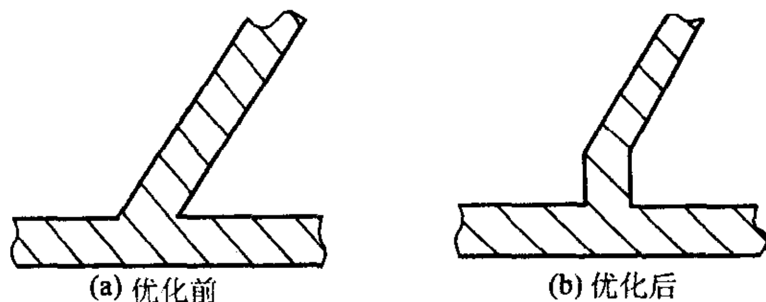


图 4.9-11　L 形接头优化实例

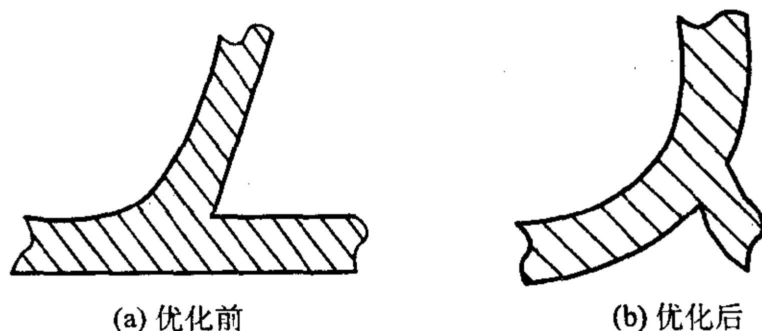


图 4.9-12　Y 形接头优化实例

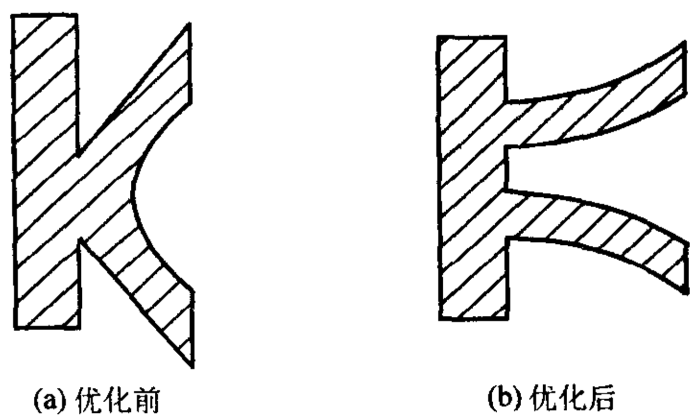


图 4.9-13　V 形接头优化实例

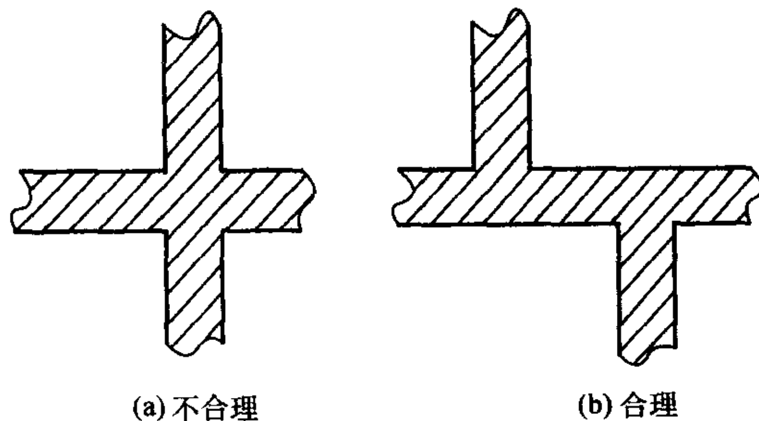


图 4.9-14　十字形接头优化实例

(4) 肋

为了增加铸件的力学性能、减小铸件质量、消除铸件的缩孔和防止铸件产生裂纹、变形、夹砂等缺陷，在铸件结构设计中大量采用肋。

设计肋时，要尽量分散与减少热节点，避免多条肋互相交叉，肋与肋和肋与壁的连接处要有圆角，垂直于分型面的肋应有铸造斜度。除此，还应考虑如下问题。

1) 应用肋来提高铸件质量和承载性能　图 4.9-15 所示的大平面壳体铸件，按原设计，在浇注时因高温金属液的烘烤，平面 *A* 处的砂型易“起皮”而产生夹砂缺陷。增设肋后该处不易“起皮”，也就不易产生夹砂缺陷。这种肋也有利于合金液充满该平面。

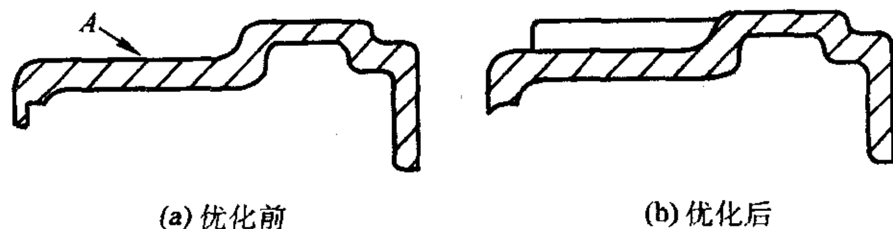


图 4.9-15　防止砂型起皮的肋的用法

2) 肋的布置应尽量减少热节点　肋的布置与铸壁一样要尽量避免肋与肋的十字形交叉（如图 4.9-16）；肋与壁相交时，如有必要，可在热节点处开孔（如图 4.9-17），以防止缩松、裂纹。

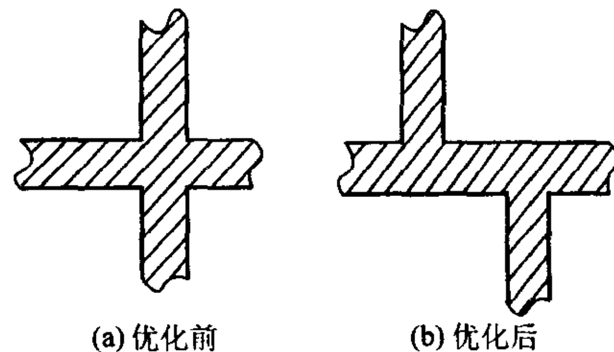


图 4.9-16　肋的布置

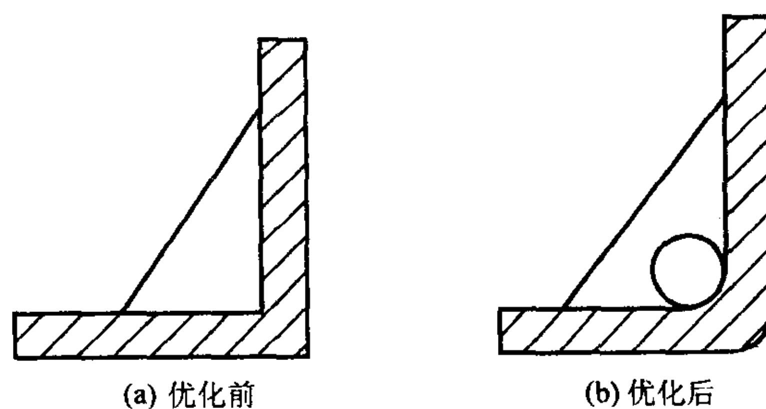


图 4.9-17　肋板加工艺孔

5.2　不同铸造合金铸件结构的优化原则

铸件结构应根据铸造合金凝固体收缩和线收缩特性、铸件形状特点进行设计。对于凝固收缩大，容易产生集中缩孔的合金，如铸钢、可锻铸铁、黄铜、无锡青铜、铝硅共晶合金等倾向于采用顺序凝固方式来设计铸件的壁厚；对于容易产生缩松的合金，并且采用冒口补缩效果不大的锡青铜、磷青铜，在设计铸件结构时常采用同时凝固方式以使缩松更分散一些，对于收缩较小的合金，如灰铸铁更倾向于采用同时凝固方式来设计铸件的壁厚。对于一些结构复杂的大型铸件，可根据对其不同部位的质量要求，分别按顺序或同时凝固方式进行设计，这样就容易保证铸件质量。对于复杂的大型铸钢件为保证其质量，可采用铸焊结构。

对于不同的铸造材料，其结构要求也不尽相同，下面就不同的铸造材料为保证获得恰当的凝固组织进行铸件结构优化的原则。

(1) 铸铁件的结构

1) 灰铸铁件的结构　灰铸铁件的壁厚不应太薄，边角处应适当加厚，防止出现白口组织使该处既硬又难于加工。

灰铸铁的力学性能对壁厚的敏感性极显著，因此不宜采用增加壁厚的办法来提高铸件的承载能力。要用合理的截面形状而不是增加界面厚度来提高承载能力。

灰铸铁几乎没有塑性，产生变形后不可能进行矫正。因此在设计灰铸铁件时，尽量按同时凝固的原则，采用对称的、能自由收缩的结构进行设计，以防止铸件产生变形和裂纹。

2) 球墨铸铁件的结构　球墨铸铁件的结构设计要力求壁厚均匀。球铁的弹性模量较高，壁厚不均造成较大的残余应力，甚至导致裂纹。

3) 可锻铸铁件的结构　铸件的壁厚不能太厚，以保证获得白口。当用冲天炉熔化时，可锻铸铁件的最大壁厚一般不应超过 25 mm；当用电炉熔炼时，最大壁厚不应超过 35 mm。这种铸件的凝固收缩较大，铸件结构的设计需要考

虑设置冒口的情况。

对于某些具有较大平面的可锻铸铁件，为防止裂纹可在其上设网格肋（网格肋的肋距为 12～17 mm，肋高 3～5 mm）。也常用加强肋防止不同厚度截面连接处产生裂纹。

4）高硅铸铁件的结构　高硅铸铁是一种耐腐蚀铸铁。它的流动性比灰铸铁差，因此铸件的壁厚一般不应小于 10 mm。而且高硅铸铁收缩大，易产生缩孔，在厚薄相差较大的铸壁连接处产生较大的热应力，易产生裂纹。导热性差，力学性能低，凝固时析出气体的倾向性较大，而且铁液表面易形成氧化膜。出于上述特点，铸件的结构力求壁厚均匀，铸件轮廓尺寸不宜过大（不应超过 600 mm×600 mm），壁的过渡要平滑，尽可能避免有阻碍收缩的凸出部分。用球面形壁代替面积较大的平面形壁，避免采用封闭形的截面，尽量采用能自由收缩的开式截面，以免收缩应力过大。

5）高铬铸铁件的结构　高铬铸铁是一种耐蚀铸铁，它的铸造性能介于灰铸铁与铸钢之间，具有如下特点：过热温度过高时，流动性才较好；凝固收缩较大，其值与白口铸铁相近；线收缩值较大（1.5%～2.0%）；形成粗大晶粒的倾向较大（特别是在含碳量较低时）；热导率较小；弹性模量比灰铸铁高。因此，这种铸铁件很容易产生缩孔和裂纹。结构设计时壁厚应尽量均匀，避免厚大截面和局部金属堆积，尽量采用能自由收缩的结构。这有利于防止缩孔、裂纹和粗晶组织。

（2）铸钢件的结构

铸钢的铸造性能与铸铁有很大差别，它的熔点高、流动性低、体收缩和线收缩大，易产生浇不足、缩孔、缩松、裂纹和变形等铸造缺陷。铸造圆角和壁的过渡尺寸比灰铸铁的大。

1）碳素钢与低合金钢铸件的结构　碳素钢与低合金钢的铸造性能差别不大，因此低合金钢铸件的结构可按碳素钢铸件的结构设计。

低碳钢的结晶温度范围窄，补缩性能好，产生分散性缩松的倾向性小，而是在铸件最后凝固的部位留下集中缩孔。易在等壁厚的铸壁上产生轴线缩松。而热裂倾向性小。

高碳钢的结晶温度范围宽，补缩性能较差，易产生缩松。其铸件也易产生热裂和冷裂。

中碳钢的铸造性能介于低碳钢和高碳钢之间，在实际应用中多采用中碳钢铸件。

对于承受高温、高内压和高疲劳或高冲击应力的铸钢件，应按顺序凝固原则设计铸件的结构，以获得健全的铸件。承受弯曲应力的铸件轴线附近的应力很小，轴线缩松对其强度的影响不大；受压应力作用的铸件也允许存在一定轴线缩松。因此，这类铸钢件可采用同时凝固形式进行壁厚设计。

某些大型铸钢件，由于不同部位承受的内压力或冲击应力不同，有的部位可按顺序凝固方式设计壁厚；有的部位可按同时凝固方式设计壁厚，以节省金属消耗，降低产品成本。

2）高锰钢铸件的结构　高锰钢铸件在热处理之前性能较脆；水韧处理时在厚截面处冷却速度慢要析出碳化物，使钢的塑性下降；它的线收缩大（2.4%～3.0%）和导热性能低（比碳钢约小 4/5）。因此裂纹倾向性大，在切割冒口时易在冒口根部产生裂纹。

为了防止裂纹和析出碳化物，应尽量按同时凝固原则设计高锰钢铸件的结构。

铸件结构应尽量能自由收缩，在某些情况下，为避免机械阻力，将高锰钢铸件分成几部分铸造，然后再组合成一个整体。

3）铸造不锈钢铸件的结构　在设计铸件结构时应考虑它的铸造性能较差，钢液中有氧化铬夹杂物，钢液表面易产生氧化铬薄膜阻碍钢液充填砂型，铸件很容易产生冷隔、夹杂和表面皱纹等缺陷，而且容易产生铸件粘砂。不锈钢中的铬钢的体收缩大，容易产生缩孔和缩松；导热性差，线收缩大，热应力大，容易出现较粗大的铸态组织，发生裂纹。因此，壁厚应尽量均匀，平缓过渡，避免阻碍收缩的凸出部分。

对承受较大压力的铸件，应按顺序凝固原则设计结构，以保证补缩效果，获得致密铸件。

（3）非铁合金铸件的结构

1）铝合金、镁合金铸件的结构　这类铸件的结构，一般力求壁厚均匀，尽量采用薄壁结构，这对防止缩孔、缩松和提高合金的力学性能均有利。

对于壁厚不均匀，需要冒口补缩的铝、镁合金铸件可加厚某些部位，实现顺序凝固以利于补缩，保证铸件质量。

2）锡青铜铸件的结构　锡青铜的结晶温度范围宽，易形成缩松，偏析倾向也大，因此铸件结构应尽量采用薄壁结构。对于气密性要求高的铸件还应避免机械加工，尽量保留铸件表面组织致密的激冷层，即铸造“黑皮层”。

高温时锡青铜较脆，强度较低，因此大铸件的凸出部分应该用较薄的加强肋加固，以免产生裂纹。铸件结构应尽量能自由收缩，防止产生裂纹。

3）无锡青铜和黄铜铸件的结构　这类合金的体收缩和线收缩都较大，易氧化；气孔、缩孔和缩松倾向性大。因此，一般按铸钢件设计原则来设计这类合金铸件。不同壁厚的过渡要平缓，壁厚差不应太大，两者之比不应大于 2.6（甚至 2.2）。

当铸件上不可避免地有厚大部分（法兰等）时，应按顺序凝固原则设计壁厚。有时可在铸件上局部“增肉”作为补缩通道，造成顺序凝固消除缩孔和缩松。为保证补缩效果，“增肉”截面厚度应不小于壁厚的 2～3 倍。

（4）铸造工艺对零件结构的要求

铸件的结构不仅应有利于保证铸件的质量，而且应考虑到模样制造、造型、造芯和清理等操作的方便，以利简化制造工艺过程，稳定产品质量，提高生产率和降低成本。铸造工艺对铸件结构的基本要求如下：简化或减少分型面；减少砂芯数量；方便起模；有利于砂芯的固定和排气；具有铸造工艺孔；有利于铸件清理；有利于满足铸件尺寸公差。

6　铸造工艺的 CAD/CAE 在铸造工艺优化中的作用

铸造工艺计算机辅助设计计算（CAD）和铸件凝固过程的数值模拟（CAE）技术已经达到了实用化程度，利用铸造工艺 CAD/CAE 同样可以达到对工艺优化的目的。具体过程如图 4.9-18 所示。

在铸件的铸造工艺 CAD 的过程中，利用计算机根据确定的工艺参数可以完成三维图形下的重量和模数计算、冒口设计计算、浇注系统设计计算和冷铁等的设计计算。铸件的铸造工艺 CAD 完成后可以得到铸件的铸造工艺三维图形。之后利用铸造工艺 CAE 对这种工艺下的铸件凝固过程进行数值模拟，并预测在凝固过程中产生的缺陷（如缩孔、缩松等）是否在铸件上。具有智能优化的数值模拟分析软件根据模拟的结果对工艺参数进行调整，进行下一次的模拟，最终达到合理的工艺，即实现了工艺优化的目的。

由于铸造工艺的影响因素较多且复杂，在优化方面往往就单一问题进行处理，如将铸件的缩松状况作为目标函数，使缩松达到最小值，这个值可视铸件生产的执行标准而定；冒口的高度和直径作为设计变量；同时也是约束条件，要小于铸型的尺寸，冒口的模数也是约束条件，要大于铸件的模

数。这一优化问题的求解过程是，模拟→检查目标函数→改变设计变量→再模拟……直至满足目标函数为止。

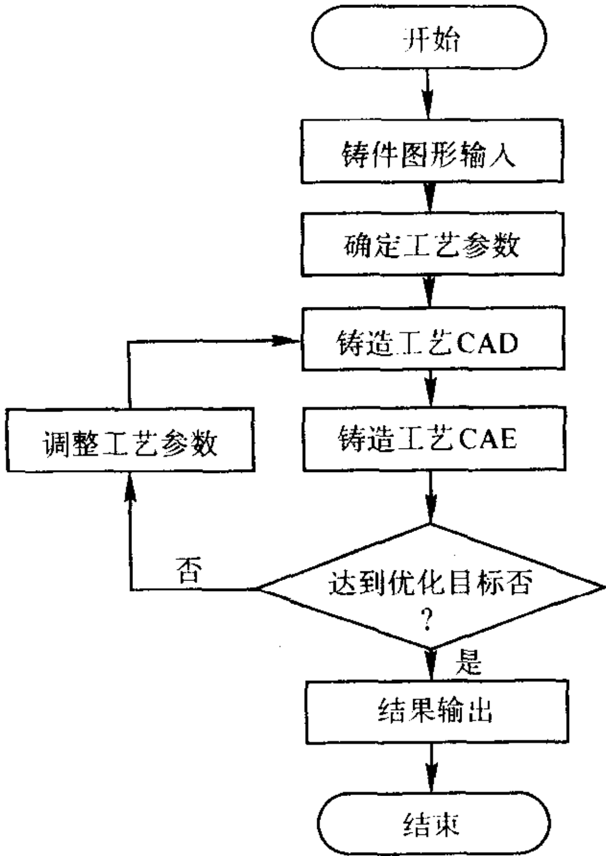


图4.9-18 铸造工艺CAD/CAE的流程图

以金属型浇注的汽车铝支架为例。在最初设计的基础上进行了6次模拟优化，每一次都在优化模块的控制下自动运行。在前几次的优化模拟中，由于铸件中有缩孔产生，则采用了增大冒口直径和高度措施。最后的铸件设计表明没有缩孔，缩孔都转移到冒口部位。

优化过程中微观缩松的程度不断发生变化，最初孔洞量为1.4%，经过优化过程，微观缩松度下降到0.42%，如图4.9-19所示，并且都转移到冒口中。

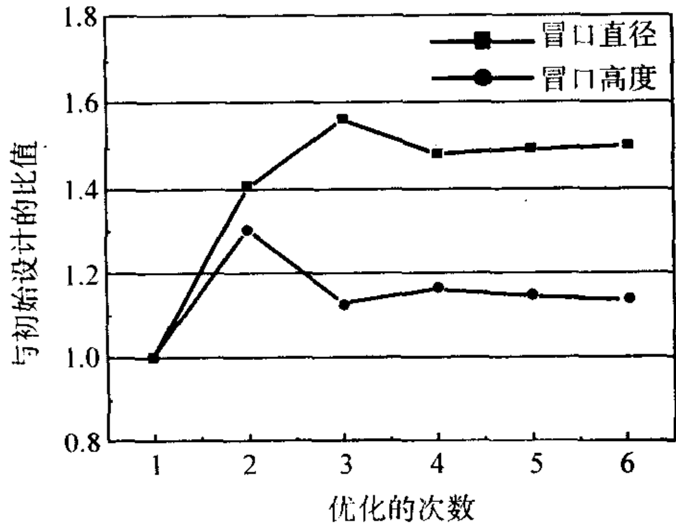


图4.9-19 优化过程中铸件冒口直径、高度变化情况

优化过程中冒口的高度和直径不断变化，最终的铸件无缩孔的冒口设计为：冒口直径、高度分别为原设计的1.14倍和1.5倍，如图4.9-20所示。这种情况下，冒口尺寸增加了，成品率增加了，微观缩松被限制在冒口中。

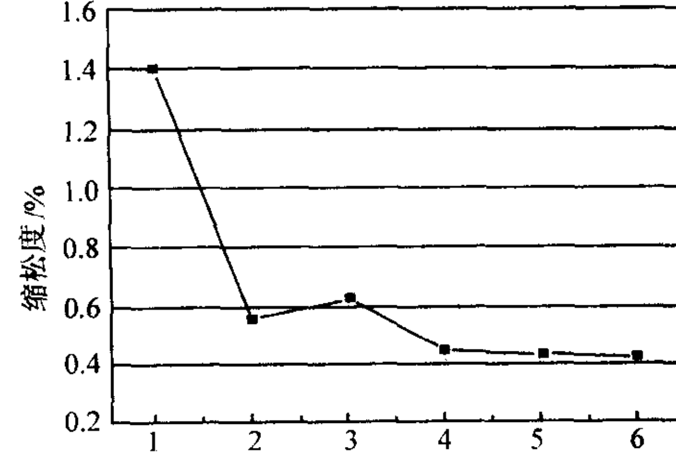


图4.9-20 优化过程中铸件的微观缩松的变化情况

铸造工艺CAD/CAE一方面代替人工进行铸造工艺的设计；另一方面代替了铸造工艺实验对工艺的检验。因此，利用铸造工艺CAD/CAE系统对铸造工艺进行设计，不仅达到了优化的目的，而且可以提高设计效率和降低设计成本。

7 熔炼过程配料的成本优化计算

在合金的熔炼过程中，配料是很重要的一个生产环节。配料的结果对合金生产的成本、合金产品的质量以及生产工艺的设计都有很大的影响。

不同的熔炼工艺和不同的原材料有不同的配料计算方法，当前最常用的配料优化方法是线性规划中的单纯形法，以及在此基础上发展的大M法、对偶单纯形法等方法，以及专门对应于块状原材料的整数规划方法。随着现代数学在工程中应用的发展，模糊算法、模拟退火算法、神经网络优化方法等方法也在优化配料中得到具体的应用。随着计算机技术的不断发展，使解决复杂的计算问题越来越简单，传统配料过程中许多计算简单的方法逐渐淘汰，而科学的配料优化方法得到了越来越广泛的应用。

7.1 线性规划方法

线性规划问题（linear programming又称LP问题）是满足以下条件的最优决策问题（做一决策使某一目标达到最优的问题）：

1）每个问题的决策方案可以用一组变量的取值来表示，这组变量称为决策变量；

2）有一定的约束条件，这个约束条件是关于变量的线性等式或线性不等式；

3）有一个可以表示为决策变量的线性函数的决策目标（称为决策的目标函数），决策的目的是使这个函数的取值达到最大或是最小。

线性规划一般的数学模型

$$\min\ (\max)\ S = \sum_{i=1}^{n} C_i x_i \tag{4.9-10}$$

$$\sum_{i=1}^{n} a_{ij} x_i \leqslant b_j\ (\text{或} \geqslant b_j\ \text{或} = b_j)\ (i = 1,\ 2,\ 3\cdots m) \tag{4.9-11}$$

式（4.9-10）是目标函数，式（4.9-11）是约束条件，x_i是决策变量。

满足所有的约束条件的决策变量的一组值，称为线性规划的一个可行解，常写成向量的形式

$$X = [x_1\ x_2 \cdots x_n]^T \tag{4.9-12}$$

所有可行解的集合称为可行域，使目标函数取得最优值的可行解，称为线性规划的最优解。求解一个线性规划的问题就是寻求线性规划的最优解的过程。而对于配料的线性规划问题，常用的求解方法有单纯形法，以及在此基础上发展的大M法、对偶单纯形法等。

7.2 配料计算中线性规划问题的建模

假设目标合金成分中有m种元素，第j种元素的最小质量分数为b_{jmin}，最大质量分数为b_{jmax}。配料需要n种原材料，配料所需的第i种原材料的质量分数为x_i，每种原料的价格为C_i，第i种原材料中含有第j种元素的质量分数是a_{ij}，η_j是第j种元素在熔炼过程中的增减系数。配料成本为S。配料的目的是在满足成分要求的基础上，使合金的成本达到最低。

配料建模的目标函数为

$$\min S = \sum_{i=1}^{n} C_i x_i \tag{4.9-13}$$

成分的约束条件

$$b_{j\min} \leqslant \sum_{i=1}^{n} a_{ij}\ (1+\eta_j)\ x_i \leqslant b_{j\max} \qquad (4.9\text{-}14)$$

在配料过程中不希望某种原材料超过某个值，则有

$$x_i \leqslant w_i \qquad (4.9\text{-}15)$$

式中，w_i 是某种原材料的上限。

非负性约束条件是

$$x_i \geqslant 0 \qquad (4.9\text{-}16)$$

归一化约束条件

$$\sum_{i=1}^{n} x_i = 1 \qquad (4.9\text{-}17)$$

式（4.9-16）和式（4.9-17）是在工程中 LP 问题的普适的约束条件。

这样得到 LP 问题

$$\min S = \sum_{i=1}^{n} C_i x_i \qquad (4.9\text{-}18)$$

$$\begin{array}{ccccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{n1}x_n & \geqslant & b_{1\min} \\
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{n1}x_n & \leqslant & b_{1\max} \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{1j}x_1 & + & a_{2j}x_2 & + & \cdots & + & a_{nj}x_n & \geqslant & b_{j\min} \\
a_{1j}x_1 & + & a_{2j}x_2 & + & \cdots & + & a_{nj}x_n & \leqslant & b_{j\max} \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{1m}x_1 & + & a_{2m}x_2 & + & \cdots & + & a_{nm}x_n & \geqslant & b_{m\min} \\
a_{1m}x_1 & + & a_{2m}x_2 & + & \cdots & + & a_{nm}x_n & \leqslant & b_{m\max} \\
x_1 & & & & & & & \leqslant & w_1 \\
 & & x_2 & & & & & \leqslant & w_2 \\
 & & & & \ddots & & & & \vdots \\
 & & & & & & x_n & \leqslant & w_n \\
x_1 & + & x_2 & + & \cdots & + & x_n & = & 1
\end{array} \qquad (4.9\text{-}19)$$

7.3 配料计算中线性规划问题的标准化

配料成本优化计算线性规划问题的标准形式如式（4.9-20）、式（4.9-21）所示

$$\min S = \sum_{i=1}^{n} C_i x_i \qquad (4.9\text{-}20)$$

$$\begin{array}{ccccc}
\sum_{i=1}^{n} a_{i1}x_i & + & (-1)^1 x_{n+1} & = & b_{1\min} \\
\sum_{i=1}^{n} a_{i1}x_i & + & (-1)^2 x_{n+2} & = & b_{1\max} \\
\vdots & & \vdots & & \vdots \\
\sum_{i=1}^{n} a_{ij}x_i & + & (-1)^{2j-1} x_{n+2j-1} & = & b_{j\min} \\
\sum_{i=1}^{n} a_{ij}x_i & + & (-1)^{2j} x_{n+2j} & = & b_{j\max} \\
\vdots & & \vdots & & \vdots \\
\sum_{i=1}^{n} a_{im}x_i & + & (-1)^{2m-1} x_{n+2m-1} & = & b_{m\min} \\
\sum_{i=1}^{n} a_{im}x_i & + & (-1)^{2m} x_{n+2m} & = & b_{m\max} \\
x_1 & + & x_{n+2m+1} & = & w_1 \\
x_2 & + & x_{n+2m+2} & = & w_2 \\
\vdots & & \vdots & & \vdots \\
x_n & + & x_{2n+2m} & = & w_n \\
\sum_{i=1}^{n} x_j & & & = & 1
\end{array} \qquad (4.9\text{-}21)$$

式（4.9-20）为目标函数，式（4.9-21）为约束方程。

$[x_1\ x_2 \cdots x_n]^{\mathrm{T}}$ 为解向量，$[x_{n+1}\ x_{n+2} \cdots\ x_{2n+2m}]^{\mathrm{T}}$ 为松弛变量或剩余变量。

7.4 使用单纯形法解决配料问题的具体步骤

用单纯形法解决 LP 问题的过程就是在凸集中寻找顶点的过程。针对线性规划的标准形式［如式（4.9-20）和式（4.9-21）］，可使用单纯形表解决此种形式的 LP 问题。

在单纯形法解决 LP 问题的具体步骤中，未必在一开始就能找到初始基本可行解，为了解决这个问题，在单纯形法的基础上发展了人工变量法，即在约束方程中添加人工变量，并相应的处理目标函数，使人工变量在迭代结束后为零，其中最典型的最为常用的是大 M 法。

大 M 法是在约束方程中添加人工变量，取一个足够大的系数 M，与人工变量相乘，加到目标函数中，从而在单纯形法迭代的过程中，保证第一步存在一个初始基本可行解，而在迭代结束后，人工变量全为零。

根据大 M 法对原来的标准形式修改如式（4.9-22）和式（4.9-23）。

$$\min S = \sum_{i=1}^{n} C_j x_j + \sum_{k=1}^{n+2m+1} M x_{2n+2m+k} \qquad (4.9\text{-}22)$$

$$\begin{array}{ccccccc}
\sum_{i=1}^{n} a_{1i}x_i & + & (-1)^1 x_{n+1} & + & x_{2n+2m+1} & = & b_{1\min} \\
\sum_{i=1}^{n} a_{1j}x_i & + & (-1)^2 x_{n+2} & + & x_{2n+2m+2} & = & b_{1\max} \\
\vdots & & \vdots & & \vdots & & \\
\sum_{i=1}^{n} a_{ij}x_i & + & (-1)^{2j-1} x_{n+2j-1} & + & x_{2n+2m+2i-1} & = & b_{j\min} \\
\sum_{i=1}^{n} a_{ij}x_i & + & (-1)^{2j} x_{n+2j} & + & x_{2n+2m+2i} & = & b_{j\max} \\
\vdots & & \vdots & & \vdots & & \\
\sum_{i=1}^{n} a_{im}x_i & + & (-1)^{2m-1} x_{n+2m-1} & + & x_{2n+4m-1} & = & b_{m\min} \\
\sum_{i=1}^{n} a_{im}x_i & + & (-1)^{2m} x_{n+2m} & + & x_{2n+4m} & = & b_{m\max} \\
x_1 & + & x_{n+2m+1} & + & x_{2n+4m+1} & = & w_1 \\
x_2 & + & x_{n+2m+2} & + & x_{2n+4m+2} & = & w_2 \\
\vdots & & \vdots & & \vdots & & \\
x_n & + & x_{2n+2m} & + & x_{3n+4m} & = & w_n \\
\sum_{i=1}^{n} x_i & + & x_{3n+4m+1} & & & = & 1
\end{array}$$

$$(4.9\text{-}23)$$

式中，M 为一个在计算过程中足够大的数；$[x_1\ x_2 \cdots\ x_n]^{\mathrm{T}}$ 为解向量；$[x_{n+1}\ x_{n+2} \cdots x_{2n+2m}]^{\mathrm{T}}$ 为松弛变量或剩余变量；$[x_{n+2m+1}\ x_{n+2m+2} \cdots\ x_{2n+4m+2}]^{\mathrm{T}}$ 作为人工变量，在迭代开始时，作为初始可行基，在迭代结束后，退化为零解。

使用单纯形表利用大 M 法解 LP 问题，是一个常用的方法，但毕竟是一个手工的方法，费时费力，随着计算机在工程应用中的日益广泛，这个问题得到了更好的解决，下面就使用程序语言解决 LP 问题作一个简要的介绍。

利用上面所述的 LP 问题的标准形式，建立一个向量表，如表 4.9-2 所示。

表 4.9-2 优化配料问题解决过程中使用的原始向量表

约束条件			松弛变量和剩余变量								人工变量										约束函数值
x_1	⋯	x_n	x_{n+1}	x_{n+2}	⋯	x_{n+2m-1}	x_{n+2m}	x_{n+2m+1}	⋯	x_{2n+2m}	$x_{2n+2m+1}$	$x_{2n+2m+2}$	⋯	$x_{2n+4m-1}$	x_{2n+4m}	$x_{2n+4m+1}$	⋯	x_{3n+4m}	$x_{3n+4m+1}$		
a_{11}	⋯	a_{1n}	−1								1										$b_{1\min}$
a_{11}	⋯	a_{1n}		1								1									$b_{1\max}$
⋮		⋮			⋱								⋱								⋮
a_{m1}	⋯	a_{mn}				−1								1							$b_{m\min}$
a_{m1}	⋯	a_{mn}					1								1						$b_{m\max}$
1								1								1					w_1
	⋱								⋱								⋱				⋮
		1								1								1			w_n
1	⋯	1																	1		1
C_1	⋯	C_n	0	0	⋯	0	0	0	⋯	0	M	M	⋯	M	M	M	⋯	M	M		S

注：向量表的第一行是为了表述方便，在建立向量表时，并不包括这一行。向量表中空处元素为零，在此省略。

将向量表 4.9-2 按照下面的格式分为五个区域，如表 4.9-3 所示，其中 x_1，x_2，⋯，$x_{3n+4m+1}$ 作为决策变量，$\begin{bmatrix} \alpha_{1\ 1} & \cdots & \alpha_{1\ 3n+4m+1} \\ \vdots & & \vdots \\ \alpha_{2m+n+1\ 1} & \cdots & \alpha_{2m+n+1\ 3n+4m+1} \end{bmatrix}$为表 4.9-2 中的决策变量、松弛变量、剩余变量、人工变量对应的系数矩阵，$[\beta_1, \beta_2, \cdots, \beta_{2m+n+1}]^{\mathrm{T}}$ 为表 4.9-2 中的约束函数值，$[\sigma_1, \sigma_2, \cdots, \sigma_{3n+4m+1}]$ 为表 4.9-2 中的最后一行，即目标函数对应的系数向量，S 为目标函数值。在迭代的过程中，同一个区域中的元素都用相同的方式进行处理，并把 $[\sigma_1, \sigma_2, \cdots, \sigma_{3n+4m+1}]$ 中的元素 σ_j（$j=1, 2, \cdots, 3n+4m+1$）作为检验数来判断迭代的进行过程。

表 4.9-3 原始向量表的压缩形式

x_1	⋯	$x_{3n+4m+1}$	
$\alpha_{1\ 1}$	⋯	$\alpha_{1\ 3n+4m+1}$	β_1
⋮		⋮	⋮
$\alpha_{2m+n+1\ 1}$	⋯	$\alpha_{2m+n+1\ 3n+4m+1}$	β_{2m+n+1}
σ_1	⋯	$\sigma_{3n+4m+1}$	S

迭代的具体步骤如下。

1）判断检验数 σ_j（$j=1, 2, \cdots, 3n+4m+1$）的大小，如果所有的检验数都小于等于零，则已经得到最优解，否则进行下一步。

2）如果存在 $\sigma_j \geqslant 0$，并且与这个 σ_j 相对应的 x_j 的系数向量小于等于零，即 $P_j = \begin{bmatrix} \alpha_{1\ j} \\ \alpha_{2\ j} \\ \vdots \\ \alpha_{i\ j} \\ \vdots \\ \alpha_{2m+n+1\ j} \\ \sigma_j \end{bmatrix} \leqslant \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$（$i=1, 2, \cdots, 2m+n+1$，均有 $\alpha_{ij} \leqslant 0$），则此问题无解，停止计算，否则，转入下一步。

3）根据 $\max(\sigma_j > 0) = \sigma_k$ 确定 x_k 为换入变量，根据 θ 规则 $\theta = \min\left(\dfrac{\beta_i}{\alpha_{ik}} \mid \alpha_{ik} > 0\right) = \dfrac{\beta_i}{\alpha_{ik}}$，得到主元素 α_{ik} 转入下一步。

4）在表中以 α_{ik} 为主元素对系数矩阵 $\begin{bmatrix} \alpha_{1\ 1} & \cdots & \alpha_{1\ 3n+4m+1} \\ \vdots & & \vdots \\ \alpha_{2m+n+1\ 1} & \cdots & \alpha_{2m+n+1\ 3n+4m+1} \end{bmatrix}$进行矩阵行变换，把 x_k 所对应的列向量 $P_k = \begin{bmatrix} \alpha_{1\ k} \\ \alpha_{2\ k} \\ \vdots \\ \alpha_{i\ k} \\ \vdots \\ \alpha_{2m+n+1\ k} \\ \sigma_k \end{bmatrix}$转化为$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$。

5）然后转到第一步，重复迭代，直至满足得到最优解的条件。

迭代完毕，所有的主元素 α_{ik} 对应的 x_k 的最优解为 β_i，其他变量 x_j 的最优解为零。这样就得到了最优解向量。这种方法用到铸铁熔炼中常常获得比较满意的结果。

编写：田学雷（山东大学）

第 10 章　焊接工艺优化设计

优化设计不仅为工程设计提供了一种重要的科学设计方法，使得在解决复杂的设计问题时，能从众多的设计方案中找到尽可能完善的、最适宜的设计方案，而且能够大大提高设计效率和设计质量，获得明显的经济效益和社会效益。

目前优化技术在焊接领域中主要应用在焊接冶金和材料优化设计、焊接产品优化设计、焊接工艺参数优化设计、焊接自动化优化设计、焊接生产规划设计、焊接生产工艺过程的优化设计、焊接生产备料的优化设计等。

1　焊接材料优化设计

1.1　焊条配方的优化设计

1.1.1　焊条优化设计的一般步骤

1）建立优化设计的数学模型　建立数学模型是设计和改进药皮配方的关键环节之一。数学模型必须正确反映出药皮组分对焊条性能的影响规律。焊条优化设计通常采用回归设计法建模，为了得到准确的数学模型，必须合理设计试验方案，严格试验条件，科学地处理试验数据。

2）确定优化方法，计算最优配方　焊条的优化问题一般属于非线性规划问题，应选用有约束优化设计方法求解。罚函数法是求解这类问题的有效方法，并且在焊条的优化设计中已得到成功的应用。优化方法确定后，编制计算程序和上机计算。计算时应尽量多选几个初始点，以便获得全局最优解。

3）分析审查计算结果　检查计算出的最优配方焊条的性能是否满足设计要求，然后根据药皮组分对各项性能的影响规律和生产成本等，综合评价所求药皮配方的合理性及可行性。如果满足以上两项要求，可转入下一步。否则，调整计算参数重新计算，直到获得满意结果。

4）试验验证和实施　对计算求出的最优配方，要进行试验验证。如果试验结果与计算结果相符，说明设计方案正确，可将优化的配方投入生产。如果误差超过允许的精度，分析原因，适当调整配方，进一步试验验证，直到满足要求。

1.1.2　焊条优化设计实例

欲研制一种超低氢高韧性低合金高强钢焊条，对熔敷金属力学性能的要求为：$\sigma_s \geqslant 580$ MPa，$\sigma_b \geqslant 650$ MPa，$\delta_5 \geqslant 25\%$，$\psi \geqslant 65\%$。采用合金元素裂纹敏感系数：

$$P_{cm} = C + \frac{Si}{30} + \frac{Mn + Cu + Cr}{20} + \frac{Ni}{60} + \frac{Mo}{15} + \frac{V}{10} + 5B \tag{4.10-1}$$

评价裂纹敏感性，为了保证熔敷金属的抗裂性能，要求 $P_{cm} \leqslant 0.175\%$。

（1）方案设计

根据对熔敷金属的性能要求，选用 $CaO\text{-}SiO_2\text{-}CaF_2$ 渣系和直径为 $\phi 4$ mm 的 H08A 焊芯，合金系统为 Mn-Ni-Cr-Mo 系，合金元素通过药皮过渡。

首先确定初步配方，进行工艺试验；然后在工艺性能基本满足的基础上，固定药皮其他组分，重点考察锰铁（Z1）、镍粉（Z2）、金属铬（Z3）、钼铁（Z4）、硅钙 + 1 号稀土（Z5）和焊芯含碳量（Z6）对熔敷金属力学性能的影响。采用一次回归正交设计安排试验计划，选用 L16（215）正交表。根据经验确定各考察因子的上下水平，并计算出它们的编码值。每个因子的水平按式（4.10-2）进行编码：

$$x_j = \frac{Z_j - Z_{0j}}{\Delta_j} \quad (j = 1,\ 2,\ \cdots,\ 6) \tag{4.10-2}$$

式中，$Z_{0j} = (Z_{2j} + Z_{1j})/2$，$\Delta_j = (Z_{2j} - Z_{1j})/2$。

编码结果如表 4.10-1 所示，$Z_1 \sim Z_6$ 的单位为质量分数。在零点进行两次重复性试验。

表 4.10-1　因子水平编码表

因子		Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
上水平	Z_{2j}	7.4	6.0	1.0	2.3	5.5	0.082
	x_{2j}	1	1	1	1	1	1
下水平	Z_{1j}	3.4	2.0	0.28	0	3.0	0.045
	x_{1j}	-1	-1	-1	-1	-1	-1
零水平	Z_{0j}	5.4	4.0	0.64	1.15	4.25	0.063 5
	x_{0j}	0	0	0	0	0	0
变动区间	Δ_j	2.0	2	0.36	1.15	1.25	0.018 5

按上述设计的方案可得到 16 种焊条的合金配方，制出的焊条按 GB/T 5118—1995 规定焊接熔敷金属试板，进行拉伸和冲击等力学性能试验。

（2）建立回归方程

采用一次正交回归分析程序，对试验结果进行统计处理，求出熔敷金属力学性能指标与编码因子之间的回归方程，并对回归系数和方程进行统计检验，结果如下：

$$\sigma_s = 63.12 + 4.90x_1 + 4.11x_2 + 1.93x_3 + 5.06x_4 + 1.59x_5 + 2.6x_6 - 0.45x_1x_4 - 0.74x_1x_5 - 0.55x_4x_6 \ (\times 9.8\ \text{MPa})$$

$$\sigma_b = 70.12 + 4.36x_1 + 3.66x_2 + 1.55x_3 + 4.22x_4 + 1.54x_5 + 2.62x_6 - 0.43x_1x_3 - 0.50x_1x_4 - 0.50x_1x_5 - 0.33x_2x_4 - 0.41x_1x_5 \ (\times 9.8\ \text{MPa})$$

$$\delta_5 = 23.05 - 1.81x_1 - 1.37x_2 - 0.85x_3 - 1.36x_4 - 0.28x_5 - 1.13x_6 - 0.30x_2x_4 - 0.50x_2x_5 - 0.26x_4x_6 \ (\%)$$

$$\psi = 67.69 - 3.05x_1 - 2.77x_2 - 2.22x_3 - 2.84x_4 - 2.31x_5 - 2.34x_6 - 1.61x_2x_5 - 1.71x_3x_6 - 1.65x_4x_6 - 1.82x_5x_6 \ (\%)$$

$$P_{cm} = 0.192 + 17.3 \times 10^{-3}x_1 + 9.88 \times 10^{-3}x_2 + 4.5 \times 10^{-3}x_3 + 15 \times 10^{-3}x_4 + 3.75 \times 10^{-3}x_5 + 11.1 \times 10^{-3}x_6 \ (\%)$$

$$A_{kV} = 5.23 - 0.93x_1 - 0.51x_2 - 0.27x_3 - 1.22x_4 - 0.66x_5 - 1.55x_6 + 0.49x_1x_4 + 0.42x_2x_4 + 0.42x_3x_6 + 0.32x_4x_6 \ (\times 9.8\ \text{J})$$

回归方程 $\psi = f(x)$ 的置信度为 90%，其余回归方程的置信度均为 99%。这说明上述方程正确反映了熔敷金属力学性能与考察因子编码值之间的定量关系。

（3）熔敷金属冲击性能的优化

为防止焊接构件发生低温脆性破坏，常要求熔敷金属在满足强度和塑性指标的前提下，尽可能提高其低温韧性。因此，将冲击力 A_{kV} 作为优化目标，把求 A_{kV} 最大值改为求其最小值，则该最优化问题的数学模型为：

$$\min f(x) = -A_{kV}$$

$$g_1(x) = \sigma_s - 580 \geqslant 0$$

$$g_2(x) = \sigma_b - 650 \geqslant 0$$

$$g_3(x) = \delta_5 - 25 \geqslant 0$$

$$g_4(x) = \psi - 65 \geqslant 0$$

$$g_5(x)=0.175-P_{cm}\geqslant 0$$
$$g_6(x)=x_1+1\geqslant 0$$
$$g_7(x)=1-x_1\geqslant 0$$
$$g_8(x)=x_2+1\geqslant 0$$
$$g_9(x)=1-x_2\geqslant 0$$
$$g_{10}(x)=x_3+1\geqslant 0$$
$$g_{11}(x)=1-x_3\geqslant 0$$
$$g_{12}(x)=x_4+1\geqslant 0$$
$$g_{13}(x)=1-x_4\geqslant 0$$
$$g_{14}(x)=x_5+1\geqslant 0$$
$$g_{15}(x)=1-x_5\geqslant 0$$
$$g_{16}(x)=x_6+1\geqslant 0$$
$$g_{17}(x)=1-x_6\geqslant 0$$

这是一个不等式约束条件下非线性规划的优化问题，宜采用外点罚函数法求解，罚函数为：

$$P(x,r_k)=-A_{kV}+r_k\sum_{1}^{17}\{\min[0,g_i(x)]\}^2 \quad (4.10\text{-}3)$$

式中，r_k 为第 k 步迭代时所取的罚因子。

取 $r_0=1$，$r_{k+1}=Cr_k$，$C=5$，收敛精度 $\varepsilon=1\times10^{-5}$。选用单纯形加速法求罚函数 $P(x,r_k)$ 的无约束极值。输入若干个初始点，把计算结果加以比较，取全局最优者。优化结果为：满足约束条件的目标函数最大值 $A_{kV}=77.0$ J；最优编码因子 $x_1=0.193\ 3$，$x_2=0.624\ 2$，$x_3=-0.989\ 8$，$x_4=-0.555\ 1$，$x_5=-0.719\ 1$，$x_6=-0.999\ 7$。

(4) 试验验证

根据优化结果制成焊条，焊后按标准进行力学性能试验，结果如表4.10-2所示。可以看出，结算结果和试验结果吻合很好。

表4.10-2 优化焊条熔敷金属力学性能

力学性能 / 类别	σ_s /MPa	σ_b /MPa	δ_5 /%	ψ /%	A_{kV} (-50℃) /J
优化值	580	656	25.02	69.17	77.0
试验值	589	650	24.00	69.48	78.5

1.2 药芯焊丝的计算机辅助优化设计

药芯焊丝配方设计是决定药芯焊丝内在质量的关键环节，而配方设计是一个多因子试验设计问题，各因子之间存在着复杂的交互作用，并没有固定的数学模型，因此，在配方设计时常采用经验法或单因子轮换法寻找理想配方。这种设计方法在设计新焊丝时往往带有盲目性。根据药芯焊丝性能的要求，首先确定熔敷金属中元素数量、质量分数及过渡系数，然后采用优化技术设计开发这类产品，不仅可以提高产品质量、缩短生产周期，而且可以降低成本。

1) 药芯焊丝中元素 i 的加入量 w_i　根据熔敷金属中 i 元素的设计质量分数及其过渡系数，药芯焊丝中 i 元素的加入量为：

$$w_i=\frac{w_{i,d}}{\eta_i} \quad (4.10\text{-}4)$$

式中，w_i 为药芯焊丝中 i 元素的质量分数；$w_{i,d}$ 为熔敷金属中 i 元素的质量分数；η_i 为 i 元素的合金过渡系数。

2) 药芯焊丝合金组分的质量分数　药芯焊丝中合金组分的质量分数为：

$$w_{i,cw}=\frac{w_i}{w_{i,c0}} \quad (4.10\text{-}5)$$

式中，$w_{i,cw}$ 为药芯焊丝中 i 组分的加入量；$w_{i,c0}$ 为 i 元素在该组分中的质量分数。

3) 组分总质量 m_t 与总体积 V_t

总质量：
$$m_t=\sum_{i=1}^{n}w_{i,cw} \quad (4.10\text{-}6)$$

总体积：
$$V_t=\sum_{i=1}^{n}\frac{w_{i,cw}}{r_i} \quad (4.10\text{-}7)$$

式中，r_i 为 i 组分的松装比。

4) 钢带断面积与配料面积比 f　根据总质量与总体积确定药芯焊丝的钢带面积与配料面积的比值 f：

$$f=\frac{100-\sum w_{i,cw}}{\rho_s K_f K_c\sum\frac{w_{i,cw}}{r_i}} \quad (4.10\text{-}8)$$

式中，ρ_s 为钢带密度；K_f，K_c 为考虑填充情况和组分粒度的系数。

5) 根据设计要求选择焊丝直径 d_{fm}、卷丝模直径 d_d、钢带厚度 δ 和钢带宽度 b。

6) 药芯焊丝外皮几何断面积与填充面积比　根据选择的钢带厚度 δ 和卷丝模直径 d_d，计算药芯焊丝外皮几何断面积与填充面积比 f'

$$f'=\frac{d_d^2-(d_d-2K\delta)^2}{(d_d-2K\delta)^2} \quad (4.10\text{-}9)$$

式中，K 为考虑卷制焊丝时壁厚变化的系数。

比较 f 与 f'，如果 $\frac{|f-f'|}{f}<5\%$，那么选择的钢带尺寸是合理的。则填充系数 K（%）为：

$$K=\frac{\sum w_{i,cw}}{f\rho_s K_c\sum\frac{w_{i,cw}}{r_i}+\sum w_{i,cw}} \quad (4.10\text{-}10)$$

100 g药芯焊丝中消耗的钢带的质量 w_{st} 为：

$$w_{st}=100-K\times100 \quad (4.10\text{-}11)$$

如果 $\frac{|f-f'|}{f}>5\%$，且 $f>f'$，则说明在该种填充系数下，药粉填不满管芯，在轧制时药粉有可能在管芯内窜动，导致药芯的成分不均。为防止出现这种现象，可采用以下两种方法改进。

① 添加铁粉或其他中性剂，增大填充系数　如果添加铁粉，则按式（4.10-12）计算填充系数 K：

$$K=\frac{1-f\rho_s K_c\left(\sum\frac{w_{i,cw}}{r_i}-\sum\frac{w_{i,cw}}{r_{ip}}\right)}{1+\frac{f\rho_s K_c}{r_{ip}}} \quad (4.10\text{-}12)$$

式中，r_{ip} 为铁粉的松装密度。

填充铁粉量为：

$$w_{ip}=100\times K-\sum w_{i,cw} \quad (4.10\text{-}13)$$

根据式（4.10-6）～式（4.10-9）重新进行核算，直至 $\frac{|f-f'|}{f}<5\%$，则设计完成。否则重新调整铁粉的加入量。

② 调整钢带尺寸重新计算 f'，直至 $\frac{|f-f'|}{f}<5\%$，再计算 K 和 w_{st}，设计完成。

1.3 焊剂优化设计

焊剂配方优化设计时要确定混料中各成分因子最佳质量分数，首先要选择恰当的试验设计方案进行试验分析；其次要根据目标要求及试验分析结果，选用科学的数学模型及参数估计方法建立目标函数，最后要在一定的约束条件下对目标函数进行寻优，确定混料中各成分因子最佳质量分数。

1.3.1 焊剂优化设计一般过程

(1) 试验设计方案的选择

焊剂配方设计属于混料试验设计。混料试验设计方法有多种，应以试验处理少、统计分析方法简便为原则进行选择，合理地设计混料试验，求取各种线性或非线性方程。实际上，焊剂配方的设计是一种受特殊条件约束的回归设计，即在混料条件约束下，合理安排试验，寻求试验指标 $f(x)$ 与成分因子 X_j $(j=1, 2, \cdots, p)$ 的定量关系的回归设计方法。其中单形重心设计是目前最适宜的多因子成分混料试验设计。

(2) 目标函数的建立

焊剂优化设计的目标函数的数学模型为：

$$f(x) = \sum_{i=1}^{p} b_i x_i + \sum_{i<j}^{p} b_{ij} x_i x_j + \sum_{i<j<k}^{p} b_{ijk} x_i x_j x_k + \cdots + \sum_{i_1<i_2<\cdots i_d}^{p} b_{i_1 i_2 \cdots i_d} x_1 x_2 \cdots x_d \tag{4.10-14}$$

式中，$b_i b_{ij}$ $(i<j)$，b_{ijk} $(i<j<k)$，$\cdots$，$b_{i_1 i_2 \cdots i_d}$ $(i_1<i_2\cdots<i_d)$ 为待估参数；x_1，x_2，$\cdots$，x_p 为成分因子 X_1，X_2，$\cdots$，X_p 的质量分数；$f(x)$ 为理论目标值；d 为统计模型的次数。

(3) 约束条件

焊剂配方混料问题的约束条件为：

$$\left.\begin{array}{l} \sum_{i}^{p} x_i = 1 \\ 0 \leqslant x_1, x_2, \cdots, x_p \leqslant 1 \end{array}\right\} \tag{4.10-15}$$

(4) 计算求解

对目标函数寻优确定焊剂中各成分因子最佳质量分数，就是确定规划的最优解问题：

$$\min(\max) f(x) = \sum_{i=1}^{p} b_i x_i + \sum_{i<j}^{p} b_{ij} x_i x_j + \sum_{i<j<k}^{p} b_{ijk} x_i x_j x_k + \cdots + \sum_{i_1<i_2<\cdots i_d}^{p} b_{i_1 i_2 \cdots i_d} x_1 x_2 \cdots x_d \tag{4.10-16}$$

这一问题是非线性的，可用线性逼近的方法求解，迭代步骤如下。

1) 把数学模型标准化　若模型是寻求目标函数的最小值，则该模型已为标准型；若模型是寻求目标函数的最大值，则需把寻求最大转化为寻求最小，即：

$$\min Z = -f(x) = \sum_{i=1}^{p} b_i x_i - \sum_{i<j}^{p} b_{ij} x_i x_j - \sum_{i<j<k}^{p} b_{ijk} x_i x_j x_k - \cdots - \sum_{i_1<i_2<\cdots i_d}^{p} b_{i_1 i_2 \cdots i_d} x_1 x_2 \cdots x_d \tag{4.10-17}$$

记标准化后的模型为：

$$\left.\begin{array}{l} \min Z = G(x) = G(x_1, x_2, \cdots, x_p) \\ \sum_{i}^{p} x_i = 1 \\ 0 \leqslant x_i \leqslant 1 \quad (i=1, 2, \cdots, p) \end{array}\right\} \tag{4.10-18}$$

2) 取定允许误差 $\varepsilon > 0$，一般取 $\varepsilon = 0.01$。

3) 选取迭代初始点。

4) 运用单纯形，确定线性规划问题：

$$\left.\begin{array}{l} \min S = \nabla G(X^{(k)})^{\mathrm{T}} X \\ \sum_{i}^{p} x_i = 1 \\ 0 \leqslant x_i \leqslant 1 \quad (i=1, 2, \cdots, p) \end{array}\right\} \tag{4.10-19}$$

的最优解 $\overline{X}^{(k)} = (\overline{x}_1^{(k)}, \overline{x}_2^{(k)}, \cdots, \overline{x}_p^{(k)})^{\mathrm{T}}$

式中，$X = (x_1, x_2, \cdots, x_p)^{\mathrm{T}}$，$\nabla G(X^{(k)}) = \left(\frac{\partial G(X^{(k)})}{\partial x_1}, \frac{\partial G(X^{(k)})}{\partial x_2}, \cdots, \frac{\partial G(X^{(k)})}{\partial x_p}\right)^{\mathrm{T}}$。

5) 检验 $\overline{X}^{(k)} = (\overline{x}_1^{(k)}, \overline{x}_2^{(k)}, \cdots, \overline{x}_p^{(k)})^{\mathrm{T}}$ 是否满足收敛准则 $|\nabla G(X^{(k)})^{\mathrm{T}}(X^{(k)} - \overline{X}^{(k)})| \leqslant \varepsilon$。若满足则停止迭代，此时 $X^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_p^{(k)})$ 就是最优解，否则转向下一步继续迭代。

1.3.2　应用实例

欲配制含 MgO、Al_2O_3、CaF_2、$CaCO_3$、SiO_2、TiO_2、ZrO_2 等组分的陶质焊剂，使焊剂具有较高软化温度 $T_{软}$、较大的线胀系数 $|\Delta\alpha|$ 与合适的熔点。

(1) 设计方案

预备试验时，焊剂各组分及其变化范围见表 4.10-3。采用对称单纯形设计，得到 43 种配方设计方案的试验方案。

表 4.10-3　焊剂组分设计变量

序号	组分	变量名	取值范围（质量分数）/%
1	MgO	x_1	10～20
2	Al_2O_3	x_2	20～30
3	CaF_2	x_3	13～23
4	$CaCO_3$	x_4	13～23
5	SiO_2	x_5	0～10
6	ZrO_2	x_6	0～10
7	TiO_2	x_7	10～20

(2) 回归方程的建立

对 43 种配方，分别配制一种焊剂，测试各种焊剂的软化温度、熔点、线胀系数，利用数理统计方法，建立软化温度、熔点以及线胀系数的回归方程为：

$$T_{软} = -1\,209x_1 + 1\,540x_2 + 385x_3 + 609x_4 + 1\,907x_5 + 1\,907x_6 + 5\,847x_1x_4 + 9\,124x_1x_7 + 6\,333x_1x_5 - 4\,861x_3^2 - 34\,156x_2x_5^2$$

$$T_{熔} = -932x_1 + 5\,966x_3 + 130x_4 + 1\,169x_6 + 6\,904x_1x_4 + 8\,308x_1x_7 + 2\,308x_1x_5 - 12\,659x_3^2 12\,027x_2x_5^2$$

$$\alpha = 262x_1 - 274x_2 + 211x_3 - 123x_4 - 492x_5 + 1\,168x_6 - 416x_7 + 129x_1x_2 - 840x_1x_3 - 185x_1x_4 - 112x_1x_5 - 170x_1x_6 + 98x_1x_7 + 209x_2x_3 + 646x_2x_4 + 1\,018x_2x_5 - 1\,023x_2x_6 + 1\,110x_2x_7 - 174x_3x_4 - 16x_3x_5 - 1\,433x_3x_6 - 128x_3x_7 + 643x_4x_5 - 1\,189x_4x_6 + 775x_4x_7 - 842x_5x_6 - 1\,410x_5x_7 - 733x_6x_7$$

对回归方程进行方差分析，以检验回归方程的显著性，线胀系数的回归方程在 $\alpha = 0.053$ 水平下显著。软化温度、熔点的回归方程均在 0.01 水平下显著。

(3) 编程计算

对得到的三个回归方程，求解满足条件式 (4.10-20) 最佳焊剂配方。

$$\begin{cases} \max\ (\alpha - \alpha_0) \\ \max T_{软} \\ 1\,200℃ < T_{熔} < 1\,360℃ \end{cases} \tag{4.10-20}$$

利用单纯形法求得最佳焊剂成分配比为：

$$x_1 : x_2 : x_3 : x_4 : x_5 : x_6 : x_7 = 10:30:20:20:1:5:10$$

将此设计代入回归方程，得到最佳焊剂物理性能的理论估计值为：

$T_{软} = 1\,281℃$，$T_{熔} = 1\,357℃$，$\alpha = 7.8 \times 10^{-6}/℃$

(4) 优化配方的试验检验

利用理论计算最佳焊剂成分配比配成焊剂，进行试验检验，结果见表 4.10-4。

表 4.10-4 优化设计焊剂配方性能检验

试验	性能			
	$\Delta\alpha/10^{-6}\cdot K^{-1}$	$T_{软}$/℃	$T_{熔}$/℃	脱渣率/%
理论值	7.8	1 281	1 357	100%
实测值	8.3	1 282	1 340	
相对误差	6.4%	0.08%	1.2	

1.4 埋弧焊焊丝和焊剂优化选配

1.4.1 选配原则

根据被焊钢材的类别及对焊接接头性能的要求选择焊丝，并选择适当焊剂相配合。一般情况下，对低碳钢、低合金高强钢的焊接，应选用与母材强度匹配的焊丝；对耐热钢、不锈钢的焊接，应选用与母材成分相匹配的焊丝；堆焊时应根据对堆焊层的技术要求、使用性能等，选择合金系统及相近成分的焊丝并选用合适的焊剂。

还应根据焊接产品的技术要求（如坡口和接头形式、焊后加工工艺等）和生产条件，选择合适的焊剂与焊丝的组合，必要时应进行焊接工艺评定，检测焊缝金属的力学性能、耐腐蚀性、抗裂性以及焊剂工艺性，以考核所选焊材是否合适。

1.4.2 埋弧焊焊丝与焊剂优化选配

（1）低碳钢埋弧焊焊丝与焊剂选配

常用低碳钢的埋弧焊焊剂与焊丝的优选组合见表 4.10-5。

表 4.10-5 常用低碳钢的埋弧焊焊丝与焊剂的选配

钢号	烧结焊剂与配用焊丝		熔炼焊剂与配用焊丝	
	焊剂	焊丝	焊剂	焊丝
Q235	SJ401、SJ403、SJ402（薄板、中厚板）	H08A、H08E	HJ431、HJ430	H08A、H08MnA
Q255				
Q275				
15、20	SJ301、SJ302、SJ502、SJ501、SJ503（中厚板）	H08A、H08E、H08MnA	HJ431、HJ430、HJ330	H08A、H08MnA
25、30				H08MnA、H10Mn2
20g、22g				H08MnA、H08MnSi、H10Mn2
20R				H08MnA

（2）低合金钢埋弧焊用焊丝与焊剂优选

埋弧焊焊接的低合金钢，主要是热轧正火钢。选用焊丝与焊剂配合时应保证焊缝金属的力学性能，应选用与母材强度相当的焊接材料，并综合考虑焊缝金属的冲击韧度、塑性及焊接接头的抗裂性。焊缝金属的强度不宜过高，通常控制在不低于或略高于母材强度，过高会导致焊缝金属韧度、塑性及焊接接头抗裂性能降低。对调质钢，为避免热影响区韧度和塑性的降低，一般不宜采用粗丝、大电流、多丝埋弧焊，采用陶质焊剂 572F－6＋HJ350 的混合焊剂（其中 HJ350 占 80%～82%），配合 H18CrMoA 焊丝可实现 30CrMnSiNi2A 的埋弧焊。

常用热轧、正火低合金钢埋弧焊焊剂与焊丝的配合选用见表 4.10-6。

表 4.10-6 常用热轧、正火低合金钢埋弧焊焊剂与焊丝优选

钢号	焊丝	焊剂	备注
09Mn2、09Mn2Si、09Mn	H081、H08MnA	HJ430、HJ431、SJ301	
16Mn、16MnR、16MnCu、14MnNb	H08Mn、H08MnA	SJ501、SJ502	用于薄板
	H08A	HJ430、HJ431、SJ301	用于不开坡口对接
	H08MnA、H10Mn2	HJ430、HJ431、SJ301	用于中板开坡口对接
	H10Mn2、H08MnMoA	HJ350	用于厚板深坡口
15MnV、15MnVCu、16MnNb、15MnVR	H08MnA	HJ430、HJ431	用于不开坡口对接
	H10Mn2、H10MnSi	HJ430、HJ431	用于中板开坡口对接
	H08MnMoA	HJ250、HJ350、SJ101	用于厚板深坡口
15MnVN、15MnVNCu 15MnVTiRE、15MnVNR	H10Mn2	HJ431	
	H08MnMoA、H08Mn2MoA	HJ350、HJ250、HJ252、SJ101	
18MnMoNb、14MnMoV、14MnMoVCu、14MnMoVg 18MnMoNbg、18MnMoNbR	H08MnMoA	SJ102	
	H08Mn2MoA、H08Mn2MoVA、H08Mn2NiMo	HJ250、HJ252、HJ350、SJ101	
X60 低合金管线钢	H08Mn2MoA	HJ431	
	H08MnMoA	SJ101	
	H10Mn2	SJ102	
X65 低合金管线钢	H08MnMoA	SJ102、SJ301	
	H08Mn2MoA	SJ101	

(3) 低温钢埋弧焊焊丝与焊剂优选

低温钢要求在较低的使用温度下具有足够的韧度及抗脆性破坏能力。因此，应选用碱性焊剂，焊丝应严格控制含碳量，S、P含量应尽量低。选用烧结焊剂配合Mn-Mo或含Ni焊丝；如采用C-Mn焊丝，应配合非熔炼焊剂，通过焊剂向焊缝过渡微量Ti、B等合金元素，以保证焊缝金属的低温韧度。焊接时采用较小的线能量，一般在28~45 kJ/cm。常用的低温钢埋弧焊焊丝与焊剂选用见表4.10-7。

表4.10-7 常用的低温钢埋弧焊焊丝与焊剂的优化选配

钢号	工作温度/℃	焊丝	焊剂
16MnDR	-40	H10MnNiMoA、H06MnNiMoA	SJ101、SJ603
DG50	-46	H10Mn2Ni2MoA	SJ603
09MnTiCuREDR	-60	H08MnA、H08Mn2	SJ102、SJ603
09Mn2VDR、2.5Ni钢	-70	H08Mn2Ni2A	SJ603
3.5Ni钢	-90	H05Ni3A	SJ603

(4) 耐热钢埋弧焊焊丝与焊剂优选

1) 低合金耐热钢埋弧焊焊丝与焊剂优选　焊丝与焊剂组合的基本原则是焊缝金属的化学成分、力学性能与母材基本一致或达到产品所要求的性能。为提高焊缝金属的抗热裂性能，焊丝的含碳量应控制在略低于母材的含碳量。常用的低合金耐热钢埋弧焊焊剂与焊丝的匹配见表4.10-8。

表4.10-8 常用的低合金耐热钢埋弧焊焊丝与焊剂的选用

钢种	钢号	焊丝	焊剂
0.5Mo	—	H08MnMoA	HJ350
0.5Cr-0.5Mo	12CrMo	H08CrMoA、H10CrMoA	HJ350、SJ103
1Cr-0.5Mo 1.25Cr-0.5Mo	15CrMo	H08CrMoA、H10CrMoA、H13CrMoA	HJ350、SJ103
1Cr-0.5MoV	12CrMoV	H08CrMoV	HJ350、HJ250、SJ103
2.25Cr-1Mo	Cr2Mo	H08Cr3MoMnA、H13Cr2Mo1A	HJ350、SJ103、SJ104
2Cr-MoWVTiB	12Cr2mOWVTiB	H08Cr2MoWVNbB	HJ250
Mn-Mo	14MnMoV 18MnMoNb	H08Mn2MoA	HJ350、SJ603、SJ101
Mn-Ni-Mo	13MnNiMoNb	H08Mn2NiMo	HJ350、SJ101、SJ603

2) 中合金耐热钢埋弧焊焊丝与焊剂选用　中合金耐热钢比低合金耐热钢具有更大的淬硬倾向，对焊接冷裂纹更为敏感，因此焊剂与焊丝的选用原则为：在保证焊接接头与母材具有相同的高温蠕变强度和抗氧化性的前提下，提高其抗裂性。厚壁工件的窄间隙焊接时应选用碱性焊剂或采用高碱度的烧结焊剂，如SJ601、SJ605、SJ103和SJ104等。焊丝选用高Cr-Ni奥氏体焊丝，能有效地防止焊接接头热影响区裂纹，或选用与母材成分基本相同的焊丝，可得到同质焊缝金属的接头，容易满足要求。

3) 高合金耐热钢埋弧焊焊丝与焊剂的选配　马氏体耐热钢埋弧焊焊丝与焊剂选配：马氏体耐热钢淬硬倾向大，防止裂纹是焊接的首要问题。应选用Cr含量与母材相同的同质焊丝，以保证高温使用性能，并选用高碱度低氢焊剂。对于常用的1Cr12、1Cr13马氏体耐热钢，采用的焊丝和焊剂配合为（H1Cr13、H0Cr14）+（SJ601、SJ605、SJ608）。

奥氏体耐热钢埋弧焊焊丝与焊剂选配：要求焊缝金属成分大致与母材成分匹配，同时应控制焊缝金属中的铁素体含量，对长期在高温下工作的焊件，焊缝中铁素体应不大于5%。奥氏体耐热钢埋弧焊的焊丝应选用低硅、低硫、低磷，或成分与母材相近。对Cr、Ni大于20%的奥氏体钢，为提高抗裂性，可选用高Mn（6%~8%）焊丝。焊剂应选用碱性或中性焊剂，以防止向焊缝增硅。常用奥氏体耐热钢埋弧焊焊丝与焊剂的选用见表4.10-9。

表4.10-9 常用的奥氏体耐热钢埋弧焊焊丝与焊剂的选用

钢　号	焊　丝	焊剂
0Cr19Ni9	H08Cr19Ni9、H0Cr21Ni10	SJ601 SJ605 SJ608 HJ260
1Cr18Ni9Ti	H1Cr19Ni10Nb、H0Cr21Ni10Ti	
0Cr18Ni11Ti、0Cr18Ni11Nb	H0Cr21Ni10Ti、H0Cr19Ni10Nb	
0Cr18Ni13Si4	H0Cr19Ni11Mo3	
1Cr20Ni14Si2	H1Cr25Ni13	
0Cr23Ni13	H1Cr25Ni13	
0Cr23Ni20	H1Cr25Ni20	
0Cr12Ni12Mo2	H0Cr19Ni11Mo3	
0Cr19Ni13Mo3	H0Cr19Ni11Mo3	

弥散硬化耐热钢埋弧焊焊丝与焊剂优选：埋弧焊可用于焊接厚度小于13 mm的弥散硬化耐热钢。如不要求焊缝金属与母材等强，可使用Cr-Ni奥氏体钢焊丝，否则必须使用特种焊丝和焊剂。特别是含Al、Ti等元素的钢，焊接时应采用无氧化性的焊剂。常用弥散硬化耐热钢埋弧焊焊丝与焊剂的选用见表4.10-10。

表4.10-10 常用弥散硬化耐热钢埋弧焊焊丝与焊剂的选用

钢　号	焊　丝	焊剂
S17400、SJ5500	H0Cr19Ni9	SJ601 SJ605 SJ608
1Cr17Ni7Al、X17H5M3、S3500	H1Cr25Ni20、ERNiCr-3、AWS5774B	
0Cr15Ni25Ti2MoAlVB、A-286 1Cr22Ni2Co20Mo3W3NbN	H1Cr25Ni12Mo3、H1Cr25Ni20、ERNiCrFe-6	

(5) 不锈钢埋弧焊焊丝与焊剂优选

1) 铁素体不锈钢埋弧焊焊丝与焊剂优选　焊接高铬铁素体不锈钢应注意晶间裂纹和脆性问题。常用高铬铁素体不锈钢埋弧焊焊丝与焊剂的优选见表4.10-11。

表4.10-11 常用铁素体不锈钢埋弧焊焊丝与焊剂的选用

钢　号	焊　丝	焊　剂
1Cr17、1Cr17Ti、1CR17Mo	H1Cr17、H0Cr21Ni10、H1Cr24Ni13、H0Cr26Ni21	SJ601、SJ701、SJ608、HJ171、HJ151
1Cr25Ti、1Cr28	H0Cr26Ni21、H1Cr26Ni21、H1Cr24Ni13	

2) 奥氏体不锈钢埋弧焊焊丝与焊剂优选 在焊接过程中铬、镍的烧损可通过焊剂和焊丝中合金元素的过渡来补充。焊接时应选择细焊丝和较小的焊接线能量。常用的奥氏体不锈钢埋弧焊焊丝与焊剂的优选见表 4.10-12。

表 4.10-12 常用奥氏体不锈钢埋弧焊焊丝与焊剂的选用

钢号	焊丝	焊剂
00Cr18Ni10N	H00Cr21Ni10	SJ601、SJ608、SJ701、SJ107、HJ172、HJ151
0Cr18Ni9、1Cr18Ni9	H0Cr21Ni10	
1Cr18Ni9Ti、0CR18Ni9Ti	H0Cr20Ni10Ti、H0Cr20Ni10Nb	
1CR18Ni12Mo2Ti	H0CR19Ni12Mo2	
0Cr18Ni12Mo2Ti	H00Cr19Ni12Mo2	
00Cr17Ni14Mo2	H00Cr18Ni14Mo2	
0Cr18Ni14MoCu2	H00Cr19Ni12Mo2Cu2	

2 焊接结构的优化设计

焊接结构设计是整个焊接结构生产中的首道工序，它决定了焊接生产的制造方向及技术要求。

焊接结构的形式主要有板式组合梁柱结构（工型、箱型、管型等），桁架组合梁柱结构（型钢或管件组合、焊接工型箱型管型组合等），容器结构（储罐、球罐等），壳体结构（车体、船体、机壳、钢结构场馆及厂房等），机体结构（基底、机座、机体等），单一管杆结构（管道、钢轨、钢筋等）。在保证经济和安全的条件下，根据使用中不同的受力性质、负荷和环境选用最优的结构材料，根据不同的结构形式选用最优的焊接方法，根据母材及相应的焊接方法选用最适宜的焊接材料、焊接工艺、焊接前后处理及质量检测方法。

另外一个问题就是在焊接结构设计中如何减少应力集中。不同的焊接接头形式、焊缝形状、结构转角或截面突变都会形成不同程度的应力集中，必须在结构设计中特别注意。

焊接残余应力在焊接结构中是不可避免的，有拉、压应力，且分布不均，但有规律，当与工作应力反向叠加时相当于减小工作应力。应利用其规律，使其变害为利，如不可能做到这点，而结构又有很高的防断要求，则需加以调整，进行减少焊接残余应力的焊后处理。

2.1 焊接变形与焊接应力

焊接时一般采用集中热源在局部加热，因此在焊件上产生不均匀温度场，不均匀的温度场使材料不均匀膨胀，处于高温区域的材料在加热过程中的膨胀量大，受到周围温度较低、膨胀量较小的材料的限制而不能自由伸展。结果焊件中出现内应力，使高温区的材料受到挤压，可使局部产生压缩塑性应变。在冷却过程中，已经受压缩塑性应变的材料，由于不能自由收缩而受到拉伸，冷却后，比周围区域相对缩短、变窄或减小，在这个区域呈现拉伸残余应力，周围区域承受压缩残余应力。存在残余应力时，也会产生焊接变形。

2.1.1 焊接残余变形种类及估算

在焊接结构生产中，残余变形不仅影响结构的尺寸精度和外观，而且可能降低其承载能力。

1) 纵向残余变形 单道焊缝的纵向收缩可由下式估算：

$$\Delta L = 0.86 \times 10^{-6} q_v L \quad (4.10\text{-}21)$$

式中，q_v 为焊接线能量，J/cm；L 为构件长度。

2) 横向收缩变形 手工焊对接接头的横向收缩量由下式估算：

$$\Delta B = 0.25 F_H / \delta \quad (4.10\text{-}22)$$

式中，F_H 为焊缝熔敷金属截面积；δ 为板厚。

T 形接头和搭接接头的横向收缩量 ΔB 随角焊缝焊脚 K 的增加而增大，随 δ 的增加而降低。其横向收缩的理论估算公式如下：

$$\Delta B = A \frac{\alpha q}{c\gamma v\delta} \quad (4.10\text{-}23)$$

式中，α 为线胀系数；q/v 为焊接线能量；c 为比热容；δ 为板厚；γ 为密度。

对于留有较大间隙且一次焊透的电渣焊接头 A 值取 1.6，对于一次焊透的电弧焊接头取 1.0～1.2，对于非焊透的情况如丁字接头、搭接接头则 A 值可由表 4.10-13 按焊接条件选取。

表 4.10-13 根据焊接条件的 A 值

焊接方法	焊接线能量 q_v/J·cm^{-1}	单位厚度线能量 q_{vs}（$=q/vs$）/J·cm^{-2}	A
交流电弧焊	约 57 500	≤46 300 >46 300	$0.06+0.203\times10^{-4}q_{vs}$ 1.0
	10 500～22 000	≤31 200 >31 200	$0.15+0.272\times10^{-4}q_{vs}$ 1.0
CO_2 气体保护焊	约 14 300	<8 400 8 400～19 300 >19 300	$0.15+0.272\times10^{-4}q_{vs}$ $-0.12+0.585\times10^{-4}q_{vs}$ 1.0
	约 11 100	<3 780 3 780～16 750 >16 750	$0.15+0.272\times10^{-4}q_{vs}$ $0.02+0.585\times10^{-4}q_{vs}$ 1.0
	约 8 800	<1 260 1 260～15 100 >15 100	$0.15+0.272\times10^{-4}q_{vs}$ $0.12+0.585\times10^{-4}q_{vs}$ 1.0

3) 弯曲变形 由焊缝纵向收缩引起的弯曲变形所造成的挠度 f 可按下式计算：

$$f = 0.86 \times 10^{-6} \times \frac{e q_v L^2}{8I} \quad (4.10\text{-}24)$$

式中，e 为焊缝塑性变形区中心（一般可取焊缝中心）与截面中性轴的距离；L 为构件长度（焊缝与构件等长）；I 为构件截面惯性矩。

由焊缝横向收缩引起的弯曲变形，如果横向焊缝在构件上分布不对称，例如图 4.10-1 构件上的短肋板焊缝，则每对肋板与翼缘之间的角焊缝的横向收缩 ΔB_2 将使构件弯曲一个角度：

$$\varphi_2 = \Delta B_2 \frac{S_2}{I} \quad (4.10\text{-}25)$$

式中，S_2 为翼缘对构件水平中性轴的静矩；$S_2 = F_2(h/2 - \delta_1/2)$；$F_2$ 为翼缘截面积。

每对肋板与腹板之间的角焊缝引起的横向收缩 ΔB_1 将使构件弯曲一个角度：

$$\varphi_1 = \Delta B_1 \frac{S_1}{I} \quad (4.10\text{-}26)$$

式中，S_1 为高度为 h_1 的一部分腹板对构件水平中性轴的静矩，$S_1 = h_1 \delta_2 e$。

若 $\varphi = \varphi_1 + \varphi_2$，则对于图 4.10-1 中构件的总挠度可按下式估算：

$$f = 5\varphi l + 4\varphi l + 3\varphi l + 2\varphi l + \varphi l$$

如果构件的中心有一肋板，则它所引起的挠度为：

$$f_0 = \varphi l/4$$

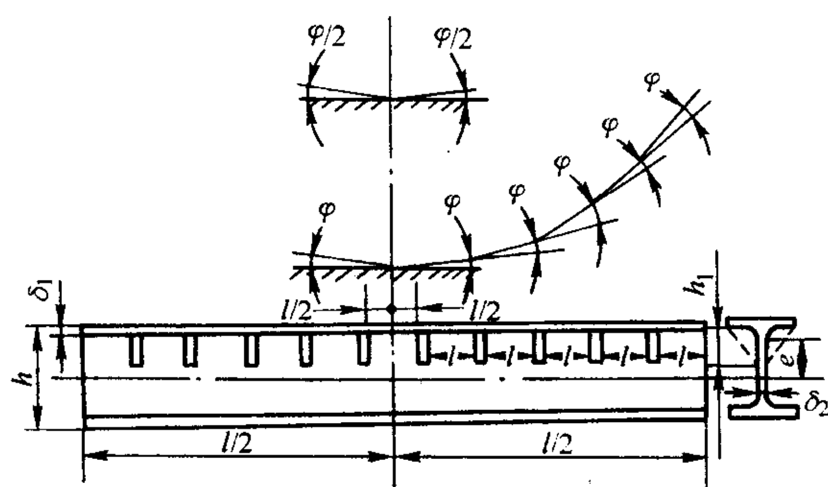


图 4.10-1　短肋板引起的弯曲变形



图 4.10-2　薄板受压失稳变形

4）角变形　对接接头的角变形 β，随坡口角度增大而增大。T形接头的角变形 β 取决于角焊缝的焊脚尺寸 K 和板厚 δ。

5）扭曲变形　与角焊缝所造成的角变形沿焊接方向逐渐增大的影响有关。

6）波浪变形　在焊接内应力的压应力的作用下，薄板可能失稳，产生波浪变形（图 4.10-2）。压应力越大，薄板的宽度与厚度之比 B/δ 越大，就越易失稳，产生波浪变形。压应力随焊缝尺寸和焊接线能量的增加而增加。角变形也会引起类似的波浪变形，但其原因与此不同。

2.1.2　焊接残余应力

1）纵向（沿焊缝方向）残余应力　由于焊缝冷却纵向收缩，引起焊缝纵向残余应力。

2）焊缝横向残余应力　焊缝横向残余应力产生的直接原因来自焊缝冷却的横向收缩，间接原因是焊缝的纵向收缩。另外表面和内部不同的冷却过程以及可能叠加的相变过程也是影响因素。

2.2　减小焊接变形与应力的方法

2.2.1　控制焊接残余变形的方法

（1）设计措施

1）选择合理的焊缝尺寸和形状　在保证结构有足够承载能力的前提下，应采用尽量小的焊缝尺寸。对仅起联系作用的和受力不大按强度计算尺寸甚小的角焊缝应按板厚选取工艺上可能的最小尺寸。

在薄板结构中采用电阻点焊来代替熔化焊缝可以减少变形，节省焊后校正工作量。

2）尽可能减少焊缝的数量　适当选择壁板的厚度，可减少肋板的数量，从而节省焊接和变形校正量。对自重要求不严格的结构，这样做即使质量稍大，也是比较经济的。采用压型结构代替肋板结构对防止薄板结构的变形十分有效。

3）合理安排焊缝位置　焊缝对称于构件截面的中性轴，或使焊缝接近中性轴，可减少弯曲变形。

（2）工艺措施

1）反变形　图 4.10-3～图 4.10-5 为反变形在各种情况下的应用。当构件刚度过大，采用上述强制反变形有困难时，可以先将梁的腹板在下料拼板时作成上挠的，然后再进行装配焊接。

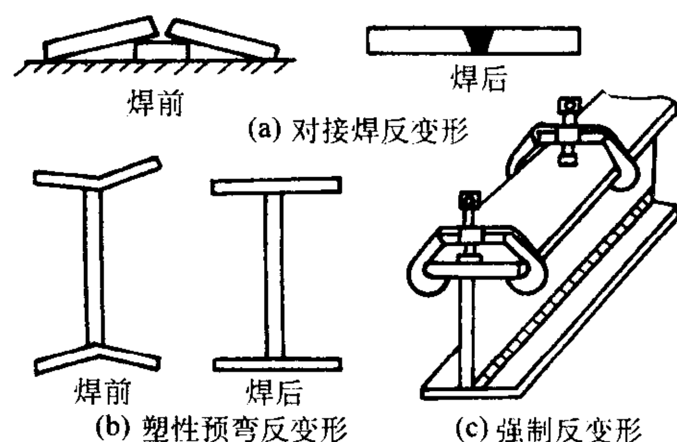


图 4.10-3　防止角变形的反变形措施

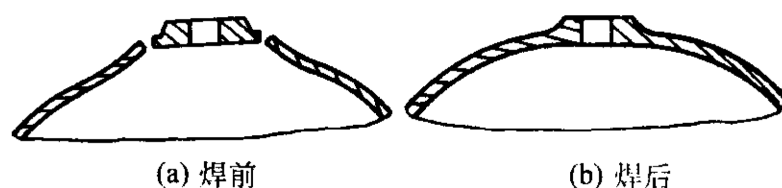


图 4.10-4　防止壳体局部塌陷的反变形

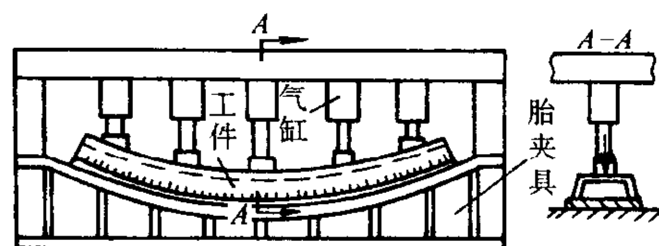


图 4.10-5　防止构件弯曲的强制反变形

在薄板上焊接骨架时，对薄板采用加热、机械拉伸或两者同时使用使其伸长，然后在薄板上装配焊接骨架，薄板预拉伸和加热后再冷却所产生的拉应力可有效地降低焊接应力，防止失稳波浪变形。

在薄板对接时可采用在焊缝两侧一定距离处适当宽度上加热，使焊缝得到拉伸，从而减少压缩塑性变形，降低残余应力，而消除波浪变形。

2）刚性固定法　对防止弯曲变形的效果不如反变形，但对角变形和波浪变形有较好的效果。焊接薄板时为防止波浪变形，在焊缝两侧压紧固定，加压位置应尽量接近焊缝并保持压力均匀。

3）选择合理的焊接方法和规范　选用能量较集中的焊接方法，如 CO_2 保护焊、等离子弧焊代替气焊和手工电弧焊进行薄板焊接，可减少变形量。

焊缝不对称的构件，可通过选用适当的焊接工艺参数，在没有反变形或夹具的条件下，控制弯曲变形。

在焊缝两侧采用直接水冷或水冷铜块散热，可限制和缩小焊接热场，减少变形，但对淬火倾向大的钢种应慎用。

4）选择合理的装配焊接次序　把结构适当地分成部件，分别装配焊接，然后再拼焊成整体。使不对称的焊缝或收缩量较大的焊缝能比较自由地收缩而不影响整体结构。

2.2.2　减小焊接应力的设计方法

（1）设计措施

1）尽量减少焊缝的数量和尺寸。

2）避免焊缝过分集中，焊缝间应保持足够的距离。

3）采用刚性较小的接头形式。

4）在残余应力为拉应力的区域内，应避免几何不连续性，以免内应力在该处集中。

（2）工艺措施

1）采用合理的焊接顺序和方向　先焊收缩量较大的焊缝，使焊缝能较自由地收缩。先焊错开的短焊缝，后焊直通长焊缝，使焊缝有较大的横向收缩余地。先焊在工作时受力

较大的焊缝，使内应力合理分布。

2）降低焊缝的拘束度　在焊接镶块的封闭焊缝或其他拘束度大的焊缝时，可采用反变形法降低焊件的局部刚度，以减小焊缝的拘束度。

3）锤击焊缝　用头部带小圆弧的工具锤击焊缝，使焊缝得到延展，降低内应力，锤击效果应保持均匀适度，避免锤击过分产生裂纹。一般不锤击第一层和表面层。

4）局部加热　在焊接结构的适当部位加热使之伸长，加热区的伸长带动焊接部位，使它产生一个与焊缝收缩方向相反的变形。在加热区冷却收缩时，焊缝就可能比较自由地收缩，从而减小内应力。

2.3 焊接结构设计

2.3.1 焊接结构设计的基本要求

1）实用性　结构必须达到产品所要求的使用功能和预期效果。

2）可靠性　结构在使用期内必须安全可靠。因此结构受力必须合理，能满足强度、刚度、稳定、抗振、耐蚀等方面要求。

3）工艺性　应该是能够焊接施工的结构，包括焊前预加工、焊后处理、所选用的金属材料具有良好的焊接性、结构的焊接与检验的可达性等。

4）经济性　制造该结构时，所消耗的原材料、能源和工时应最少，综合成本低。

2.3.2 焊接结构设计基本原则

1）合理选择和利用材料　所选用的材料必须同时能满足使用性能和加工性能的要求。有特殊性能要求的部位，可采用特种金属，其余采用满足一般要求的廉价金属。尽可能选用轧制的标准型材和异型材，可减少备料工作量和焊缝数量。在划分结构的零、部件时，要考虑到备料过程中合理排料的可能性，以减少余料。

2）合理设计结构形式　尽量采用简单、平直明快的构造形式，减少短而不规则焊缝。要有利于实现机械化和自动化焊接，避免采用难以弯制或冲压的具有复杂空间曲面的结构。

3）减少焊接量　利用冲压件代替一部分焊件；结构形状复杂、角焊缝多且密集的部位，可用铸钢件代替；必要时，可适当增加壁厚，以减少或取消加强肋板等。对于角焊缝，在保证强度要求的前提下，尽可能用最小的焊脚尺寸。对于对接焊缝，在保证焊透的前提下选用填充金属量最少的坡口形式。

4）合理布置焊缝　有对称轴的焊接结构，焊缝宜对称布置，或靠近对称轴处，这有利于控制焊接变形；要避免焊缝汇交和密集；在结构上使重要焊缝连续，让次要焊缝避开高工作应力处、应力集中部位、机械加工面和需变质处理的表面等。

2.4 焊缝金属与母材强度匹配的优化选择

焊缝金属强度与母材强度匹配有三种，焊缝金属强度比母材高的称为高匹配，比母材低的称为低匹配，等于母材时称为等强匹配。焊接接头的力学性能与母材和焊缝二者之间的强度如何匹配有关。一般来讲，结构抗塑性断裂的承载能力正比于材料的强度，结构抗脆性断裂的承载能力正比于材料的韧度。保证结构安全运行的条件是实际结构的载荷应同时小于抗塑性断裂和抗脆性断裂的承载能力。图4.10-6是材料强度与韧度匹配关系图，可见保证结构运行的最低安全限度在材料的强度和韧度匹配关系图中应为带有一定斜率的直线 S。

对于中低强度钢，由于母材和焊缝强度都有较高的韧度储备，可按等强匹配原则，选择高匹配的焊接接头显然是合理的，因为焊缝韧度比母材即使有所降低，但也不会影响整个结构的安全性。而且当横向载荷使结构发生弹塑性变形时，由于焊缝区比母材区发生较小的应变（图4.10-7a），焊缝区相当于受到保护，高匹配焊接接头的抗脆断安全性甚至更高。

对于超高强钢应根据等韧度原则，选择焊缝韧度不低于母材金属的低匹配焊接接头。若根据等强度原则，选用高匹配焊接接头，可能导致焊缝韧度不足而引起低应力脆性破坏。

对于高强度钢和大型厚板，焊接时易产生焊接裂纹和脆性断裂。为避免焊接裂纹，可采用低匹配接头。对于高韧度的低匹配焊接接头，在弹性应力区焊缝抗脆性断裂能力显然高于母材。但在弹塑性区，由于焊缝比母材承受更大的工作应变（图4.10-7b），将导致低匹配接头抗脆断能力下降。因此在结构的高应变区使用低匹配焊缝时，应使焊缝具有比母材足够高的韧度，才能保证结构整体的抗脆断能力。

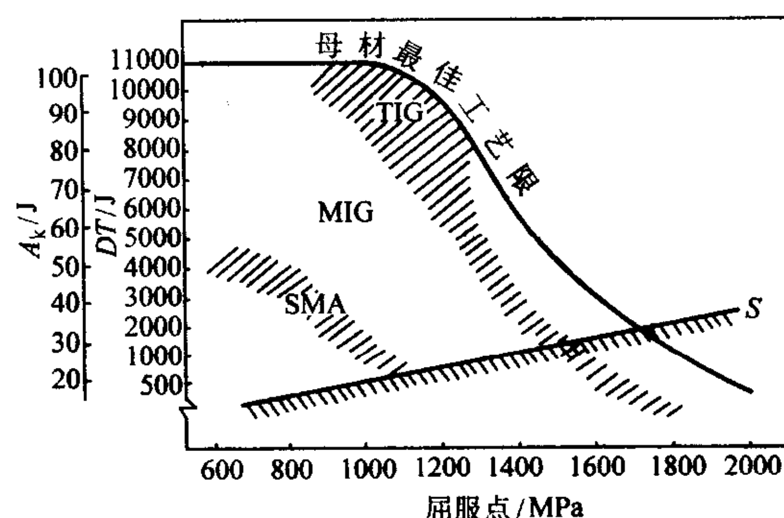


图4.10-6　材料的强度与韧性匹配关系图

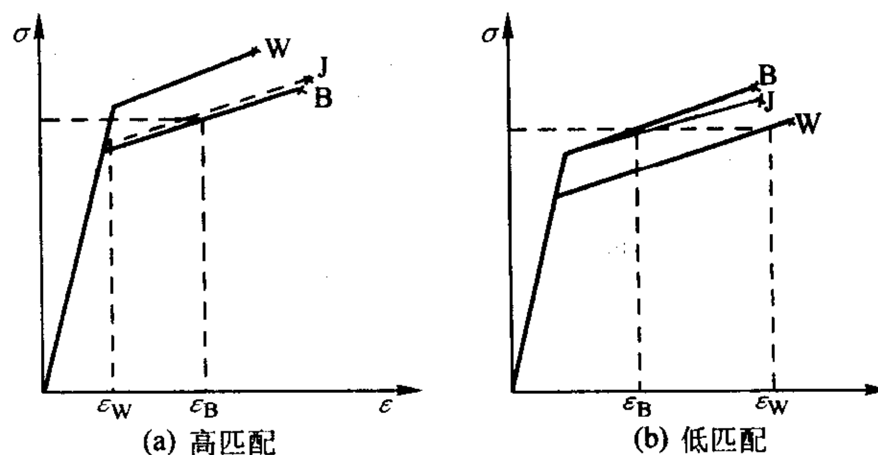


图4.10-7　对接接头不同匹配的应力与应变关系：
W—焊缝金属；B—母材；J—接头

2.5 焊接结构优化设计实例

1）问题的提出　大型薄壁箱型结构广泛应用在机械、建筑、船舶以及航空航天等领域，一般采用焊接结构。典型的如桥式起重机的主梁，具有承载能力大，水平刚度和抗扭刚度大，工艺性好，制造方便，封闭性好等特点，已被广泛应用。设计箱型梁，就是根据其实际工况，经过计算分析综合，得到一组梁的截面参数。在设计过程中，如何有意识地选择梁的截面参数，使梁具有较高的性能，同时最合理有效地利用材料，尽可能降低成本，是箱型梁设计的主要目标。

优化设计的特点是将实际问题数学模型化，然后利用数学规划的方法求解。

2）设计变量　图4.10-8所示为箱型梁截面。在实际设计中，一般根据情况给定梁的宽度 B，这样选择的参数为梁的腹板高度，上下盖板的厚度，左右腹板的厚度。对于纵向

加肋等，由于其尺寸相对很小，根据板高和板厚等因素，考虑局部稳定而配置的，所以在此不作为设计变量。因此，设计变量为：

$$x=\begin{Bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{Bmatrix} \tag{4.10-27}$$

式中，x_1 为腹板高度；x_2 为上盖板厚度；x_3 为左腹板厚度；x_4 为右腹板厚度；x_5 为下盖板厚度。

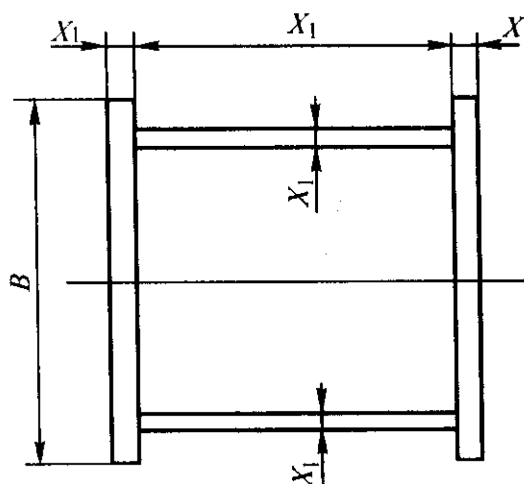


图 4.10-8　箱型梁截面

3）目标函数　箱型梁在工作时受到弯矩、剪力作用，起重机桥架的偏轨型梁又受到一定的扭矩作用，但主要是受弯矩作用。一般来说，梁的抗弯模量越大，其承载能力越高。因此在梁的的结构优化设计中，力求箱型梁的抗弯模量最大；另外在满足性能和要求的前提下，又要节省材料、降低成本。因此，目标函数选取梁的抗弯模量最大，梁的质量最小即梁的截面积最小为目标函数，即：

$$\begin{cases}\max f_1(x)=\dfrac{B(x_1+x_2+x_3)^3-(B-x_3-x_4)x_1^2}{6(x_1+x_2+x_5)}\\ \min f_2(x)=B(x_2+x_5)+x_1(x_3+x_4)\end{cases} \tag{4.10-28}$$

4）约束条件　对每个设计变量都给出尺寸界限的约束，即对梁的腹板高度和每块板的厚度给出上下界限值。此外还有上下盖板的厚度相等，盖板的厚度大于腹板的厚度等约束。对于起重机偏轨箱型梁还有主腹板厚大于副腹板厚的约束。可建立约束条件为：

$$\begin{cases}-x_2+x_3+a_1\leqslant 0\\ -x_3+x_4+a_2\leqslant 0\\ a_3\leqslant x_1\leqslant a_4\\ a_5\leqslant x_i\leqslant x_6\quad i=2,3,4,5\end{cases} \tag{4.10-29}$$

其中，a_i（$i=1,\cdots,6$）为非负数，是界限值，如 a_3 和 a_4 为梁腹板高的下限和上限。

5）编程计算（略）。

3　焊接电源的优化设计

3.1　弧焊电源优化设计过程

焊接电源的优化设计过程大致包括建立优化设计的数学模型、确定计算方法、编制计算程序、上机计算、计算结果分析和实用化处理。

（1）优化设计数学模型的建立

最优化设计，首先要明确问题的设计变量及约束条件，并确定合理的目标函数。

1）设计变量　在焊接电源的设计计算中，电源的技术性能和经济指标与很多参数有关，其中有些参数允许在一定的范围内进行调整以获得较好的设计效果，这些参数通常可取作优化设计变量。还有一些参数由于选择余地不大，一般根据经验取为常数。

2）目标函数　合理确定优化目标函数十分重要，不同的优化目标将得到不同的最优方案。从焊接电源的优化目标来看，希望设计出的产品在满足有关标准的前提下具有最高的效率、最低成本和最小质量。显然，这是个多目标优化设计问题。多目标优化设计目前处理起来比较困难，实际设计中一般仍按单目标问题求解。通常的做法是取上述三者之一作为优化目标，而其他两个作为约束条件。具体选择要视情况而定。

3）约束条件　焊接电源优化设计问题的约束条件比较复杂，确定起来比较困难。一般要根据设计变量的取值范围、电源的结构要求、性能要求以及经济指标等几个方面加以选择和确定。如果选择不当或确定错误，则不会达到期望的最优化目标。

（2）计算方法的选择

焊接电源的设计要满足规定的各项技术经济指标和制造工艺要求，所以其优化设计是一个有约束的优化问题。常用方法有罚函数法、网格法及随机试验法等。其中罚函数法适用于中型优化设计问题；网格法和随机试验法适用于小型优化设计问题。

（3）编程计算

可采用高级语言进行编制计算程序。编制的程序应具有一定的通用性。设计所用的数表和曲线应根据优化设计的要求，编制出相应的子程序，以便查找和检索。

（4）优化结果分析

对优化结果进行仔细的分析是非常重要的，一方面可以查明优化设计过程的正常发展情况及最终结果的合理性；另一方面可以获得许多有价值的规律性的东西。

（5）结果的实用化处理

计算机按优化准则计算出的优化设计方案，数据精度虽然很高，却不完全符合工程上的要求。此外，工程上的尺寸也不能都采用计算出的那样高的精度。因此，把计算数据运用于实际工程之前，必须根据材料规格和工艺上的可行性予以圆整和实用化处理。当然实用化处理结果可能偏离最优点，但只要比现有生产的焊机好，仍具有优化设计的意义。

3.2　晶闸管弧焊电源优化设计实例

（1）数学模型的建立

1）设计变量的确定　设计变量是在设计过程中进行选择并最终必须确定的各项独立参数。根据设计变量的选择原则，确定的弧焊电源主变压器 14 个优化设计变量见表 4.10-14。主变压器结构示意图见图 4.10-9。

表 4.10-14　弧焊电源主变压器的设计变量

序号	设计变量	代表符号	序号	设计变量	代表符号
1	一次导线宽/cm	a_{Cu_1}	8	二次绕组单侧盘数	m_2
2	一次导线厚/cm	b_{Cu_1}	9	铁心窗口宽度/cm	a
3	二次导线宽/cm	a_{Cu_2}	10	铁心柱宽度/cm	b
4	二次导线厚/cm	b_{Cu_2}	11	铁铜比	K
5	一次绕组每盘匝数	n_1	12	铁心窗口高/cm	h
6	一次绕组单侧盘数	m_1	13	铁心柱厚度/cm	l
7	二次绕组每盘匝数	n_2	14	磁通密度/T	B

设计变量为：

$$\begin{aligned}x&=[B_m,K,m_1,n_1,m_2,n_2,a_{Cu_1},b_{Cu_1},a_{Cu_2},b_{Cu_2}\ l,b,h,a]^T\\&=[x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10},x_{11},x_{12},x_{13},x_{14}]^T\end{aligned} \tag{4.10-30}$$

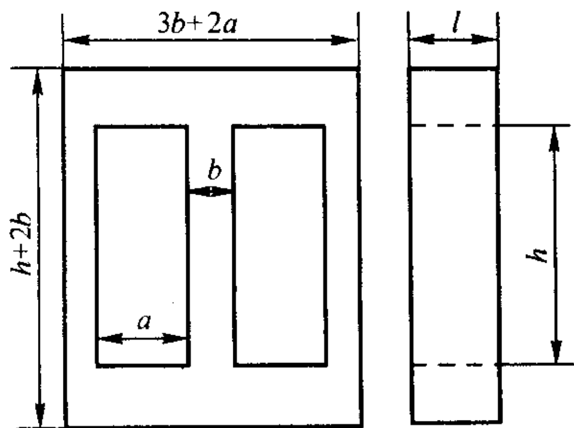


图4.10-9 主变压器结构示意图

2）目标函数的确定 弧焊电源主变压器的优化设计是个多目标函数的优化设计，采用加权组合法把各分目标函数组合为“统一目标函数”，从而使几项目标都达到最佳。

主变压器的优化目标函数模型为：

$$\min f(x)=c_1f_1(x)+c_2f_2(x)+c_3f_3(x)$$
$$=(c_1+c_2)(G_{Cu}+G_{Fe})+c_3(P_{Cu}+P_{Fe}) \quad (4.10\text{-}31)$$

式中 c_1、c_2、c_3 分别是 $f_1(x)$、$f_2(x)$、$f_3(x)$ 的加权值；

$$P_{Cu}=1.05\times\rho I_1^2\times\frac{m_1n_1l_{m_1}}{a_{Cu_1}b_{Cu_1}}+1.05\times\rho I_2^2\times\frac{m_2n_2l_{m_2}}{a_{Cu_2}b_{Cu_2}}$$

$$P_{Fe}=7.34\times10^{-3}bl\ [3h+6b+4a]\ f\ (B_m)$$

$$G_{Cu}=2.67\times10^{-2}\ [l_{m_1}n_1m_1a_{Cu_1}b_{Cu_1}+2l_{m_2}n_2m_2a_{Cu_2}b_{Cu_2}]$$

$$G_{Fe}=6.885\times10^{-3}bl\ [3h+6b+4a]$$

$$l_{m_1}=2\ (b+l+4\times0.5)\ +\pi m_1\ (b_{Cu_1}+0.16)$$

$$l_{m_2}=2\ (b+l+4\times0.5)\ +\pi\ \{2m_1\ [\ (b_{Cu_1}+0.16)\ +1]\ +m_2\ (b_{Cu_2}+0.16)\}$$

式中，I_1 为一次侧线圈额定电流，A；I_2 为二次侧线圈额定电流，A；l_{m_1} 为一次侧线圈平均匝长，cm；l_{m_2} 为二次侧线圈平均匝长，cm；ρ 为铜导线的电阻率，Ω·cm；$f\ (B_m)$ 为硅钢片损耗功率，$W\cdot kg^{-1}$。

3）约束条件 主变压器最优化设计的约束条件有空载电流、空载电压、铁心柱截面积、一次侧线圈导线截面积、二次侧线圈截面积、导线宽度比、铁心柱宽度比、窗口尺寸、窗口宽度比、一次侧线圈匝数和层数、二次侧线圈匝数和层数、电流密度范围、线圈温升及各变量取值范围等。

(2) 构造目标函数

采用混合罚函数法构造目标函数，其罚函数为：

$$P(x,r_k,M_k)=f(x)+M_k\sum_{i=1}^{p}[h_i(x)]^2+r_k\sum_{u=1}^{m}\frac{1}{g_u(x)} \quad (4.10\text{-}32)$$

式中，x 为变压器结构参数数组；P 为罚函数；M_k 为罚函数因子；$h_i\ (x)$ 为不等式约束；$g_u\ (x)$ 为等式约束；$f\ (x)$ 为目标函数。

当 $r_k\to\infty$ 时，$\min P\ (x)\ =\min f\ (x)$

此时，$\min f\ (x)$ 就是最优化的目标函数，所对应的 x 就是变压器最佳性能的结构参数。

(3) 计算程序和计算过程

将确定的初始点代入程序，利用优化程序进行寻优，使设计的电源在满足性能要求的基础上各项参数达到最优。通过优化设计计算出的结构参数还不能作为生产弧焊电源的实际参数，必须通过靠线规划程序对优化出来的导线进行靠线规划处理，使设计出来的结构参数能够进行实际生产。验证计算程序对靠线规划处理过的结构参数进行验证，以使电源结构参数满足性能和结构要求。

4 焊接方法及焊接规范优化设计

4.1 焊接方法的分类

焊接方法的种类繁多，而且新的方法仍在不断涌现。按焊接工艺特征焊接方法分为三大类，即熔化焊、固相焊和钎焊。基本焊接方法的分类见图4.10-10。

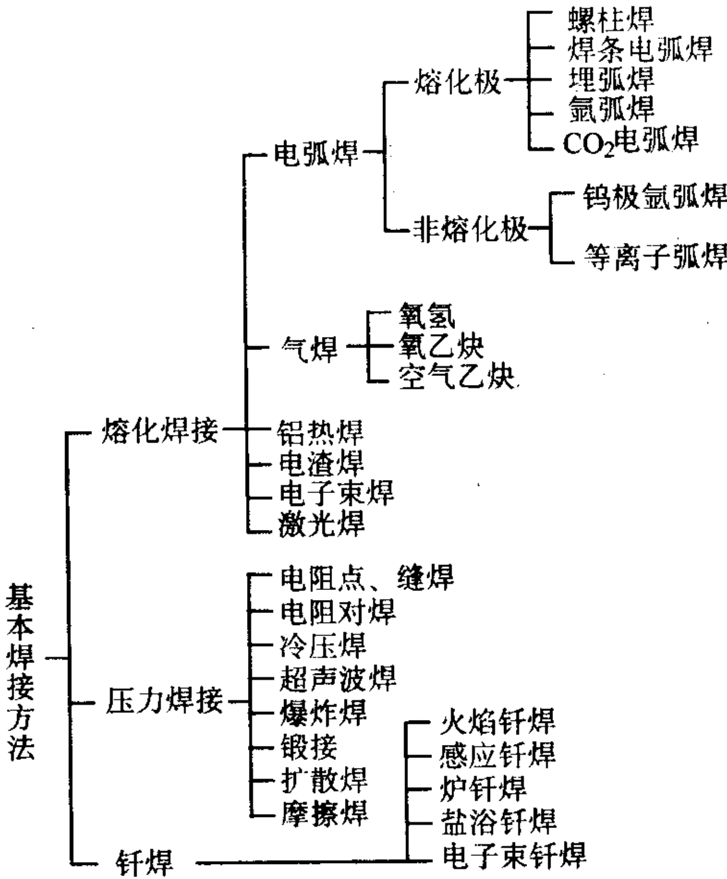


图4.10-10 基本焊接方法的分类

4.2 焊接方法的优化选择

(1) 选择的基本要求

选择焊接方法时必须符合以下要求：能保证焊接产品的质量优良可靠，生产率高，生产费用低，能获得较好的经济效益。

影响焊接方法选择的因素如下。

1）产品结构类型 不同结构的产品由于焊缝的长短、形状、焊接位置等各不相同，因而适用的焊接方法也会不同。结构类产品中规则的长焊缝和环焊缝宜用埋弧焊。手弧焊用于打底和短焊缝焊接。机械类产品可根据其结构特点，选用气体保护焊（一般厚度）、电渣焊、气电焊（重型构件宜于立焊）、电阻焊（薄板件）、摩擦焊（圆形断面）或电子束焊（高精度要求）。半成品类的产品的焊接接头往往较规则，宜采用机械化的焊接方法。微型电子器件的接头主要要求密封、导电性、受热程度小等，宜用电子束焊、激光焊、超声波焊、扩散焊、钎焊和储能焊。

2）工件厚度 工件厚度可在一定程度上决定所适用的焊接方法。图4.10-11为各种焊接方法适用的厚度范围。

3）接头型式和焊接位置 根据产品的使用要求和所用母材的厚度和形状，设计的产品可采用对接、搭接、角接等几种类型的接头型式。其中对接适用于大多数焊接方法。

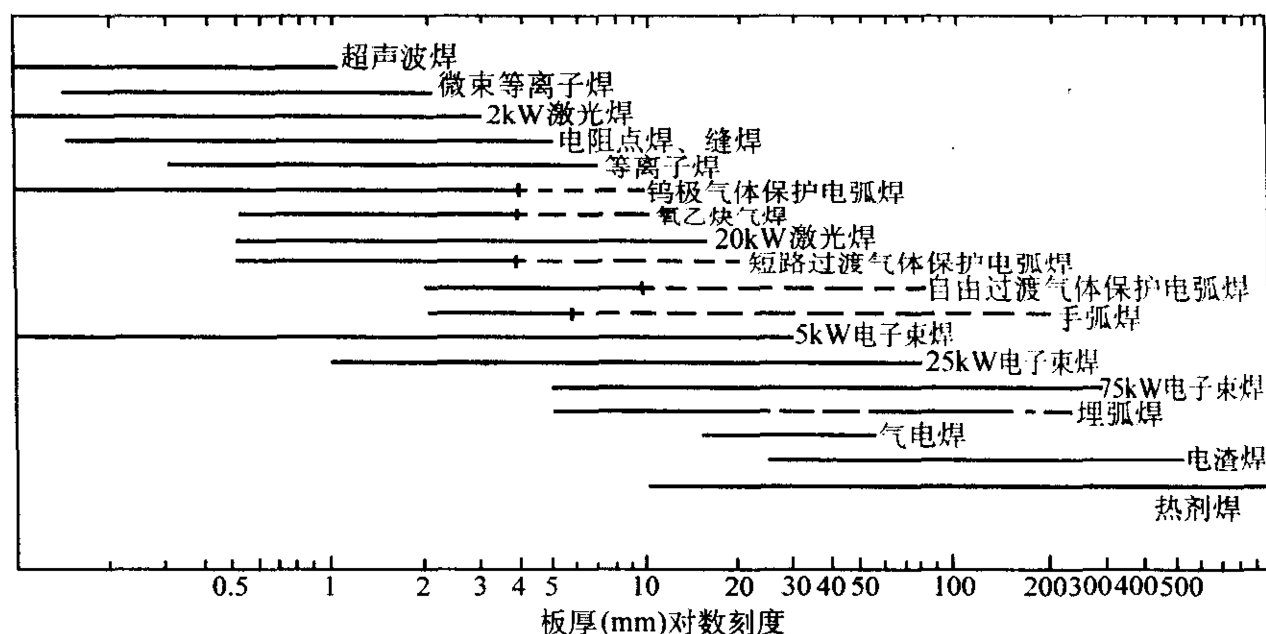


图 4.10-11　各种焊接方法适用的厚度范围

产品中各个接头位置往往根据产品的结构要求和受力情况决定。这些接头可能需要在不同的焊接位置焊接，包括平焊、立焊、横焊、仰焊及全位置焊接等。平焊是最容易、最普遍的焊接位置，因此焊接时应尽可能使产品接头处于平焊位置，这样就可以既保证良好的焊接质量，又获得较高的生产率。对于立焊接头宜采用熔化极气体保护焊（薄板）、气电焊（中厚板），当板厚超过 30 mm 时可采用电渣焊。

4）母材性能

① 母材的物理性能　母材的导热性能、导电性能、熔点等物理性能会直接影响其焊接性及焊接质量。

对于焊接导热系数较高的金属，如铜、铝及其合金，应选择热输入大，具有较高焊透能力的焊接方法，以使被焊金属在最短的时间内达到熔化状态，并使工件变形最小。相反，对于热敏感材料，应选择热输入较小的焊接方法，如激光焊、超声波焊等。

对于电阻率较高的金属宜采用电阻焊。而对于钼、钽等高熔点的难熔金属，宜采用电子束焊接方法。对于物理性能相差较大的异种金属，宜采用不易形成脆性中间相的焊接方法，如各种固相焊、激光焊等。

② 母材的力学性能　各种焊接方法对焊缝金属及热影响区的组织和性能的影响不同，因此也会不同程度地影响产品的使用性能。如电渣焊、埋弧焊时由于热输入较大，从而使焊接接头的冲击韧度降低。电子束焊的焊接接头热影响区较窄，与一般电弧焊相比，其接头具有较好的力学性能。因此电束焊对某些金属，如不锈钢或经热处理的零件是很好的焊接方法。有的焊接方法可以通过控制热输入从而控制熔深、熔合比和热影响区来获得力学性能与母材接近的接头。

③ 母材冶金性能　对于普通碳钢和低合金钢采用一般的电弧焊方法都可进行焊接。钢材的合金含量，特别是碳含量越高，焊接性越差，可选用的焊接方法种类就越有限。

对于较活泼的有色金属材料，如铝、镁及其合金，不宜选择 CO_2 气体保护焊、埋弧焊，而应选择惰性气体保护焊，如 TIG、MIG 等。对于不锈钢，通常可采用焊条电弧焊、TIG 或 MIG 等。特别是 TIG，其保护效果好，焊缝成分易控制，可以满足焊缝的耐蚀性能的要求。对于钛、锆这类金属，由于其易吸收较多的有害气体，焊后容易变脆，应采用高真空电子束焊接。

对于熔化焊不容易焊接的冶金相容性较差的异种金属，可考虑采用钎焊、扩散焊或爆炸焊等焊接方法。

（2）优化指标选择

在焊接生产过程中，当母材类型、板厚、焊缝形状、焊缝尺寸、焊缝位置、接头形式等一定时，采用不同的焊接方法，接头的质量、焊接效率和生产成本是不同的，可采用优化设计方法，根据具体情况确定出一最优焊接方法。

焊接方法的优化选择，可采用评价函数法进行优化设计，评价函数的基本原理是先将各优化指标依一定关系进行组合，从中选择一个单目标问题的目标函数，被称为评价函数。这个评价函数对各种方案进行评价，最终确定最优方案。

评价函数法的关键是合理地构造一个能恰当地反映所研究问题的评价函数。工程上常用的评价函数构造方法有平均加权法、线性加权法、虚拟目标法等。在构造评价函数之前，需要先统一各指标的量纲，量纲的统一可采用量纲为 1 相对化方法，也可采用专家评分的方法。

（3）评价函数的建立

采用线性加权法构造评价函数，其数学表达式为：

$$\phi = c_1 k_1 + c_2 k_2 + c_3 k_3 \tag{4.10-33}$$

式中，c_1、c_2、c_3 分别为质量、生产效率、成本加权值；k_1、k_2、k_3 分别是质量因子、生产效率因子、成本因子。

评价函数 ϕ 值越大，则焊接方法越优。

（4）计算程序

确定最佳焊接方法的程序流程图见图 4.10-12。

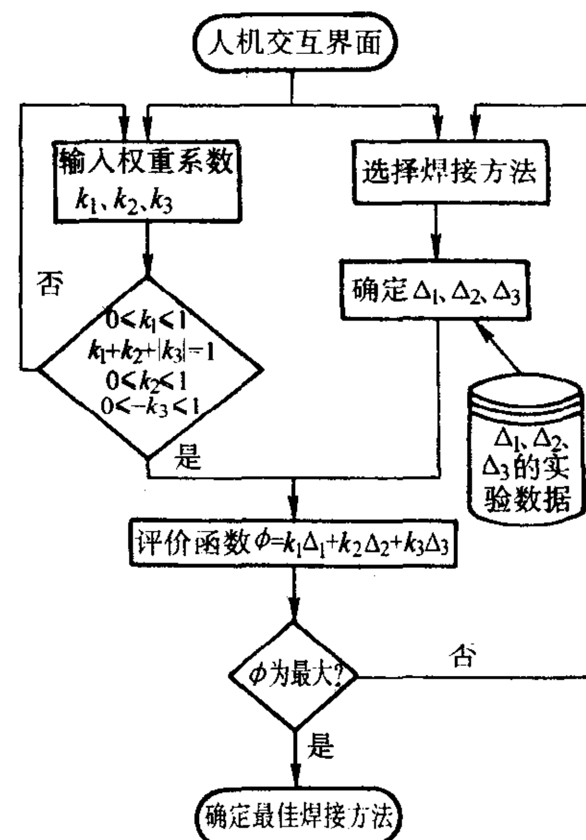


图 4.10-12　确定最佳焊接方法的程序流程图

（5）实例分析

焊接母材为 16MnR，焊缝质量权重 $k_1 = 0.6$，生产效率权重 $k_2 = 0.2$，成本权重 $k_3 = -0.2$，可选焊接方法有焊条电弧焊、埋弧自动焊、CO_2 气体保护焊，板厚 12 mm，接头形式为对接、平焊，焊缝长度 2 m 以上。

优化分析结果为：焊条电弧焊评价函数 $\phi=0.38$；埋弧自动焊评价函数 $\phi=0.58$；CO_2 气体保护焊评价函数 $\phi=0.4$。

因为 ϕ 最大值对应的焊接方法为最优，所以应采用埋弧自动焊，计算结果与实际应用一致。

4.3 CO_2 气体保护焊焊接规范的优化设计

（1）变量设计

CO_2 气体保护焊焊接低碳钢时，在焊丝成分一定的条件下，焊接接头质量主要取决于焊缝成形。而焊缝成形和焊接生产率与焊接规范存在密切关系。国家标准对焊缝的横截面几何尺寸，如焊缝宽度 e、焊缝的余高 g、构件的装配间隙 b 等（见图 4.10-13），都规定了允许的变化范围。在角焊缝中，为防止咬边和溢流，要求焊缝宽度 $e=\sqrt{2}K$，K 为焊脚尺寸。S_e 值按照在间隙最小时能保证得到要求的有效熔深，在间隙最大时不烧穿垂直板的原则选取。在保证焊缝几何尺寸和损失系数较小的条件下，寻求焊接参数电流 I、电压 U、送丝速度 v_f 使生产率（焊接速度 v）的最大值。

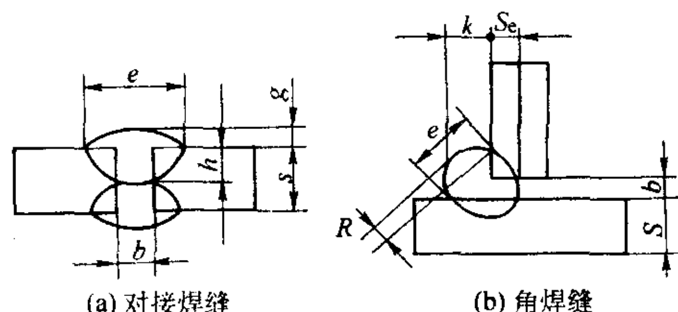


图 4.10-13 焊缝的横截面尺寸

（2）数学模型的建立

采用试验规划的方法建立数学模型。

试验间隙 b 对熔深 h 的影响数学表达式为：

$$h=f\ (U,\ v_f,\ v,\ d)\ +ab \tag{4.10-34}$$

式中，d 为焊丝直径。对接时，取 $a=1.8$，角焊缝 $a=1.2$。

将试验得到的数据进行统计分析，建立回归数学模型：

$$Y=\theta_0+\theta_1U+\theta_2v_f+\theta_3v+\theta_4d+\theta_5Uv_f+\theta_6Uv+\theta_7Ud+\theta_8v_fv+\theta_9v_fd+\theta_{10}vd+\theta_{11}U^2+\theta_{12}v_f^2+\theta_{13}v^2+\theta_{14}d^2 \tag{4.10-35}$$

式中，Y 为焊缝几何尺寸、损失系数 ψ 和焊接电流 I 等指标；θ_i（$i=1, 2, \cdots, 14$）为相应各项的回归系数。

（3）优化过程

由回归方程 $I=f\ (U,\ v_f,\ v,\ d)$ 求出 $v=\varphi\ (I,\ v_f,\ U,\ d)$，并代入其他回归方程，消去变量 v。

$$\text{设}\quad X=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}I\\v_f\\U\\d\end{pmatrix} \tag{4.10-36}$$

则该最优化问题的数学模型为：

$$\text{对接}\quad\left.\begin{array}{l}\max v=\varphi\ (x_1,\ x_2,\ x_3,\ x_4)\\h_1\leqslant h\ (x_1,\ x_2,\ x_3,\ x_4)\ \leqslant h_2\\e_1\leqslant e\ (x_1,\ x_2,\ x_3,\ x_4)\ \leqslant e_2\\g_1\leqslant g\ (x_1,\ x_2,\ x_3,\ x_4)\ \leqslant g_2\\g/e\leqslant 0.35\\\psi\ (x_1,\ x_2,\ x_3,\ x_4)\ \leqslant 10\%\\a_{1i}\leqslant x_i\leqslant a_{2i}\quad (i=1,\ 2\cdots,\ 14)\end{array}\right\} \tag{4.10-37}$$

$$\text{角接}\quad\left.\begin{array}{l}\max v=\varphi'\ (x_1,\ x_2,\ x_3,\ x_4)\\S_{e_1}\leqslant S_e\ (x_1,\ x_2,\ x_3,\ x_4)\ \leqslant S_{e_2}\\\sqrt{2}K_{\min}\leqslant e\ (x_1,\ x_2,\ x_3,\ x_4)\ \leqslant\sqrt{2}K_{\max}\\g_1\leqslant g\ (x_1,\ x_2,\ x_3,\ x_4)\ \leqslant g_2\\g/e\leqslant 0.5\\\psi\ (x_1,\ x_2,\ x_3,\ x_4)\ \leqslant 12\%\\a_{1i}\leqslant x_i\leqslant a_{2i}\quad (i=1,\ 2\cdots,\ 14)\end{array}\right\} \tag{4.10-38}$$

式中，h_1、S_{e_1} 为间隙 b 最小时的熔深；h_2、S_{e_2} 为间隙 b 最大时的熔深；e_1、e_2 为焊缝的最小和最大宽度；g_1、g_2 为对应最大间隙和最小间隙的焊缝余高；a_{1i}、a_{2i} 为参数 x_i 的允许变化区间。

在对接焊缝时，取 $b_{min}=b_{max}-\frac{1}{6}S$；在角焊缝时，取 $b_{min}=0$。

给定不同的原始数据和约束条件，进行优化计算。

5 焊接工程成本分析优化设计

5.1 焊接工程多阶段决策问题优化设计

焊接工程中常遇到多阶段决策问题，如焊接生产计划的安排调度等。这类问题的特点是，可将整个问题按时间或空间的顺序分成若干阶段，且只有对其中任意一个阶段作出决策后，才能确定下一阶段的状态，依次对每个阶段作出决策，整个问题的状态就完全确定了。对于这种多阶段决策过程的最优化问题，动态规划是一种有效的求解方法。

（1）数学模型

动态规划求解的基本思路是，把整个问题分成若干个子问题，然后依次求出每个子问题的最优决策，从而使整个问题达到最优。其最优化原理为：一个过程的最优决策具有这样的性质，即无论其初始状态和初始决策如何，从这一决策所导致的新状态开始，以后的一系列决策也是最优的。

最优化原理的迭代公式为：

$$f_K\ (X_K)\ =opti\ \{g_K\ (X_K,\ U_K)\ +f_{K+1}\ (X_{K+1})\} \tag{4.10-39}$$

式中，X_K 为第 K 段的状态变量；U_K 为第 K 段的决策变量；$f_K\ (X_K)$ 为第 K 段状态为 X_K 时的最优值；$f_{K+1}\ (X_{K+1})$ 为第 $K+1$ 段状态为 X_{K+1} 时的最优值；$g_K\ (X_K,\ U_K)$ 为第 K 段状态为 X_K 时，决策变量为 U_K 时的函数值。

Opti 根据实际取 min 或 max。

（2）实例分析

图 4.10-14 所示为一跨度 $l=24$ m，承受均布载荷的简支梁。载荷密度 $q=9\,800$ N/m，超载系数 $n=1.2$，允许的挠度与跨度比 $f/l=1/200$，梁的材料为 3 号钢，梁的截面积为对称的工字钢，用焊缝连接。为了满足强度和刚度条件，经计算跨度截面的尺寸见图 4.10-14，其中腹板的尺寸是按支座处的抗剪强度确定的。此外，规定腹板的截面形状、尺寸和翼板的宽度沿梁的长度是固定不变的，而翼板的厚度是可以变化的。现将半跨梁长分为 4 段，每段 3 m。根据抗弯强度的要求，计算得到每段翼板的厚度 δ_0。如表 4.10-15 所示。

如果各段翼板都取最小厚度，虽然可以节省金属材料，但各段连接焊缝金属的费用却增加了。这就有可能使增加的焊缝费用超过节省的金属材料费用。相反，如果各段翼板都采用 30 mm 厚的钢板，虽然可省去焊缝费用，但金属材料的费用却增加了。因此就构成了一个最优化问题，即如何选配梁的翼板厚度才能使梁的价格最低。

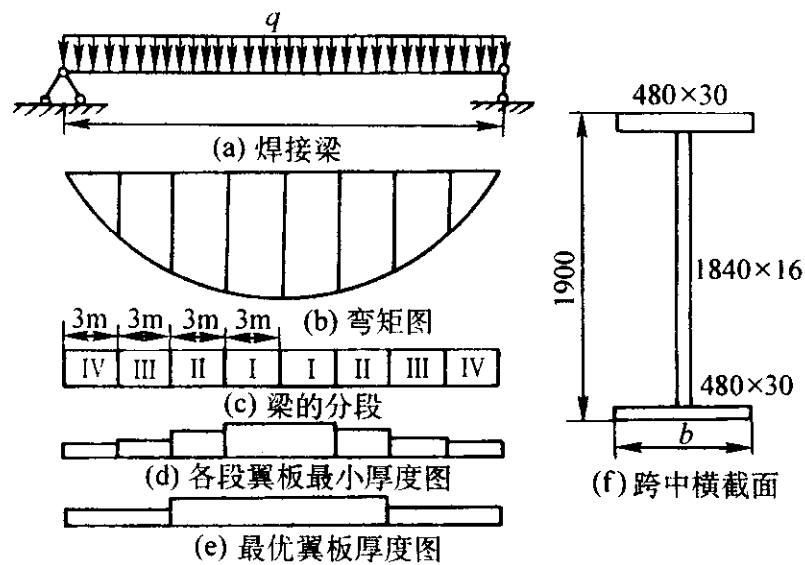


图 4.10-14 焊接梁

表 4.10-15 各段翼板的最小厚度

分段号	I	II	III	IV
δ_0/mm	30	28	20	16

设已知钢材的单价为 $C_1=0.625$ 元/kg，焊缝的单价为 $C_2=0.25$ 元/m³，一条焊缝的准备与施焊费用为 $C_3=25$ 元。假设任意一段翼板的厚度选为 δ（$\delta\geqslant\delta_0$），则该段增加的钢材费用为：

$$\Delta C=C_1\gamma\times300\times48\ (\delta-\delta_0)\ =70\ (\delta-\delta_0)\qquad(4.10\text{-}40)$$

式中，γ——钢的密度，取 7.8 g/cm³。

根据上式可以算出各段翼板取不同厚度时，增加的钢材费用，见表 4.10-16。

表 4.10-16 各段翼板加厚增加的钢材费用

段号	II	III		IV		
δ_0/mm	28	20		16		
δ/mm	30	30	28	30	28	20
ΔC/元	14	70	56	98	84	28

连接两块翼板的焊缝横截面如图 4.10-15 所示。计算熔敷金属的横截面时，可以近似地取它等于影线部分的面积。在翼板宽度 $B=480$ mm 的情况下，一条焊缝的总制造费用为：

$$C_H=C_2F_H\times48+C_3=12F_H+25\qquad(4.10\text{-}41)$$

其中 F_H 可由接头的几何尺寸计算出来。用上式可计算出不同厚度翼板接头的制造价格，见表 4.10-17。

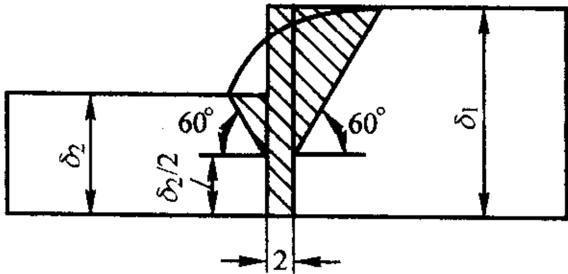


图 4.10-15 不同厚度翼板的焊接接头

表 4.10-17 不同厚度翼板接头的制造价格

δ_1/mm	30				28		20
δ_2/mm	30	28	20	16	20	16	16
F_H/mm²	190	190	205	219	179	191	101
C_H/元	47.8	47.8	49.63	51.30	46.50	47.88	37.13

这是一个三阶段决策过程的寻优问题。图 4.10-16 为焊接梁多阶段决策过程图。其中圆圈内的数字为各状态变量 δ_i（翼板厚度）的数值。根据设计要求和钢板厚度的规格，第IV段梁翼板厚度可有 4 种选择，即 16 mm、20 mm、28 mm、30 mm；第III段有 3 种选择；第II段有 2 种选择；而第I段只有一种选择，即 $\delta_1=30$ mm。圆圈之间连线上的数字为由于翼板厚度增加和施焊增加的费用额。

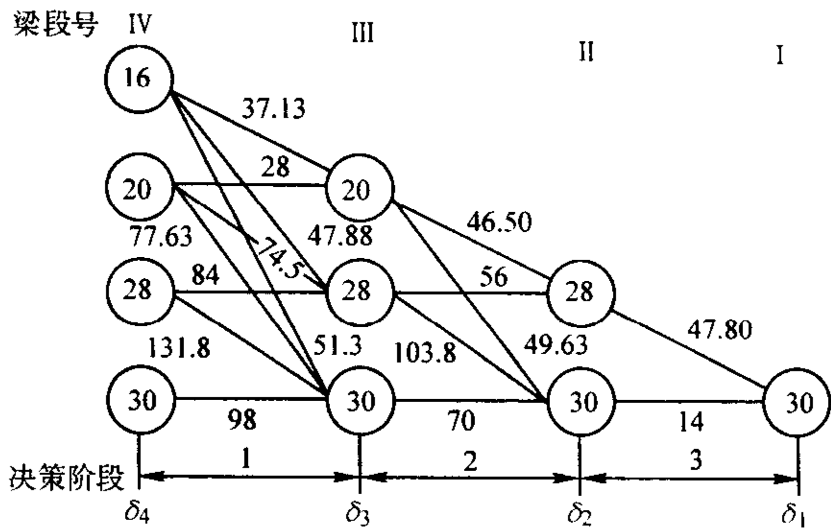


图 4.10-16 焊接梁多阶段决策过程图

由决策过程的最后一个阶段开始计算。由 $\delta_1=30$ mm 到 $\delta_2=28$ mm，只有一条路线，其最小费用为 f_3（28）$=47.8$ 元；由 $\delta_1=30$ mm 到 $\delta_2=30$ mm 也只有一条路线，其最小费用为 f_3（30）$=14$ 元。

再计算第 2 阶段，这时有 3 个状态变量，即 $\delta_3=20$ mm、28 mm、30 mm。若选取 $\delta_2=28$ mm、30 mm 作为决策变量值，则由式 $f_K(X_K)=opti\{g_K(X_K,U_K)+f_{K+1}(X_{K+1})\}$ 得：

$$f_2(20)=\min\begin{Bmatrix}g_2(20,30)+f_3(30)\\g_2(20,28)+f_3(28)\end{Bmatrix}=\min\begin{Bmatrix}49.63+14\\46.5+47.8\end{Bmatrix}=63.63$$

这说明若 $\delta_3=20$ mm，选取 $\delta_2=30$ mm 费用最小，为 63.63 元，其最优路线用箭头表示。

同理有

$$f_2(28)=\min\begin{Bmatrix}g_2(28,28)+f_3(28)\\g_2(28,30)+f_3(30)\end{Bmatrix}=\min\begin{Bmatrix}56+47.8\\103.8+14\end{Bmatrix}=103.8$$

$$f_2(30)=g_2(30,30)+f_3(30)=70+14=84$$

但是，从实际出发，当 $\delta_3=\delta_2=\delta_1=30$ mm 时，在全跨范围内厚度为 30 mm 的翼板总长度达到 18 m，超过了钢板规格的最大长度 12 m。因此需要增加一条焊缝，由表 4.10-17 可知增加费用 47.8 元，而对于半跨梁需要增加一半价格。因此这个方案的实际费用为：f_2（30）$=84+47.8/2=107.9$ 元。

最后计算第 1 阶段，得

$$f_1(16)=\min\begin{Bmatrix}g_1(16,20)+f_2(20)\\g_1(16,28)+f_2(28)\\g_1(16,30)+f_2(30)\end{Bmatrix}=100.76$$

$$f_1(20)=\min\begin{Bmatrix}g_1(20,20)+f_2(20)\\g_1(20,28)+f_2(28)\\g_1(20,30)+f_2(30)\end{Bmatrix}=91.63$$

$$f_1(28)=\min\begin{Bmatrix}g_1(28,28)+f_2(28)\\g_1(28,30)+f_2(30)\end{Bmatrix}=187.80$$

$$f_1(30)=\min\{g_1(30,30)+f_2(30)\}=205.50$$

比较 f_1（16）~ f_1（30）可知，最小费用为 91.63 元，其相应的最优化设计为 $\delta_3=\delta_4=20$ mm，$\delta_1=\delta_2=30$ mm。

因此说明了在一般情况下，结构材料的质量是反映经济效益的重要指标，但是在某些情况下，如果只考虑材料的消

耗，而忽视焊接费用，就有可能得到材料消耗虽然最少，而整个造价却不是最低的结果。

5.2 焊接成本优化分析

（1）焊接成本的影响因素

影响整个焊接过程最终成本的因素有：

1）钢板准备（切割、开坡口等）、定位和矫正等辅助工时费用；

2）焊接燃烧电弧工时费用；

3）中心起弧、清理焊缝和消除应力等辅助工时费用和管理费用；

4）焊条、焊剂及保护气体等的焊接材料费用及电费。

其中，焊缝所需的填充金属量直接影响焊接燃弧工时与焊接材料的消耗，对焊接成本影响较大。

（2）焊接成本的估算

焊接加工成本 C_w 由焊接燃烧工时 t_r 费用、辅助工时（包括被焊工件集中、矫正和定位的时间 t_1 和换焊条、重新起弧、清渣以及清理焊缝的时间 t_2）费用以及焊条、保护气体等的焊接材料费用和电费 C_m 组成。其中：

$$t_r = \sum_{i=1}^{N} M_i/E_i = \sum_{i=1}^{N} L_iA_i\rho/E_i \qquad (4.10\text{-}42)$$

$$t_1 = \sum_{i=1}^{n} C_{1i}\alpha_i G_i^{0.5} X_i^{0.5} \qquad (4.10\text{-}43)$$

$$t_2 = \sum_{i=1}^{N} C_{2i}\alpha_i^{0.5} S_i^{1.5} L_i \qquad (4.10\text{-}44)$$

$$C_m = \sum_{i=1}^{N} M_iK_m/\eta_m = \sum_{i=1}^{N} L_iA_i\rho K_m/\eta_m \qquad (4.10\text{-}45)$$

因此，焊接加工成本 C_w 为：

$$\begin{aligned} C_w &= \frac{K_w(t_1+t_2+t_r)}{\eta_w} + C_m \\ &= K_w\left[\sum_{i=1}^{n} C_{1i}\alpha_i\sqrt{G_iX_i} + \sum_{i=1}^{N} C_{2i}S_i^{1.5}L_i + \sum_{i=1}^{N}\frac{L_iA_i\rho}{E_i}\right]/\eta_w + \sum_{i=1}^{N} L_iA_iK_m/\eta_m \end{aligned} \qquad (4.10\text{-}46)$$

式中，K_w 为企业单位工时成本系数；η_w 为工时利用率；N 为焊件焊缝总数；C_{1i} 为与辅助工时 t_1 有关的企业特定的工时系数；α_i 为难度系数，根据焊接结构的难易程度而定；G_i 为焊接第 i 条焊缝时被焊接零部件的总质量，当一次定位有多道焊缝焊接时，仅计算焊接首道焊缝时的焊接件总质量，其余赋零值；S_i 为板厚；n 为需集中、定位后焊接的焊缝数；X_i 为焊接第 i 条焊缝时需集中、定位后焊接的零件数；C_{2i} 为辅助工时 t_2 中与焊缝金属体积有关的工时系数；L_i 为第 i 道焊缝长度；A_i 为第 i 道焊缝截面积；ρ 为焊缝金属密度；E_i 为单位时间内焊缝金属填充量；K_m 为单位体积焊缝金属材料的费用系数；η_m 为焊缝金属材料利用率。

（3）焊接的成本特征

由式（4.10-46）可知，焊接加工成本及其主要相关参数可表示为如下形式：

$$C_w = f\,(L,\ A,\ G,\ \rho/E) \qquad (4.10\text{-}47)$$

焊缝截面积 A 可表示为焊缝特征参数的函数即：

$$A = f\,(S,\ b,\ \alpha\ (\beta),\ P,\ H,\ R)$$

式中，S 为焊接件板厚；b 为根部间隙；α（β）为坡口夹角；P 为钝边尺寸；H 为坡口深度；R 为根部半径；参数 S、b、α（β）、P、H、R 等组成了焊缝形式 w_t，可由 S 值从国家标准及企业内部制订的焊接规程中查到，即 $w_t = f(s)$；因而 A 可表示为 S 的函数，即 $A = A\,(S)$。

焊接效率 E 可表示为 $E = \pi d^2 \rho v/4$

式中，d 为焊丝直径；v 为焊条速度。

E 进一步可由焊接电流确定，而焊条直径 d 与焊接速度 v、焊接电流 I 各有特定的相关关系，且 d 是焊接形式和壁厚的函数，$d = d\,(w_t,\ S)$。因此 ρ/E 可表示为 S 的函数。

于是，焊接加工成本的全部特征参数都可由参数 S、L 及 G 表达。通常成本特征由结构特征、材料特征和精度特征三部分组成。对于焊接结构，由于有一定精度要求的特征一般均在焊接工序之后完成，因此在考虑焊接成本时一般可以不考虑精度特征；于是可把 S、L 和 G 视为焊接的成本特征。

根据成本计算结果分析成本结构，进行经济性评价。

编写：邹增大（山东大学）
王新洪（山东大学）
曲仕尧（山东大学）

第11章　产品数字化设计与成形工艺仿真优化技术

随着信息技术的发展，数字化概念已经应用到各个领域，出现了数字化设计、数字化制造、数字化装备、数字化企业、数字化军队、数字化城市等概念。

在材料成形领域，有限元数值模拟技术、成形过程优化设计理论与方法、模具计算机辅助设计与制造、快速成形技术、反求技术、专家系统的发展和应用，使传统的成形工艺与模具设计正经历着数字化的变革。

产品设计数字化并且与成形工艺仿真及优化技术集成，可以快速完成产品设计，并有效地确保所设计的产品具有最优的结构和可行的工艺性，显著降低产品开发周期、制造周期和相关成本。

对于某种新产品，基于数字化的产品设计与成形工艺仿真集成系统如图4.11-1所示，大致分为以下过程。

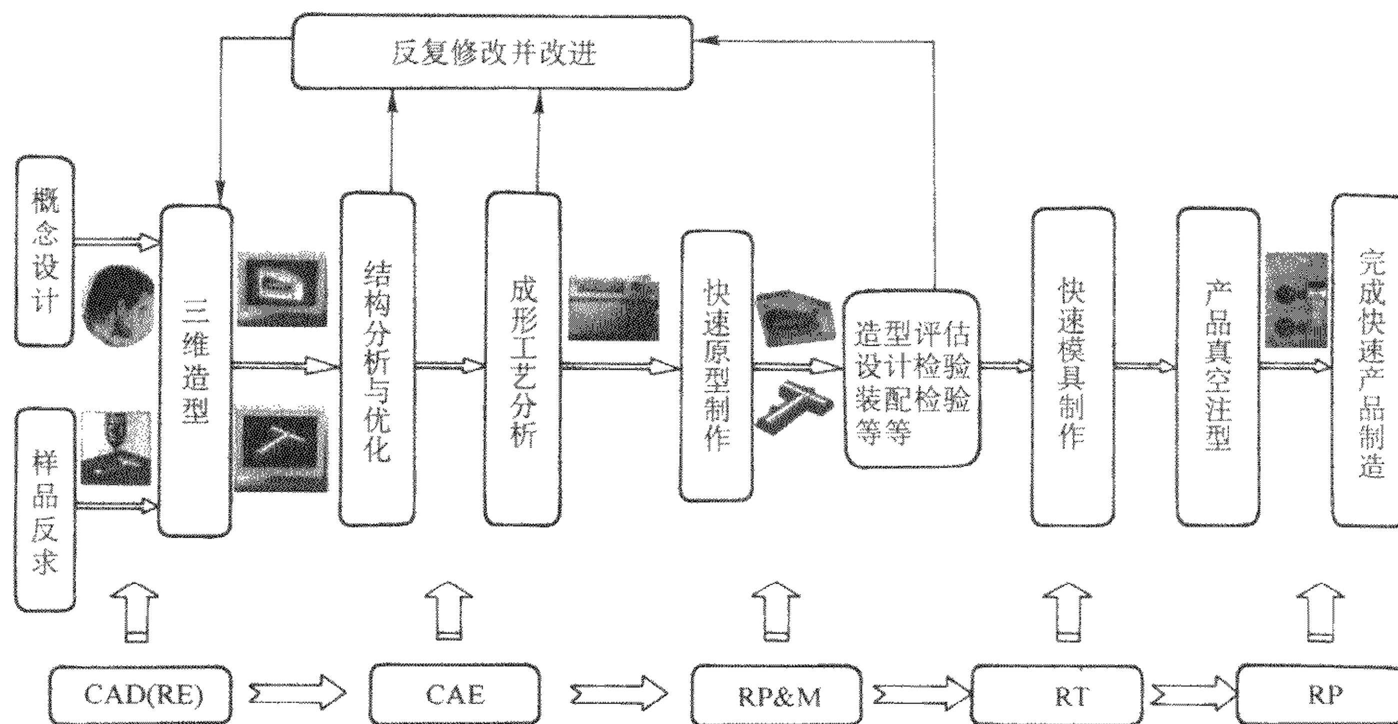


图4.11-1　产品数字化结构设计与材料成形工艺仿真优化技术系统

1）三维造型　从概念设计开始，或根据某个类似样品经反求开始，通过CAD软件或数据采集与拟合软件进行产品外观、结构设计及其零部件的三维造型设计。

2）结构分析与优化　如果构件为运动部件或受力部件，则需根据运动或受力状态，建立数学模型，分析构件的运动状态或受力状态，对结构进行分析和优化，将经优化后构件的三维数据传输到快速成形设备，制作原型，再次进行外观评估、设计检验、装配检验等，直至产品达到结构形状和运动或受力等的要求。在产品结构形状设计过程中，还应当考虑实际生产时材料成形的工艺性，从成形工艺的难度、模具形状的复杂程度及其工序数目的多少等方面综合考虑，从而在外观、结构、运动或受力状态、成形工艺、模具开发成本等多方面达到综合平衡。

3）快速成形　将三维造型设计数据以STL文件的格式传输至快速成形与制造设备，迅速制造出任意复杂形状的三维实体模型或零件，用于产品或零件的可视化检验、装配检验、干涉检查及功能试验等，某些特殊材料制成的原型还具有较高的强度或光弹特性，可用于空气动力学试验、光弹应力分析、传热试验等。设计人员在对模型进行各种检验、模拟试验分析的基础上，进行可行性评估，根据评估的结果决定对原设计方案是否进行修改，若需要修改产品设计，则进行相应的修改，直至满足设计要求。

4）小批量产品制造　对于塑料制品，还可以利用快速原型作为母模，快速翻制软质模具，例如，硅橡胶模具等，配合真空注型设备与工艺，进行小批量制品的快速制造；对于金属板料制品，利用快速原型作为母模，快速翻制硬质模具，例如，金属喷涂模具等，进行小批量板料制品的制造，以配合整套产品的评价和市场试销等。

5）成形过程工艺和模具设计　当确定了产品的设计方案和市场前景后，就需进行实际生产的工艺和模具设计，在生产工艺和模具设计过程中，通过建立成形过程数值模拟模型，利用有关数值模拟软件，进行成形过程的数值模拟，验证修改成形工艺和模具设计方案，优化材料成形工艺和模具设计方案，以保障生产用工艺与模具设计的可行性。

6）生产用模具加工、试模与投产　当整个设计开发环节结束后，进行批量生产用模具的加工制造、试模、生产，进而完成产品的整个设计、开发和生产过程。

在上述设计开发过程中，数字化设计与制造贯彻于产品的整个开发过程，有效地提高了产品的开发速度、开发质量，减少了开发成本和风险。

1　产品数字化设计方法与平台

目前，基于数字化的产品快速设计有两种主要途径：一种是根据产品的要求或直接根据二维图样在CAD软件平台上设计产品三维模型；另一种是在仿制产品时用扫描机对已有的产品实体进行扫描和拟合，得到三维数字模型（见图4.11-2）。

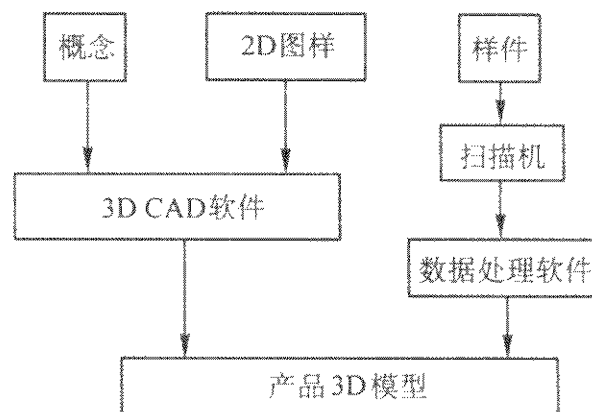


图4.11-2　产品数字化设计基本途径

1）用计算机辅助设计软件进行产品三维设计 目前产品的设计已经大面积地直接采用三维计算机辅助设计软件来构造产品的三维模型，很少通过二维图样设计后，再进行三维设计。对于某种新产品的设计，通过基础的构思或概念，利用三维设计软件，可迅速地设计出三维形状与结构，通过计算机上的三维装配检验、外观评价等，不断修改三维设计方案。当然，也可以设计不同的方案，并通过不同设计方案的比较，选择合适的设计方案。目前应用较多的三维计算机辅助设计软件见表4.11-1。

表4.11-1 国外部分通用的CAD/CAM系统

软件名称	开发公司	国别
Pro/Engineering	Parametric Technology Co.	美国
Unigraphics	Unigraphics Solutions Co.	美国
CATIA	Dassault Systems Co.	法国
Cimatron	Cimatron Co.	以色列
I-DEAS	Structural Dynamics Research Co.	美国
CADDS5	Computervision Co.	美国
DUCT	Delcam Co.	英国
CADAM	Dassault Co.	法国
AutoCAD	Autodesk Co.	美国
SPACE-E	日立造船情报系统株式会社	日本

2）采用反求技术进行产品快速设计 新产品开发过程中的另一条重要路线是样件的反求。反求是对存在的实物模型或零件进行测量，并根据测量数据重构出实物的三维CAD模型，进而对实物进行分析、修改、检验和制造的过程。反求工程主要用于已有零件的复制、损坏或磨损零件的还原、模型精度的提高及数字化模型的检测等。反求的主要方法有三坐标测量法、投影光栅法、激光三角形法、核磁共振和CT法以及自动断层扫描法等。反求工程中较大的工作量就是离散数据的处理。一般来说，反求系统中携带具有一定功能的数据拟合软件，或借用常规的CAD/CAM软件UGII、Pro/E、Delcam等，也有独立的曲面拟合与修补软件，如Surfacer等。采用反求工程方法进行产品快速设计，需要对样品进行数据采集和处理，具体内容见图4.11-3。

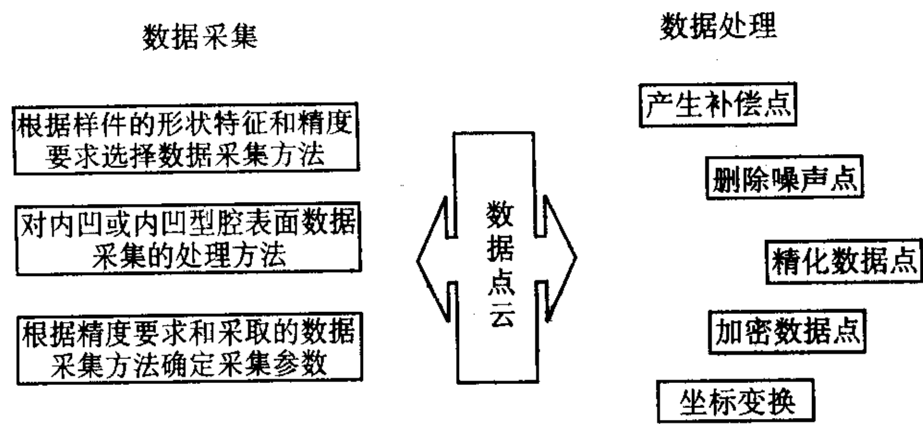


图4.11-3 反求工程中的数据采集与处理技术

2 产品结构分析与优化技术

（1）工程设计中人工试凑法与类比分析法存在的问题

长期以来，工程设计工作者一直沿用经验类比设计方法来处理工程实际问题，即反复通过“设计—分析—再设计”过程，得到一个较为满意的设计方案。

但是在产品开发过程中，有关产品的结构、功能、操作性能、生产工艺、装配性能，甚至维护性能等问题都需要在开发过程前期得以解决。通常，人们借助理论分析、CAD系统和各种比例的实物模型，或参考先前产品的开发经验来解决有关新产品开发中所出现的各种问题。但这种试凑与类比的设计过程，不仅需要花费较长的设计时间，而且仅限于少数候选方案的分析比较。由于设计过程缺乏严格的、科学的定量分析工具，一般很难得到近乎最优的设计方案，特别是对于影响因素很多的复杂工程设计问题，更是如此。

（2）产品结构优化设计理论与方法

1）产品结构优化设计理论与方法 随着计算机技术的发展和应用，20世纪50年代发展起来的以线性规划与非线性规划为主要内容的数学工具——数学规划被应用于解决工程设计问题，形成了工程优化设计理论与方法。

工程优化设计研究与实践的不断深入，使传统的产品设计方法发生了根本性的变革。特别是近年来，随着有限元法、可靠性设计、计算机辅助设计理论与方法的发展，以及与优化设计方法的结合应用，使产品设计过程逐步向自动化、集成化、智能化发展。

工程优化设计，本质上是根据优化设计理论，采用优化设计算法，运用计算机高质量高速度地完成设计任务。

① 将产品结构设计问题转化为数学模型，即用数学表达式描述产品设计问题。

② 按照数学模型的特点选择优化设计方法及其计算程序，运用计算机求得最优解，即最优设计方案。

因此，产品结构优化设计包括建立数学模型和运用优化方法求解两个方面的重要内容。但是对于在复杂工况下工作的工件来说，要建立符合功能、形状、强度和刚度等多目标的优化模型非常困难。即使在诸多简化的基础上，建立了相应的模型，并最终求得“最优解”，其结果也仅能作为产品设计的重要参考和理论依据。

2）产品结构设计 产品的结构设计既要满足产品的外观及加工制造等要求，更需要满足产品的使用要求。产品设计经外观及加工可行性评估后，能否满足在给定工况下的使用要求，需要进行结构分析。

结构分析的主要内容有，结构件的强度、刚度、稳定性及疲劳寿命等。结构分析的目的主要是进行安全性评估，进而实现结构的优化，根据优化的结果，修改初始设计，使产品不仅具有可行的外观和结构，而且满足强度、刚度、稳定性及其疲劳寿命等要求。

目前，应用较多的结构分析软件有ANSYS、NASTRAN、MARC、ADINA、SAP系列等。

3）产品优化设计 产品的结构优化设计是在一定约束条件下设计满足目标功能的优化几何尺寸、材料特性等结构性能的设计方法。但目前的优化设计分析仅涉及结构尺寸，大多不具备优化几何形状或拓扑形状的功能。

在产品设计过程中，随着对产品的功能要求、外观要求、装配要求的增多，产品的设计自由度越来越小。如果一味追求设计的最优化，可能要付出昂贵的代价，甚至产品设计的失败。因而，按照产品开发的技术经济要求，并不一定要使产品设计在各方面达到最优。

为此，可以结合当今出现的现代先进设计制造技术，运用基于虚拟设计和快速成形技术，实现产品设计开发的快速响应和提高产品的性能，降低设计成本和开发风险。同时，在设计初期即能发现产品使用中可能出现的问题，及时纠正设计，从而实现产品设计的“一次性”成功。

4）产品装配性设计和使用性设计 由于有关装配、操作和维修的问题，往往只会在产品开发的后期或在最终产品试车过程中、甚至在投入使用一段时间后才能暴露出来，尤其是有关维修的问题往往是在产品已经售出很长时间以后才能被发现。

为了解决这些问题，有时产品就不得不返回到设计构造

阶段以便进行必要的设计变更，这样的产品开发程序既低效、耗时，费用又高。

产品本身的功能可以通过模拟方法测试和检验，然而要预先测试和检验人机交互性能是非常困难的。虚拟设计技术可以提供解决这类问题的方法。即便是在设计的初期阶段，计算机产生的最初模型也可以放入虚拟环境进行实验。这样不仅可以使产品的外表、形状和功能得到模拟，而且有关产品的人机交互性能也能得到测试和检验，使产品的缺陷在当时的设计阶段就能被及时发现并加以解决。

(3) 基于虚拟设计和快速成形的产品结构分析和优化设计技术

基于虚拟设计和快速成形的产品结构分析和优化设计技术的核心是追求产品设计的快（开发速度快）、高（设计产品的性能高）、低（设计风险、成本低）、准（一次性成功）。它实际上实现了虚拟设计、快速成形技术和计算机辅助分析技术的有机集成，在产品开发设计中取得了非常好的技术经济效益和极大的社会示范效应。

3 产品结构的快速成形与快速模具制造技术

从20世纪90年代开始，市场环境发生了巨大变化。企业产品开发步入了“三个三”（产品设计周期三周，试制周期三个月，生产周期三年）。企业的发展战略已经从60年代“如何做得更多”、70年代“如何做得更便宜”、80年代“如何做得更好”发展到90年代的“如何做得更快”。因此，快速响应市场需求，已成为制造业发展的重要走向。于是，上世纪80年代末期，快速原型制造技术和基于快速原型的快速模具制造技术一经产生，因其在产品开发和单件小批量产品制作中显著的经济效益和时间效益，而受到广泛关注并得到迅速发展。

快速成形（也称快速原型）制造技术（Rapid Prototyping & Manufacturing，RP&M 或 RP）是借助计算机、激光、精密传动和数控等现代手段，将计算机辅助设计（CAD）和计算机辅助制造（CAM）集成于一体，根据计算机上构造的三维模型，能在很短时间内直接制造产品样品而无需传统的机械加工机床和模具的先进制造技术。该项技术创立了产品开发的新模式，使设计师以前所未有的直观方式体会设计的感觉，感性而迅速地验证和检查所设计产品的结构和外形，从而使设计工作进入了一种全新的境界，改善了设计过程中的人机交流，缩短了产品开发的周期，加快了产品更新换代的速度，降低了企业投资新产品的风险。

而以 RP&M 原型作母模来翻制模具的快速模具制造技术（Rapid Tooling，RT），进一步发挥了快速原型制造技术的优越性，可在短期内迅速推出满足用户需求的一定批量的产品，大幅度降低了新产品开发研制的成本和投资风险，缩短了新产品研制和投放市场的周期，在小批量、多品种、改型快的现代制造模式下具有强劲的发展势头，基于快速原型工艺方法制造各类简易经济快速模具已成为 RP&M 应用的热点问题。

3.1 快速原型制造技术及其分类

快速原型制造技术的基本原理是基于离散的增长方式成形原型或制品。快速原型制造技术的早期根源可以追溯到100多年前的地形学工艺领域，当时的学者采用叠层蜡片制作了立体地貌图，后来许多学者都根据这种“增长”型的制造方式提出了针对不同材料的不同的建造技术，陆续出现了近百种有关的专利技术，有力地促进了该项技术的发展。20世纪80年代以来，快速原型制造工艺商品化系统的研制加紧了步伐，自从1988年美国 3DSystems 公司推出第一台商品化快速原型制造设备 SLA－250 以来，目前与快速成形工艺方法相对应的已经商品化的快速成形设备不下10余种，同时还陆续出现了许多在实验室层面上比较成熟的快速原型工艺方法。

快速成形制造技术从广义上讲可以分成两类：材料累积和材料去除。但是，目前人们谈及的快速原型制造方法通常指的是累积式的成形方法，而累积式的快速原型制造方法通常是依据原型使用的材料及其构建技术进行分类的，如图4.11-4所示。

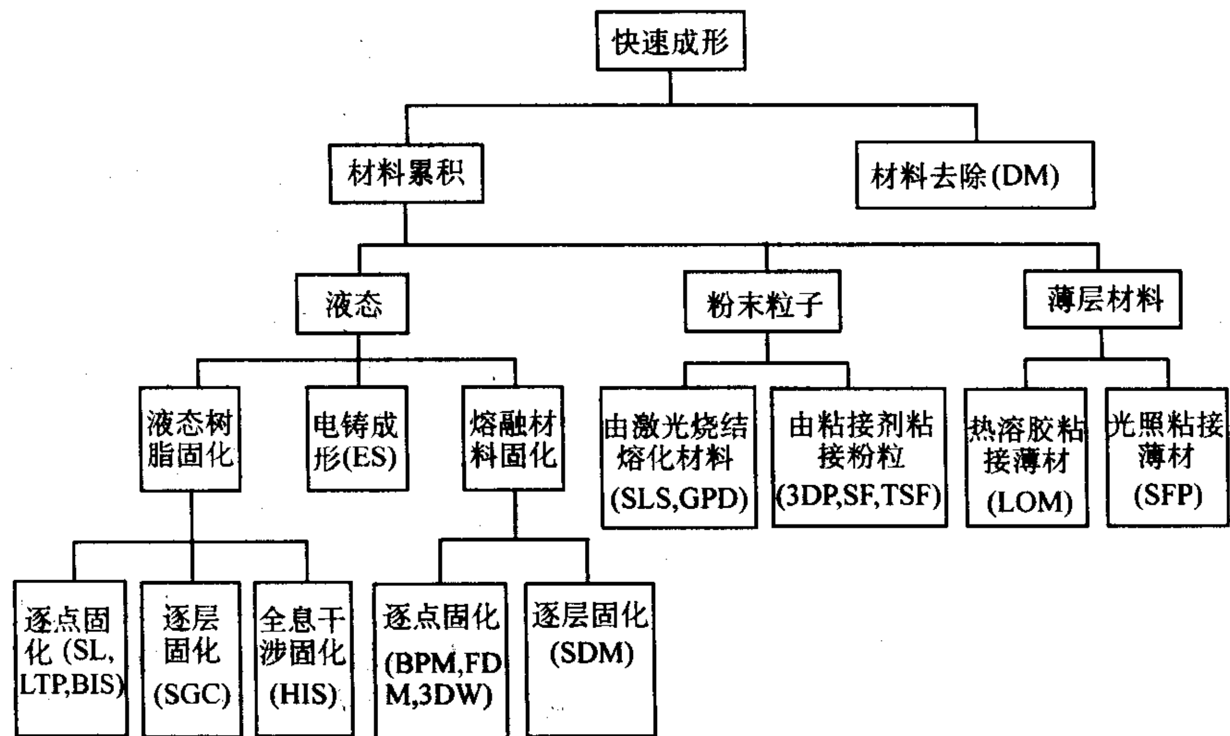


图4.11-4 快速成形工艺方法的分类

3.2 快速模具制造技术及其分类

快速原型由于其制造方法要求的使用材料的限制，并不能够完全替代最终的产品。因此，在新产品功能检验、投放市场试运行以获得用户使用后的反馈信息以及小批量生产等方面，仍需要由实际材料制造的产品。因此，利用快速原型作母模来翻制模具并生产实际材料的产品，便产生了基于快速原型的快速模具制造技术（Rapid Tooling，RT）。

基于 RP 的快速模具制造方法一般分为直接法和间接法两大类。直接制模法是直接采用 RP 技术制作模具，在 RP

技术诸方法中能够直接制作金属模具的是选择性激光烧结法(SLA)。用这种方法制造的钢铜合金注射模，寿命可达 5 万件以上。但此法在烧结过程中材料发生较大收缩且不易控制，故难以快速得到高精度的模具。目前，基于 RP 快速制造模具的方法多为间接制模法。间接制模法是指利用 RP 原型间接地翻制模具。依据材质不同，间接制模法生产出来的模具一般分为软质模具（Soft Tooling）和硬质模具（Hard Tooling）两大类。

软质模具因其所使用的软质材料（如硅橡胶、环氧树脂等）有别于传统的钢质材料而得名，由于其制造成本低和制作周期短，因而在新产品开发过程中作为产品功能检测和投入市场试运行以及国防、航空等领域单件、小批量产品的生产方面受到高度重视，尤其适合于批量小、品种多、改型快的现代制造模式。目前提出的软质模具制造方法主要有硅橡胶浇注法、金属喷涂法、树脂浇注法等。

软质模具生产制品的数量一般为 50 ~ 5 000 件，对于上万件乃至几十万件的产品，仍然需要传统的钢质模具，硬质模具指的就是钢质模具，利用 RP 原型制作钢质模具的主要方法有熔模铸造法、电火花加工法、陶瓷型精密铸造法等。图 4.11-5 给出了基于 RP 的快速模具方法的分类及应用。

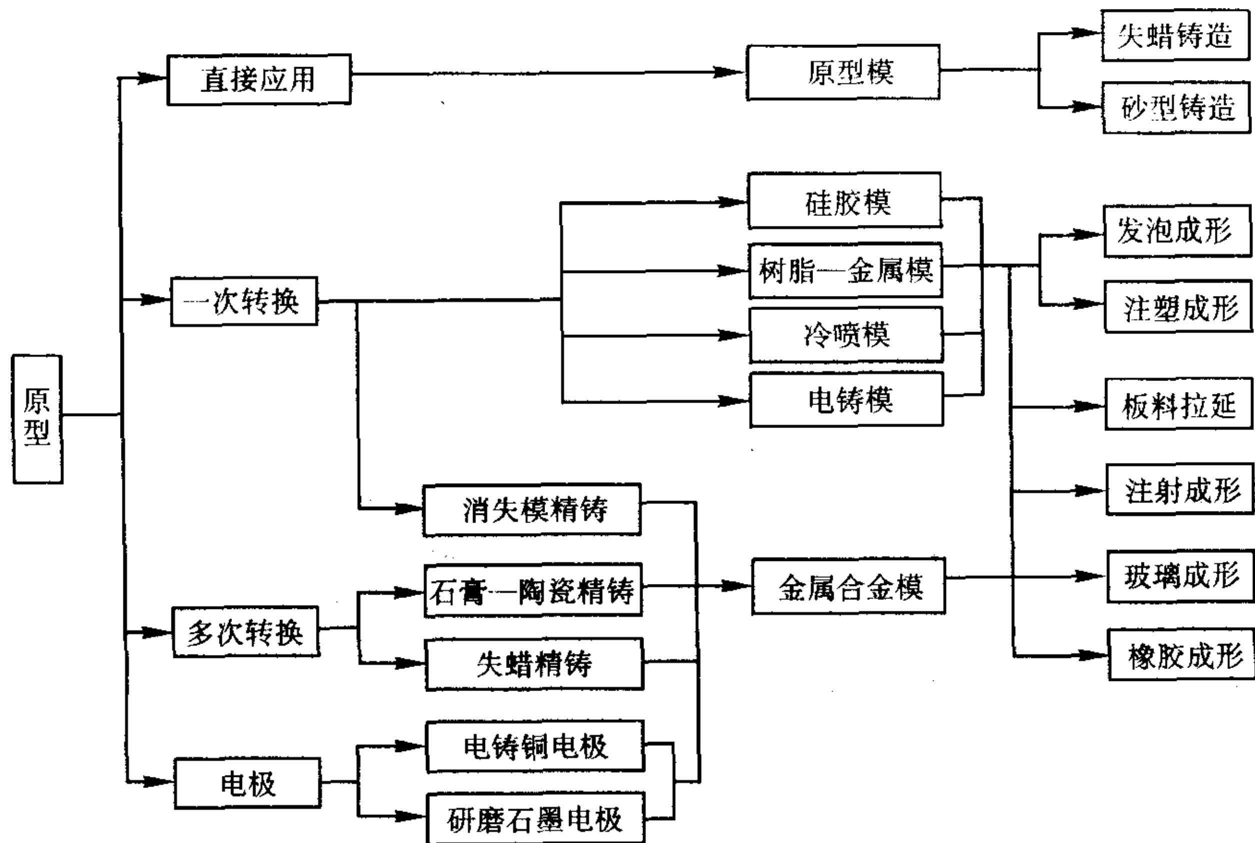


图 4.11-5　基于快速原型的快速模具制造方法

运用 RT 技术突出的特点就是其显著的经济效益，它与传统的数控加工模具方法相比，周期和费用都降低 1/3 ~ 1/10 左右，如表 4.11-2 所示。近年来，工业界对 RT 的研究开发投入了日益增多的人力和资金，RT 的收益由此也获得了巨大增长。据 SME 统计，近年来 RT 服务的收益年增长率均高于 RP 系统销售，如 1996 年比 1995 年增长 62.7%，而 RP 设备销售额的增长为 42.6%。

表 4.11-2　基于 RP&M 原型快速制作模具与传统机加工法制作模具的比较

制模方法	制作成本/美元	制作周期/周	模具寿命/件
硅橡胶浇注法	5	2	30
金属树脂浇注法	9	4 ~ 5	300
电弧热喷涂法	25	6 ~ 7	1 000
镍蒸发沉积法	30	6 ~ 7	5 000
传统机加工方法	60	16 ~ 18	250 000

4　产品构件成形工艺与模具虚拟仿真优化设计技术

随着有限元等离散数值模拟技术的逐渐成熟和计算机软硬件的迅速发展，材料成形过程模拟在原有的数值分析方法的基础上得到了更加广泛的应用。针对各类成形工艺中材料的流动变形所遵循的基本原理，逐步形成了完整的处理技术和模拟方法，并相继出现了各种商品化的工艺仿真软件平台。

材料成形数值模拟与优化技术及其工艺过程仿真软件平台的广泛应用，强有力地改造着传统的依赖于经验式的工艺设计模式，为产品的数字化工艺设计、工装模具设计与数字化制造提供了重要技术支撑和决策咨询。

目前，产品成形的工艺仿真软件，体积成形仿真软件有 DEFORM、FORGEIII、QFORM 等，板料成形仿真软件有 DYNAFORM、OPTRIS、PAM - STAMP 等，注塑成形仿真软件有 MOLDFLOW、C - MOLD 等。

(1) 体积成形过程仿真技术和模拟平台

对于锻造、挤压、镦挤、拉拔和轧制等金属材料体积成形工艺，材料发生较大的塑性变形，而弹性变形相对极小，可忽略不计。因此，通常采用的研究工具是刚塑性或刚黏塑性有限元法。

刚（黏）塑性有限元法遵循 Markov 变分原理，即假设变形体由刚塑性材料构成，服从 Mises 屈服准则及其相关联的流动准则（即 Levy—Mises 本构关系），在所有满足动可容条件的速度场中，真实的速度场使变形体的总势能取最小值。

根据不同的体积不变条件，其处理方法有 Lagrange 乘子法、体积可压缩法和罚函数法。在对实际成形工艺实施刚塑性有限元模拟时，需要处理刚性区、初始速度场生成、摩擦条件和边界条件的施加、网格划分与再划分及计算结果的可视化等问题。

20 世纪 90 年代以来，国际上出现了许多通用体积成形刚 - 黏塑性有限元数值模拟的商业软件，这些软件以其友好易用的界面和可靠的性能在世界各地的科研院所及锻造企业中得到了广泛的应用（见表 4.11-3）。

表 4.11-3　商用锻造过程数值模拟软件系统

软件系统	维数	国别	开　发　商
EESY-2-FORM EESY-FORM	2 3	德国	CPM Gesellschaft fur Computer-anwendung
DEFORMPC DEFORM2D DEFORM3D DEFORMHT	2 2 3 2	美国	Scientific Forming Technologies Corporation
FORGE2 FORGE3 FORGE3PC	2 3 3	法国	Transvalor SA
MARC Autoforge	2 3	美国	MARC Analysis Research Corporation
MSC Superforge	3 (2)	荷兰	Technical Project Coordinator
Qform	2	俄罗斯	Quantor Ltd
CASFORMPC	2	中国	山东大学模具工程技术研究中心

金属材料体积塑性成形模拟商品化软件一般能够模拟各种体积成形过程，如锻造、挤压、拉拔、镦挤和轧制过程，能够预测缺陷生成，验证和优化工艺与模具设计方案，能够模拟等温、非等温、单工序和多工序成形过程，适合于刚塑性、刚黏塑性、刚体材料模型，冷、温、热成形和液压机、锤、摩擦压力机、机械压力机等设备。在成形工艺和模具设计中，使用模拟软件进行过程仿真，可实现工艺与模具设计方案的虚拟验证。

金属材料体积塑性成形模拟商品化软件一般包括以下模块（见图 4.11-6）。

1）前处理模块　供用户迅速方便地输入参数，其几何参数可以立刻以图形形式显示出来。对于三维软件，还可提供由三维 CAD 软件几何模型数据输入接口。

2）后处理模块　供用户处理和显示结果的图形，一般包括变形过程显示，应力、应变、应变速率、温度等彩色云图和等值线显示，速度场速端图显示及其载荷-行程曲线显示。

3）网格自动生成和再生成模块　生成二维四节点单元、三维四面体和三维六面体的有限单元网格。要求网格密度容易控制，单元大小过渡均匀，生成计算过程迅速，完全自动实现。一般采用动态内存，以使网格数目不受限制。

4）有限元分析模块　包括二、三维成形过程等温和非等温分析，适合于任意模具形状。

5）材料数据库　提供通过实验获得的完备和准确的材料数据，包括金属材料流动应力、变形速率、等效应变和温度的关系曲线或数据。材料数据库应具有修改、删除和添加的功能。模拟结果数据库提供了模拟结果数据处理的统一接口。

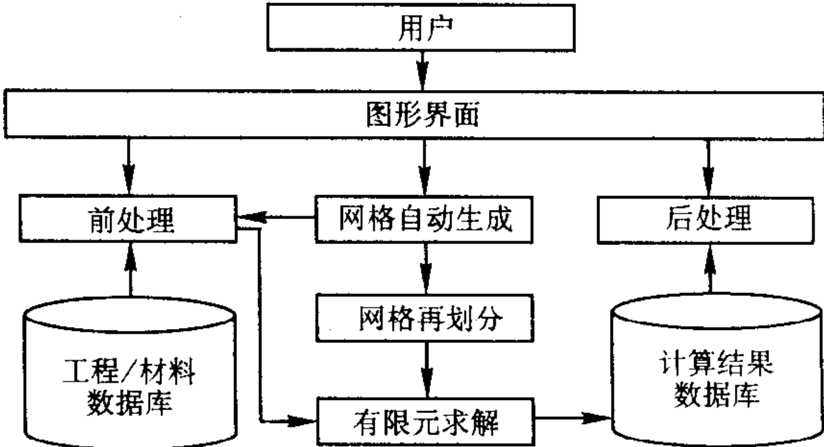


图 4.11-6　体积成形过程有限元模拟系统基本框架

（2）板料成形仿真技术与软件

板料成形技术被广泛用于压力容器、汽车、船舶、飞行器的外壳等大型金属薄壁结构件以及各种形状的日常用品的制造。传统的模具设计是一个反复的研制过程，旷日持久而又投资巨大。随着工业化的日益发展，产品更新换代周期越来越短，新材料的不断出现，传统的设计方法已明显不适应现代工业的发展要求。

20 世纪 70 年代，开始有人提出板料成形分析的计算机数值模拟。近 10 年来，一个融计算机图形学、塑性力学、有限元法，能够指导板料成形设计的 CAE 系统已逐步建立起来。

一个典型的板料成形 CAE 系统构成见图 4.11-7。这样的系统的目的在于以图形与数据的形式，动态地演示出成形过程的发展预测工件中发生的应力、可能出现的折皱和破裂、成形力以及回弹等，使得设计人员可以在成形尚未实现的生产准备阶段，预见成形中可能出现的问题，确定成形所需的工艺过程和技术参数，验证并优化模具设计。

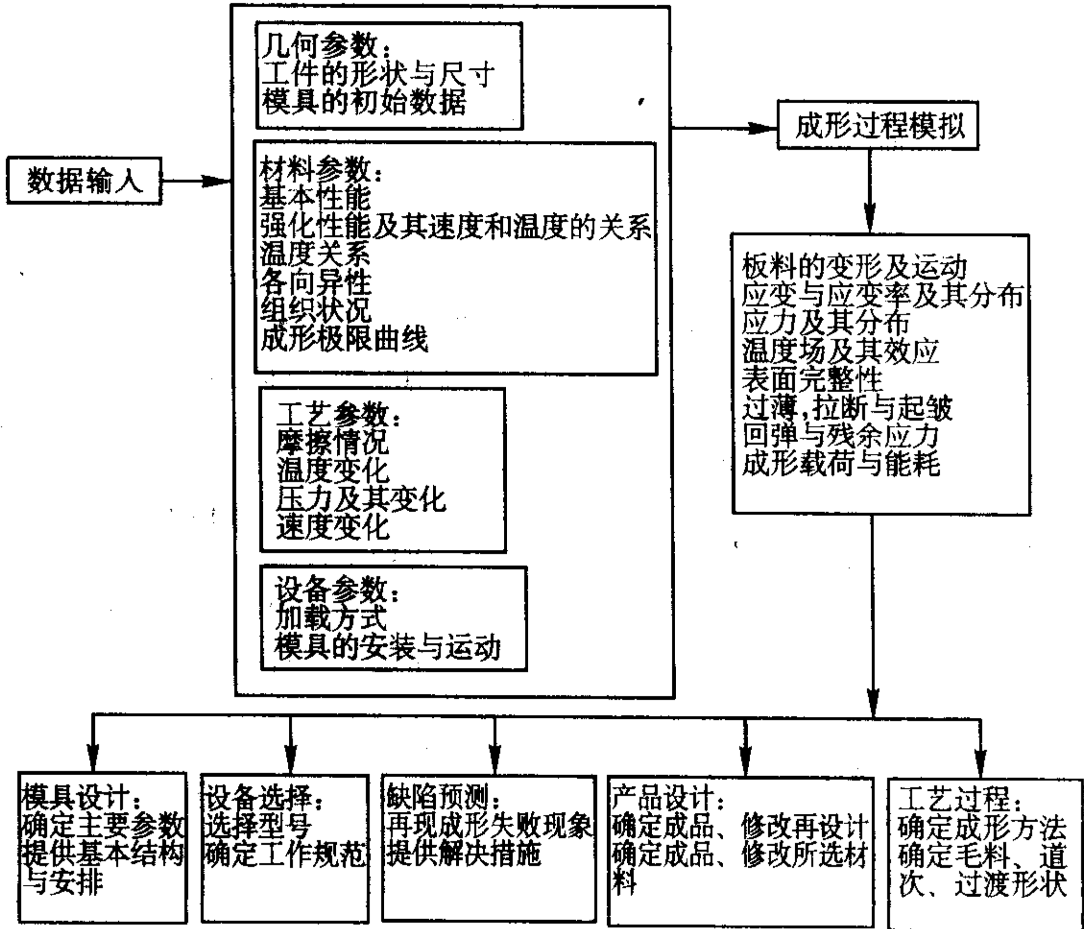


图 4.11-7　板料成形 CAE 系统

板料成形数值模拟技术经历了 20 余年的发展，20 世纪 90 年代初逐步走向实用化。随着计算机软硬件技术的发展，众多融合了计算机图形学、有限元技术、塑性成形数值模拟的软件开始出现，得到工业部门的重视和应用。目前能够进行板料成形数值模拟的主要软件见表 4.11-4。

表 4.11-4　部分板料成形数值模拟软件

软件名称	开发商	国家
LS - DYNA3D	Livemore 软件公司	美国
PAM - STAMP	ESI 公司	法国
ABAQUS/Explicit	HKS 公司	美国
OPTRIS	DYNAMIC SOFTWARE 公司	法国
SHEET - 3	Ohio 州立大学	美国
ROBUST	大阪大学	日本

(3) 注塑成形分析技术与软件

计算机辅助工程（CAE）技术已成为塑料产品开发、模具设计及产品加工的最有效工具。采用 CAE 技术，可以完全代替试模，CAE 技术提供了从制品设计到生产的完整解决方案，在模具制造之前，预测塑料熔体在型腔中的整个成形过程，帮助研判潜在的问题，有效地防止问题发生，大大缩短了开发周期，降低生产成本。此外，采用注塑成形 CAE 技术还可以实现优化塑料制品设计、优化塑料模设计及优化注塑工艺参数等。

国内外在注塑分析软件开发中取得商品化的研究成果，见表 4.11-5 所示。

表 4.11-5　注塑 CAE 软件及研发单位

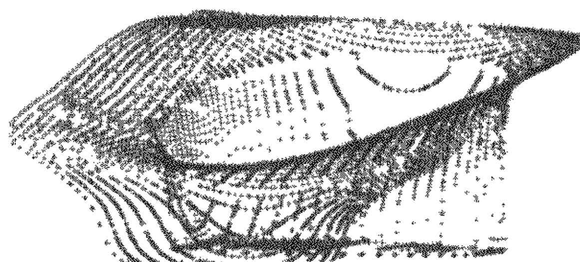
注塑 CAE 软件	开发单位	地　址
Moldflow	美国 MoldflowCorp.	http://www.moldflow.com
C - mold	美国 Cornell 大学	
Z - mold	国家橡塑模具工程研究中心	http://www.zzu.edu.cn/nerc
华塑 CAE	华中科技大学模具技术国家重点实验室	http://www.hust.edu.cn/chinese/research/modlab.htm http://www.e - mold.com.cn/products.htm

5　产品设计与制造的集成技术应用实例

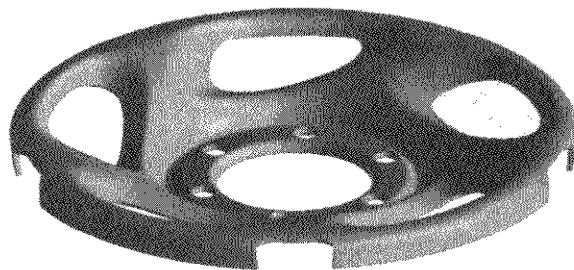
(1) 吉普车车轮 RE/CAD/RP/结构分析与优化

山东大学模具工程技术研究中心利用 RENISHAW 反求系统对车轮样件进行数据反求，并根据要求进行外观及结构的改进设计。

吉普车车轮的 5 个风孔均布，因而只测量其中之一。测量的点云数据及根据测量数据设计的车轮见图 4.11-8。在车轮外观设计过程中，为便于用户评估，对每一种设计都进行了 LOM 原型的制作（见图 4.11-9）。根据评估意见和车轮的径向疲劳、弯曲疲劳的有限元仿真（见图 4.11-10），对车轮的外观及结构进行了数次改进，同时，考虑了车轮冲压成形工艺、工艺顺序等对其力学性能的影响，以及车轮形状与冲压工艺和模具套数、模具结构复杂程度等之间的关系，最终确定了合理的设计方案，确保车轮不仅具有生动、令人赏心悦目的外观，而且能够通过疲劳试验，从而实现满足车轮功能、外观等多方面要求的优化设计。

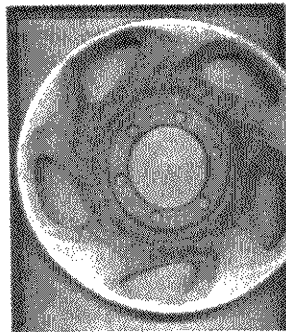


(a) 风孔数据点云

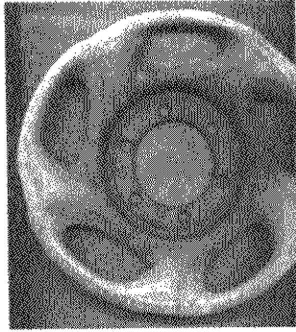


(b) 车轮三维造型

图 4.11-8　车轮反求

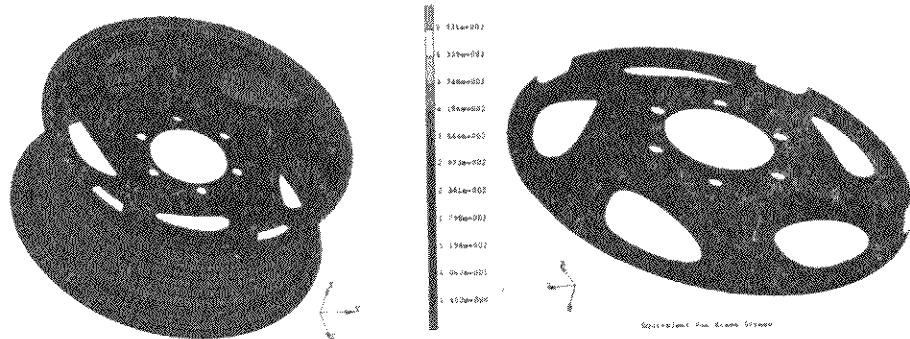


(a) 第一次设计



(b) 经改后的设计

图 4.11-9　车轮外观及结构设计原型



(a) 车轮有限元模型　(b) 轮辐的整体应力分布

图 4.11-10　车轮的有限元分析

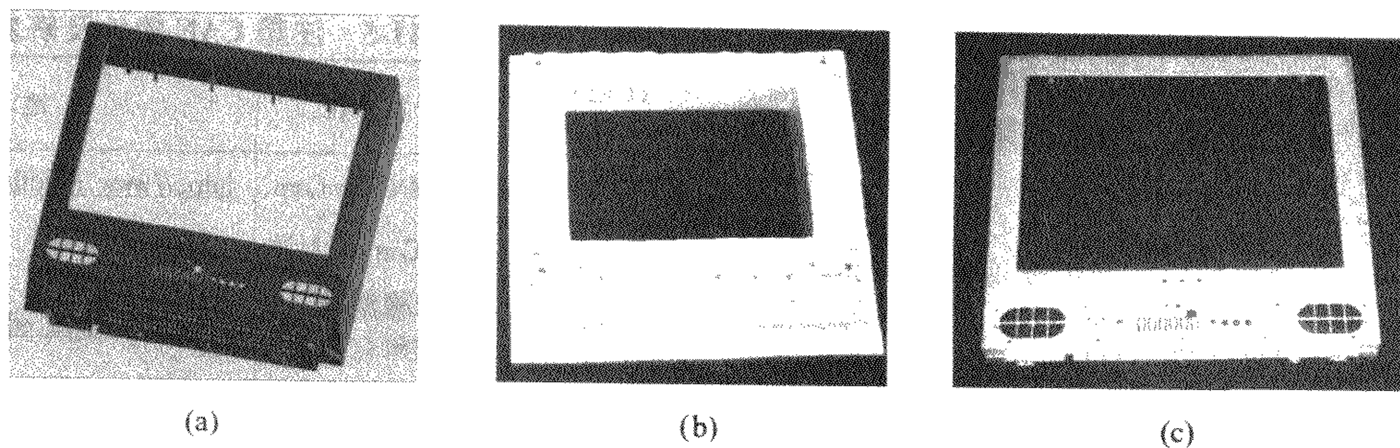
(2) 电视机壳体三维设计与快速产品制造

图 4.11-11 所示为山东大学模具工程技术研究中心为山东松下影像公司开发的 21 in 电视机壳体。其中图 4.11-11a 为采用 UGII 设计的电视机壳体的三维造型，图 4.11-11b 为采用快速原型及真空注型技术快速制作的硅橡胶模具的上模，图 4.11-11c 为利用硅橡胶模具采用真空注型快速制作的类似于工程塑料的电视机壳体的树脂件。整个产品从设计到样件的制作仅仅用了 10 天的时间，显著地缩短了产品的开发周期，同时，通过与其他部件的安装，迅速地将该型号的电视机样机推向了市场。

(3) 锻造/挤压过程模拟实例

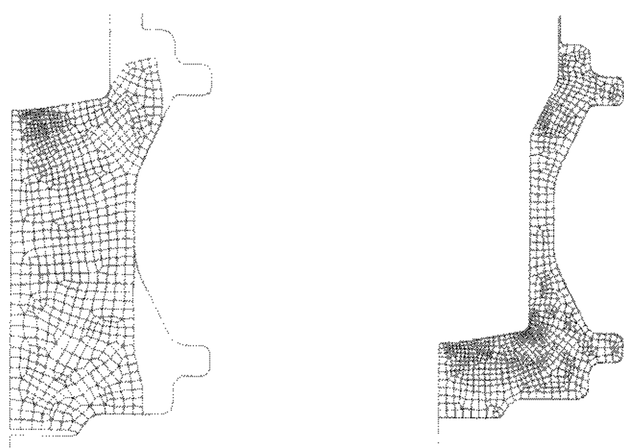
1) 轴对称锻件挤压成形　该挤压过程为非等温成形过程，模具的速度不是恒定的，工件的流动应力是根据实验测得的数据，模具和工件材料的热力参数也不是恒定的。采用山东大学模具工程技术研究中心的 CASFORM 体积成形有限元分析软件对某产品的整个成形过程进行了非等温模拟（见图 4.11-12）。

2) 缺陷预测　锻件在成形过程中由于坯料形状尺寸、成形工艺方案以及模具结构等原因，会出现充不满和折叠等缺陷，缺陷严重时会造成产品的报废，如果检测不到折叠等



(a)　(b)　(c)

图 4.11-11　电视机壳体产品的快速开发

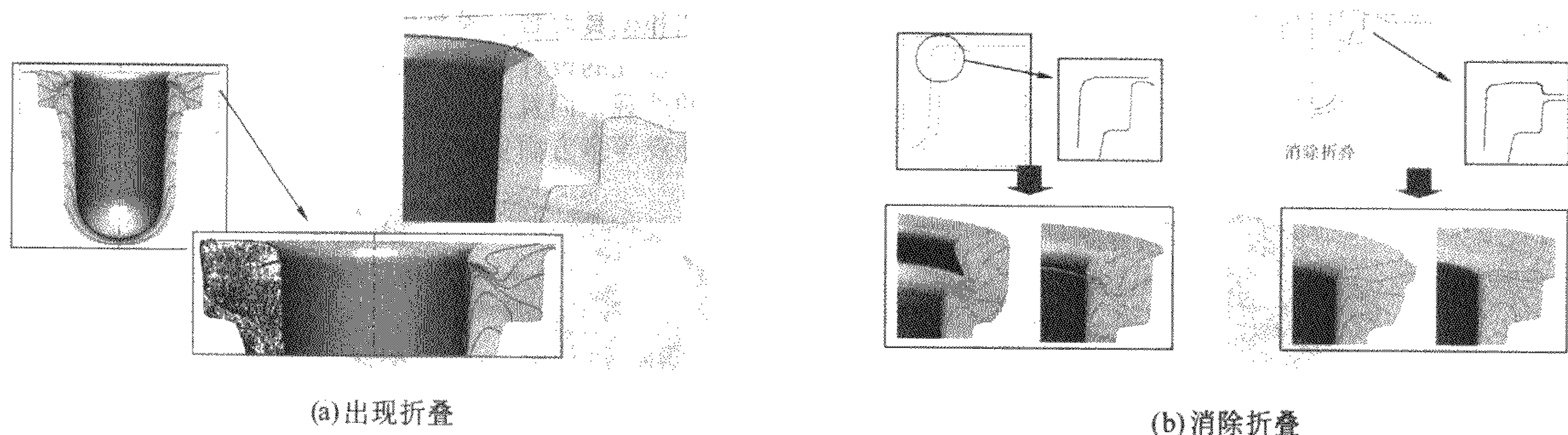


(a)行程为24.4mm时的网格变形图　(b)行程为80mm时的网格变形图

图 4.11-12　利用 CASFORM 软件模拟的不同行程时的网格变形图

缺陷而随后投入到使用中，折叠实际上就是裂纹的一个扩展源，在使用过程中会造成预想不到的事故。所以，折叠等缺陷必须要避免。

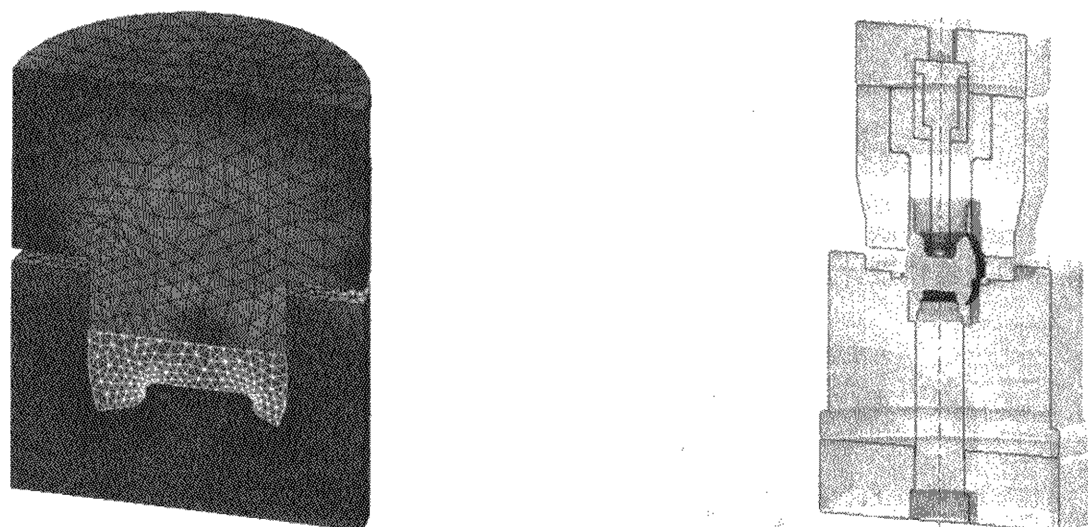
有限元模拟技术能够模拟锻件成形过程中材料的整个流动过程，能够预报回流和汇流等折叠缺陷，并通过修改工艺和模具设计方案，寻找出消除折叠的具体措施。图 4.11-13 给出了采用 Qform2D 软件预报锻件成形中的折叠缺陷和消除缺陷的改进措施。有限元数值模拟技术可以有效地在模具制造之前，在计算机上实现虚拟制造，及时发现工艺和模具设计中的问题并加以修改，避免造成不必要的时间、成本等的浪费，确保设计与制造的可靠性和准确性。



(a)出现折叠　(b)消除折叠

图 4.11-13　Qform 软件预报锻件成形缺陷及消除缺陷的措施

3）模具应力分析　在锻造成形过程中，锻模传递和承受着成形过程中的载荷，结构越复杂的模具其受力状况越复杂。对锻件成形过程中的模具进行应力状况分析和磨损分析，根据分析结果改善模具结构，可以提高模具的使用寿命。图 4.11-14 给出了利用 Qform3D 分析的某一凸轮锻件在成形过程中的模具应力状况。模拟发现，采用整体结构的模具受力状况不好，进而改进设计为楔形镶块模加预应力箍套，这样可大大提高模具寿命。

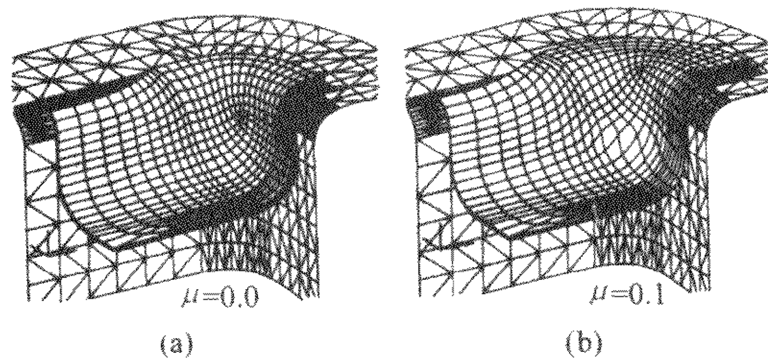


(a)凸轮锻造结束时上模的等效应变情况　(b)模具结构的改进

图 4.11-14　Qform3D 给出的凸轮锻件在成形过程中的模具分析

（4）板料成形过程模拟实例

板料成形工艺仿真较典型的应用实例是方盒件拉延。模具采用三角形表面单元离散，板料尺寸 170 mm × 110 mm × 0.75 mm，由于对称，分析时取 1/4。该算例主要是研究摩擦条件对材料流动、压边力、冲压力的影响。模拟分为两种情况：摩擦因数分别为 0 和 0.1。图 4.11-15 显示了冲头行程达到 30 mm 时板料的网格变形。在无摩擦条件下，突缘的变形较均匀，而在有摩擦情况下，摩擦力阻碍了方盒四角部分突缘材料的流动。通过模拟，一方面掌握金属流动情况，另一方面，及时发现设计问题并加以修改，验证工艺和模具设计的可行性。



(a)　(b)

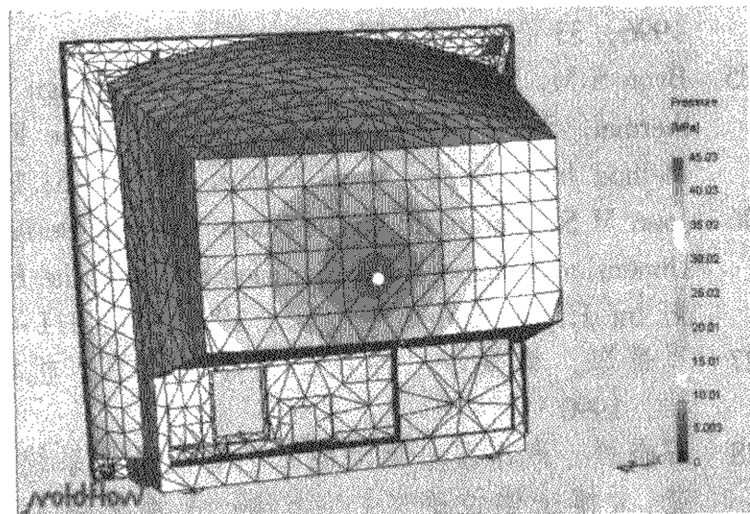
图 4.11-15　冲头行程 30 mm 时的网格变形

（5）注塑过程模拟实例

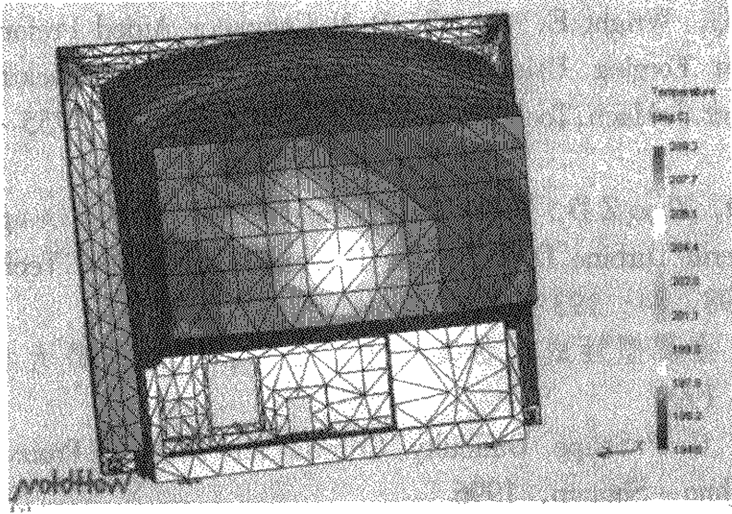
电视机后盖为电视机壳体中相当重要的注塑件，模具制造成本和周期较长。为了制定合理的工艺参数、确保模具制造的成功和减少试模与修模过程，采用注塑 CAE 系统进行相应的仿真分析是十分必要的。通过采用 Moldflow 对电视机后盖流动充填过程中的压力分布、温度分布、填充时间、产生的气穴以及熔接痕等进行分析，确定了适当的浇口数量，通过调整浇口位置及壁厚，达到了流动平衡，获得了较好的充填效果，保证了零件质量，避免了试模与修模的过程，缩短了制模周期，降低了成本。图 4.11-16 给出了其中的压力分布与温度分布分析结果。

本章主要介绍了产品数字化设计及材料成形工艺仿真与优化集成技术，首先给出了以数字化设计与制造为主线的集成系统的构成，简要介绍了各单元技术的国内外现状、设计软件平台及其在产品开发设计中的作用，结合具体产品开发实例说明了数字化设计与制造技术的应用前景和作用。

材料成形中的数字化产品开发设计与制造技术正处于蓬勃发展的时期，已经或正在发挥着显著作用，真正迎合了现代产品开发快捷、可靠、准确、低廉的要求，适应了目前小批量、多品种的产品开发要求。与国外相比，我国在数字化设计、工艺过程仿真和集成技术方面仍然存在着较大的差距，我们不仅要加强自主版权的数字化设计平台的开发，而且要加强自主材料成形过程模拟软件平台的开发；不仅要高度重视单元技术的不断提升，而且要加强单元技术的有效集成，以发挥各单元技术的特长和集成技术的综合效益。同时，加强数字化设计与制造技术的推广应用，建立相应的技术支撑体系和培训体系，使数字化设计与制造技术得到广泛应用，发挥其在我国经济建设中的巨大作用。



(a) 压力分布



(b) 温度分布

图 4.11-16　电视机后盖注塑工艺分析

材料加工工程分析和仿真软件在国外已经具有了良好的市场和应用，我国尚未很好地进行自主开发和应用。预计在 5 ~ 10 年内，我国会有较大的需求和市场。计算机辅助工程软件是一个典型的专家级软件，开发者必须具有坚实的基础理论知识和广阔的专业知识，需要长期、坚持的开发工作，因此，计算机辅助工程软件的开发不是一般人员所为，也正因为此，我国计算机辅助工程软件的开发不仅一直落后于国外，而且一直落后于我国其他种类软件的开发水平和数量。计算机辅助工程软件包括软件本身开发、市场推广与技术支持，单就软件开发而言，主要包括前后处理、网格自动划分和分析程序，其中前后处理和网格自动划分部分是仿真软件开发的重要前提和通用平台，如果可靠的 2D 和 3D 前后处理与网格自动划分程序完备，各类工程问题的数值模拟商品化软件的开发进程就会大大提高，类别也会大幅度拓展，这应当值得我们高度重视。

编写：赵国群（山东大学）
王广春（山东大学）

参考文献

1 曾昭华，傅祥志．优化设计．北京：机械工业出版社，1992

2 赵国群，阮雪榆，关廷栋．预成形设计的新方法－反向模拟技术．模具技术，1991，(3)：1～7

3 赵国群，阮雪榆，关廷栋．锻造过程的反向模拟及预成形设计．模具技术，1992（4）：1～6

4 Zhao G Q，Wright E，Grandhi R V. Preform Design in Bulk Forming Process. Proceeding of 5th International Conference on Technology of Plasticity. Columbus，Ohio，U.S.A，1996，453～456

5 Zhao G Q，Wright E，Grandhi R V. Forging Preform Design with Shape Complexity Control in Simulating Backward Deformation. Int. J. Mach. Tools Manufact.，1995，35（9）：1225～1239

6 Zhao G Q，Wright E，Grandhi R V. Preform Design in Forging Processing Using Nonlinear Finite Element Method. Transactions of NAMRI/SME，1994，12：17～24

7 Zhao G Q，Wright E，Grandhi R V. Computer Aided Preform Design in Forging Using the Inverse Die Contact Tracking Method. Int. J. Mach. Tools Manufact.，1996，36（7）：755～769

8 Zhao G Q，Zhao Z D，Wang T H，Grandhi R V. Preform Design of a Generic Turbine Disk Forging Process. J. Mater. Proc. Technol.，1998，84：193～210

9 彭颖红．金属塑性成形仿真技术．上海：上海交通大学出版社，1999

10 Teterin G P. Shape Difficulty Criteria for Forgings. Proizv：Kuznechno－Shtamp，1996

11 Badrinarayanan S，Zabaras N. A Sensitivity Analysis for the Optimal Design of Metal Forming Process. J. Mater. Proc. Technol.，1994，129：83～104

12 Badrinarayanan S，Zabaras N. Preform Design in Metal Forming. Processings of NUMIFORM95，Ithaca，NY，1995，533～538

13 Badrinarayanan S，Zabaras N. A Sensitivity Analysis for the Design of Metal Forming Processes. Computer Meth. Appl. Mech. Eng.，1996，129：319～348

14 Zhao G Q，Wright E，Grandhi R V. Preform Die Shape Design in Metal Forming Using an Optimization Method. Int. J. Numer. Methods Eng.，1997，40：1213～1230

15 Zhao G Q，Wright E，Grandhi R V. Preform Sensitivity Analysis Based Preform Die Shape Design for Net－Shape Forging. Int. J. Mach. Tools Manufact.，1997，37（9）：1251～1271

16 Zhao G Q，Huff R，Hutter A，Grandhi R V. Sensitivity Analysis Based Preform Die Shape Design Using the Finite Element Method. J. Mater. Eng. Perform.，1997（6）：303～310

17 Zhao G Q，Ma X W，Zhao X H，Grandhi R V. Studies on Optimization of Metal Forming Processes Using Sensitivity Analysis Methods. Journal of Materials Processing Technology，2004，147（2）：217～228

18 马新武．体积成形过程数值模拟与优化技术及其系统开发研究：[博士学位论文]．济南：山东大学，2002

19 赵新海．基于有限元模拟的锻造过程优化设计方法研究：[博士学位论文]．济南：山东大学，2001

20 Lee S K，Ko D C，Kim B M. Optimal Die Profile Design for Uniform Microstructure in Hot Extruded Product. Int. J. Mach. Tools Manufact.，2000，40：1457～1478

21 Gao Z Y，Grandhi R V，Microstructure Optimization in Design for Forging Processes. Int. J. Mach. Tools Manufact.，2000，40：691～711

22 王广春，管婧，马新武，赵国群．金属塑性成形过程的微观组织模拟与优化技术研究现状．塑性工程学报，2002，1：1～5

23 Joun M S，Hwang S M. Optimal Process Design in Steady－state Metal Forming by Finite Element Method—I. Theoretical Considerations. Int. J. Mach. Tool. Manufac.，1996，33（1）：51～61

24 Joun M S，Hwang S M. Optimal Process Design in Steady－state Metal Forming by Finite Element Method—II：Application to Die Profile Design in Extrusion. Int. J. Mach. Tool. Manufac.，1996，33（1）：63～70

25 Byon S M，Hwang S M. Process Optimal Design in Non－Isothermal，Steady－State Metal Forming by the Finite Element Method. Int. J. Numer. Mech. Engng，1999，46：1075～1100

26 Joun M S，Hwang S M. Die Shape Optimal Design in Three－Dimensional Shape Metal Extrusion by the Finite Element Method. Int. J. Numer. Meth. Engng.，1998，41：311～335

27 肖景容，姜奎华．冲压工艺学．北京：机械工业出版社，1990

28 石晓祥．汽车覆盖件冲压工艺智能设计系统关键技术研究：[博士学位论文]．上海：上海交通大学塑性成形工程系，2001

29 陈立周编著．稳健设计．北京：机械工业出版社，2000

30 姚华．基于特征的拉深工艺智能设计系统研究：[博士学位论文]．上海：上海交通大学塑性成形工程系，1999

31 杨红波，陈军，阮雪榆．基于BP神经网络的板料折弯件下料尺寸的计算方法．塑性工程学报，2004，11（3）：69～72

32 张先宏．镁合金热变形过程实验研究与数值模拟：[博士学位论文]．上海：上海交通大学塑性成形工程系，2003

33 中国机械工程学会锻压学会 编．锻压手册：第2卷，冲压．北京：机械工业出版社，1993

34 赵震．面向创新设计理论体系的智能冲压工艺设计KBE技术研究：[博士学位论文]．上海：上海交通大学塑性成形工程系，2002

35 余松敏．折弯机自由折弯时滑块进深的计算和实用公式．锻压技术，1994，19（3）：23～25

36 Guo Y Q，Batoz J L，Naceur H et al. Recent Development on the Analysis and Optimum Design of Sheet Metal Forming Parts Using a Simplified Inverse Approach. Computers and Structures，2000，78：133～148

37 Knopf－Lenoir C，Naceur H，Batoz J L et al. Modeling and Optimum Design of Drawbeads in Sheet Metal Forming. Advanced Technology of Plasticity，Vol. III，Proceedings of the 7th ICTP，Nuremberg，Germany，Sept. 19－24，1999：2055～2060

38 Doege E, Elend L E, Ropers C. Pliable Blank Holder Systems for the Optimization of Process Conditions in Deep Drawing. Advanced Technology of Plasticity, Vol. III, Proceedings of the 7th ICTP, Nuremberg, Germany, Sept. 19 - 24, 1999: 177~182
39 谢汐，孙成智，张卫刚．材料参数及工艺参数对盒形件成形质量的影响．金属成形工艺，2003，21（1）：41~43
40 Geiger, Vollertsen F. The Mechanism of Laser Forming. CIRP Annals, 1993, 42 (1): 301~304
41 Arnet H, Vollertsen F. Extending Laser Bending for the Generation of Convex Shapes. Journal of Engineering Manufacture, 1995, 209: 433~442
42 Vollertsen F. An Analytical Model for Laser Bending. Lasers in Engineering, 1994, 2: 261~276
43 Yan C L, Chan K C, Lee W B. Laser bending of leadframe materials. Journal of Materials Processing Technology, 1998, 82: 117~121
44 Chen G F, Xu X F. Experimental and 3D Finite Element Studies of CW Laser Forming of Thin Stainless Steel Sheets. Journal of Manufacturing Science and Engineering, 2001, 123 (2): 66~73
45 Hu Z, Labudovic M, Wang H, Kovacevic R. Computer Simulation and Experimental Investigation of Sheet Metal Bending Using Laser Beam Scanning. International Journal of Machine Tools & Manufacture, 2001, 41: 589~607
46 Vollertsen F, Holzer S. 3D - thermomechanical Simulation of Laser Forming. Simulation of Materials Processing: Theory, Methods and Applications, Shen & Dawson, 1995: 785~791
47 关振中．激光加工工艺手册．北京：中国计量出版社，1998
48 季忠，王忠雷，焦学健等．板料激光弯曲成形工艺参数优化设计．锻压技术，2002，(6)：38~41
49 谭真，郭广文．工程合金热物性．北京：冶金工业出版社，1994
50 陈建军，李明哲，隋振，苏世忠．基于遗传算法的板材分段成形过渡区的优化设计．吉林工业大学自然科学学报，2001，31（2）：12~16
51 Li M Z, Liu Y H etc. Multi - point Forming: a Flexible Manufacturing Method for a 3 - D Surface sheet. Journal of Manufacturing Processing Technology, 1999, 87: 277~280
52 陈建军，李明哲，隋振等．板材多点成形技术研究综述．塑性工程学报，2001，8（2）：46~48
53 中国机械工程学会，中国模具设计大典编委会．中国模具设计大典：第2卷，轻工模具设计．南昌：江西科学技术出版社，2003
54 周华民．塑料注射成型三维真实感流动保压过程模拟及实验研究：[博士学位论文]，武汉：华中科技大学材料学院，2001
55 申长雨等著．塑料模具计算机辅助工程．郑州：河南科学技术出版社，1998
56 Mok S L, Kwong C K, Lau W S. Review of Research in the Determination of Process Parameters for Plastic Injection Molding. Advances in Polymer Technology, 1999, 18 (3): 225~236
57 Flemings M C. Fluidity of Metals - techniques for Producing Ultra - thin Section Castings. The British Foundryman. 1964, 11 (7): 124~130
58 中国机械工程学会铸造专业学会．铸造手册：第4卷，造型材料．北京：机械工业出版社，1992
59 李先民．冒口的优化设计．四川轻化工学院学报，2002，15（6）：37~40.
60 Zhang J H. Structure Characteristics of Microalloying as - Cast Martensitic Cast Iron. Acta Metallurgica Sinica (English Edition). Series B: Process Metallurgy & Miscellaneous. 1991, 4 (3): 120~128.
61 中国机械工程学会铸造专业学会．铸造手册：第5卷，铸造工艺．北京：机械工业出版社，1994
62 中国机械工程学会铸造专业学会．铸造手册：第1卷，铸铁．北京：机械工业出版社，1993
63 中国机械工程学会铸造专业学会．铸造手册：第2卷，铸钢．北京：机械工业出版社，1991
64 中国机械工程学会铸造专业学会．铸造手册：第3卷，铸造非铁合金．北京：机械工业出版社，1993
65 王培刚，宋强，田学雷．基于AutoCAD2000的铸铝件冒口CAD. 铸造，2003，52（1）：24~27
66 荆涛，刘祥德．三维大型铸钢件铸造工艺的CAD/CAE. 大型铸锻件，1999，(3)：29~33
67 周舰，荆涛．球铁件铸造工艺CAD/CAE系统的研究与应用．铸造技术，2002，23（6）：359~361
68 张雷，姚山等．基于三维参数化造型技术的铸造工艺设计．铸造，1999，(7)：24~27
69 Schmidt D C. Eliminating Modeling 'Trial and Error' with Casting Process Optimization. Mod Cast. 2001, 91 (8): 37~39
70 郭祖培，郭京蕾．微机在铸造熔炼配料中的应用．铸造，2000，149（6)：364~366
71 易传江，阮响明，谭秋等．铝及铝合金熔铸配料决策系统研究．铝加工，2001，24（1）：18~19
72 周永孝，桂卫华，周先德等．基于模糊算法的炉料粒级分布优化系统．计算技术与自动化，2001，20（3）：37~40
73 李智，姚驻斌，张望兴等．模拟退火算法在冶炼Mn - Si合金优化配料中的应用．钢铁研究，2000，(5)：36~39
74 李智，姚驻斌，张望兴等．基于神经网络的混匀配料优化方法．钢铁研究，2000，(4)：10~12
75 孟丽莎．运筹学．北京：兵器工业出版社，1996
76 薛毅．最优化原理与方法．北京：北京工业大学出版社，2001
77 陈祝年主编．焊接设计简明手册．北京：机械工业出版社，1997
78 张丙范主编．焊接最优化基础．天津：天津大学出版社，1990
79 中国机械工程学会焊接学会编．焊接手册：第3册．北京：机械工业出版社，2001
80 吴志生，刘翠荣．晶闸管整流弧焊电源主变压器的优化设计．电焊机，2001，31（9)：30~32，41
81 胡庆彬，黄石生．弧焊逆变器变压器的优化设计．焊接学报，1998，19（2）：98~103
82 隋少华．混料优化设计理论在焊剂研究中的应用．鞍钢技术，1998，(7)：38~40
83 王省离．焊接组合双轴对称工字形截面翼缘改变单跨简支梁优化设计程序．江西水利科技，1995，21（2）：74~88
84 彭兴黔．焊接双轴对称工字型钢梁截面的优化设计．钢结构，2001，16（1）：38~41
85 李午申，巩孝军等．药芯焊丝的计算机辅助设计．机械工程学报，2001，37（7）：44~46
86 方立东，宋健等．焊接工艺参数优化计算机辅助系统的

设计与研究．江苏船舶，2002，19（6）：29～30

87　潘双夏，冯培恩，厉正平．基于成本优化的焊接结构CAD/CAPP一体化模型．浙江大学学报（工学版），1997，31（2）：174～181

88　邹增大主编．焊接材料、工艺及设备手册．北京：化学工业出版社，2002

89　赵国群，王广春，栾贻国，李辉平．基于LOM的产品快速设计与快速制造技术的应用．中国机械工程，2000，11（10）：1128～1130

90　王广春，赵国群．快速成形与快速模具制造技术及其应用．北京：机械工业出版社，2004

91　王运赣．快速成形技术．武汉：华中理工大学出版社，1999

92　马新武，赵国群．体积成形有限元模拟软件CASFORM的开发研究．锻压技术，2003，（1）：9～13

第 5 篇

材料失效分析

主 编 钟群鹏 李鹤林 张 峥

编 写 钟群鹏 付国如 孙永庆
骆红云 张 峥

审 稿 中国材料工程大典编委会

中国机械工程学会
中国材料研究学会
中国材料工程大典编委会

第1章　概　　论

1　失效分析预测预防的地位与作用

从人类认识客观世界的历史长河来说，人的认识是有限的，而客观时间是无限的。失效是人们的主观认识与客观事物相互脱离的结果，失效发生与否是不以人们的主观意志为转移的，因此，失效是绝对的，而安全则是相对的。失效分析是人们认识客观物理本质和规律的逆向思维探索，是对正向思维研究的不可缺少的重要补充，是变失效（失败）为安全（成功）的基本关键，是人们深化客观事物认识的知识源泉。失效分析、改进提高、再失效分析、再提高发展，如此往复循环、螺旋上升、发展飞跃，就是人类科学技术发展历史，乃至于社会发展历史的全过程，因此，广义地说，人类的科学技术发展史、社会发展史就是与广义失效不断作斗争，变失效（失败）为安全（成功）的历史。当今，科学技术是第一生产力，高科技的发展已成为国民经济和国防科技发展的主要关键和依托，而高科技的发展也依赖于高科技发展中的失效分析预测和预防，因此，高科技的发展更需要失效分析预测预防技术的进一步强化，并将失效分析预测预防列为高科技的发展领域之内。总之，失效分析预测预防是从失败入手着眼于成功和发展，是从过去入手着眼于未来和进步的科学技术领域，并且正向失效学这一分支学科方向发展。重视这一分支学科的发展，有意识地运用它已有的成就来分析、解决和攻克相关领域中的失效（失败、故障）问题，是人们走上成功，科技发展少走弯路的捷径之一。

失效分析预测预防的地位和作用可以从经济管理、社会生活、工程技术和科技进步等四个方面进行分析。

在经济管理上，失效分析是防止失效事故的再发生、减少经济损失或人员伤亡的必由之路；是对国内外经济纠纷进行仲裁或判决、索赔的科学技术依据；是创建优质名牌产品、提高产品的质量和更新换代的重要途径；是为各级领导提供反馈和进行技术经济规划与决策的重要参考依据。

在社会生活上，失效分析是促进安全生产、保护生产力的有效武器；是贯彻“安全第一”、“预防为主”方针的一项重要工作；是保证社会安定、经济持续发展的重要保证；是社会主义市场经济体制优越性的重要体现。

在工程技术上，失效分析是机电产品维修工作的技术基础和前提条件；是可靠性工程的基础技术工作之一；是发展和完善安全工程技术的重要实践基础；是修改和完善产品和技术标准的科学实践依据。

在科技进步上，失效分析是认识客观世界事物本质的重要知识源泉；是发展新学科、新理论、新技术、新材料、新工艺、新方法的重要窗口和“产婆”；是发展科学技术，包括发展高科技的不可避免的机遇和挑战；是从失败入手着眼于成功和发展的科技领域；是从过去入手着眼于未来和进步的科技领域；可以认为是科技是第一生产力的最活跃的因素之一。

2　与安全和失效有关的术语定义

2.1　安全的定义、内涵和外延

失效与安全术语的概念和定义十分密切，因此，先讨论有关安全的定义、内涵和外延问题。

无危则安，无损则全。“安全”顾名思义，指没有危险，不受威胁，不出事故，没有受伤，完整无损，平安健康。安全的反义词是灾害。灾害是对人类生命、财产和生存条件造成危害性后果的各种变异现象的总称。

从科学的含义上看，“安全”可以认为是一种状态，是指满足人和物不受损伤、身心健康和完整完满的一种环境、物态和状态；也有人认为，“安全”是一种能力，是指人类对自身利益——包括生命、健康、财产、资源、生存空间（领土、领海、领空）、信息、无形资产、商业机会、传统、文化、社会结构、运行机制和秩序等的捍卫、维护和控制的能力。将上述“物态论”和“能力论”相结合，安全是通过“能力”达到的一种“状态”。

安全的内涵和外延可以有狭义的（传统的）和广义的（非传统的）安全之分：狭义的安全（又可称为传统的安全）包括国家安全（即国防安全）、社会安全和公共安全。国家安全是指在国家、民族、领土层面上的安全，侧重于在国家和民族之间的冲突和对抗中捍卫国家利益问题；社会安全是指社会秩序、运行机制层面上的安全，侧重于社会犯罪、破坏和威胁中维护社会利益问题；公共安全是指在生产经营层面上的安全，侧重于对安全生产事故的控制问题。广义的（非传统的）安全是人类进入21世纪后，由于社会、经济及科学技术的发展引发的新的安全问题，包括环境安全、技术安全、城市安全、经济安全等。其中环境安全又可分为生态、人口、资源（能源、土地、水资源、矿产）、气候等相关的安全问题，环境安全的本质是由于经济高速发展引发的人与自然不能和谐发展的问题；技术安全是指新技术、新产品的出现，包括生物技术、基因技术、新的化工产品及大型工程给人和自然造成的风险和威胁；城市安全是由于城市规模和生产的集约化发展带来的突出的安全问题；而经济安全则是由于世界经济的一体化，使得社会经济的关联性、依赖型和“脆弱”性大大的增加带来的附加的安全问题，例如动力安全（核安全）、信息安全、金融安全等。

鉴于安全的内涵及其外延不断的深化和扩大，安全的概念成为一个动态的、发展的，带有全局性、关键性和战略性的问题，安全与社会经济发展之间的关系越来越紧密、依赖和互为制约，因此与安全有关的安全文化、安全科学、安全技术、安全管理、安全经济、安全伦理等也将应运而生、蓬勃发展，其重要性也日益突出，安全不再是人类被动的追求目标状态，而且将对经济社会、政治发展起反转的促进作用。

那么，安全生产工作就是对生产经营活动中的生产事故（失效）、事故隐患和事故风险进行分析、识别、评价和控制全过程的监察、监督、管理和科技活动。

而失效分析预测预防是对生产经营活动中的安全事故、事故隐患和事故风险进行分析诊断、预测预防、研究采取补救和根治措施的技术活动和管理活动。

因此，可以认为，失效分析预测预防＋监察、监督——→安全生产工作。

2.2　与失效有关的术语定义

1）失效　机电产品丧失功能的现象称为失效。

2）失效分析预测预防　分析诊断失效的模式、原因和机理，研究采取补救预测和预防措施的技术活动和管理活动。

3）失效学　研究机电产品失效的诊断、预测和预防理论、技术和方法的交叉综合的分支学科。失效学与相关学科

的边界还不够明确，它是一个发展中的新兴学科。

4）风险　风险是失效的可能性与失效后果的乘积。风险评估就是对系统发生失效的危险性进行定性和定量的分析。

5）失效件和废品　失效件是指进入商品流通领域后发生的故障，而废品则是指进入商品流通领域前发生的质量问题。废品分析采用的方法常与失效分析方法一致。

6）失效和可靠　失效是可靠的反义词。机电产品的可靠度 $R(t)$ 是指时间 t 内还能满足规定功能产品的比率，即 $n(t)/n(0)$。累积失效概率 $F(t)$ 就是时间 t 内的不可靠度，即 $F(t)=1-R(t)=[n(0)-n(t)]/n(0)$。

7）失效和事故　失效强调过程，而事故则突出后果。如由于涡轮叶片的疲劳断裂失效，导致某型号的×等事故。

8）失效分析和状态诊断　失效分析是指事后的分析，而状态诊断是针对可能的主要失效模式、原因和机理方面事先的，即在线、适时、动态的诊断。

9）失效分析和安全评定　失效分析是指事故后的失效模式、原因和机理诊断，而安全评定是指事故前，按"合于使用"原则的安全与否的评价。

3　失效的分类

失效可以按以下三种方法分类，即从技术观点、质量管理的观点和经济法的观点进行分类。

按技术观点进行分类便于对失效进行机理研究、分析诊断和采取预防对策；按质量管理的观点进行分类便于管理和反馈；按经济法的观点进行分类则便于事后处理。本书将主要讨论从技术观点的分类。

（1）按失效分析技术观点分类

从失效分析的技术观点进行分类主要是按失效模式和失效机理分类。失效模式是指失效的外在宏观表现形式和规律。失效机理则是引起失效的微观的物理化学变化过程和本质。按失效模式和失效机理相结合对失效进行分类就是宏观与微观相结合、由表及里地揭示失效的物理本质和过程，因而它是一种研究的重要分类方法。

机械失效分类见表5.1-1。在表5.1-1中，左边为失效模式或失效机理，右边为断口（表面）形态，左右之间的实线为通常情况下的对应关系，虚线为特定情况下的对应关系。需要指出的是，单一模式或单一机理的失效较少见。大多数的失效是多因素、多种机理及复杂模式的"复合型失效"。表5.1-1右边各种机理之间的联线表示它们之间的交互作用。关于各种失效模式和失效机理的定义和基本概念参考有关文献。

在按"失效分析的技术观点"分类中，还有人按失效零件或部件的类型和按引起失效的工艺环节分类的。前者便于应用于机械零、部件的可靠性设计和评估，后者则有利于机械零、部件的质量控制。

（2）按质量管理和可靠性工程观点分类

质量管理乃是人们为了保证、改进和提高质量而从事的计划、组织、协调、审查、检验等一系列控制活动。失效则是产品质量控制网"三性"（健全性、严密性和散布性）发生偏差（弱网、漏网和无网）的反映，失效分析就是对质量控制网"三性"的实地检验（图5.1-1）。因此，失效分析是质量管理的重要组成部分之一。可靠性工程是运用系统工程的思路和方法，权衡经济利弊，把设备（或系统）的失效率降低到可接受的程度，从而进行合理的设计（可靠性设计）和管理的一门技术。因此，可靠性工程实质上就是预测、预防和控制失效的技术工作和管理活动，而失效分析则是可靠性工程的技术基础。失效从质量管理和可靠性工程的观点分类就是要从失效发展过程和速度、失效的工程含义（即失效的整体性、可修复性和相关性）来分类。具体分类如下：

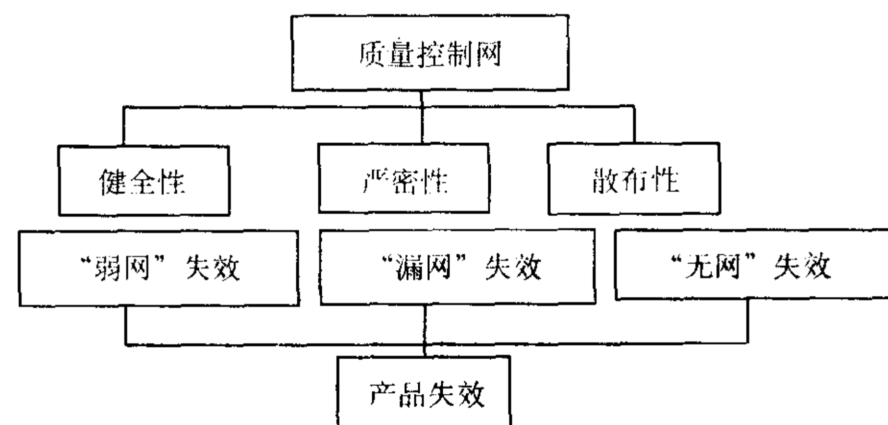


图5.1-1　产品失效与质量控制网"三性"关系

1）按产品失效的发展过程分类　如果我们以失效率——单位时间内发生失效的比率——来描述产品失效的发展过程，那么，在不进行预防性维修的情况下，设备、元件的失效率 λ 与其工作（使用）时间 t 之间关系如图5.1-2所示的典型失效率曲线。因为这种曲线的形状与浴盆相似，故称为"浴盆曲线"。

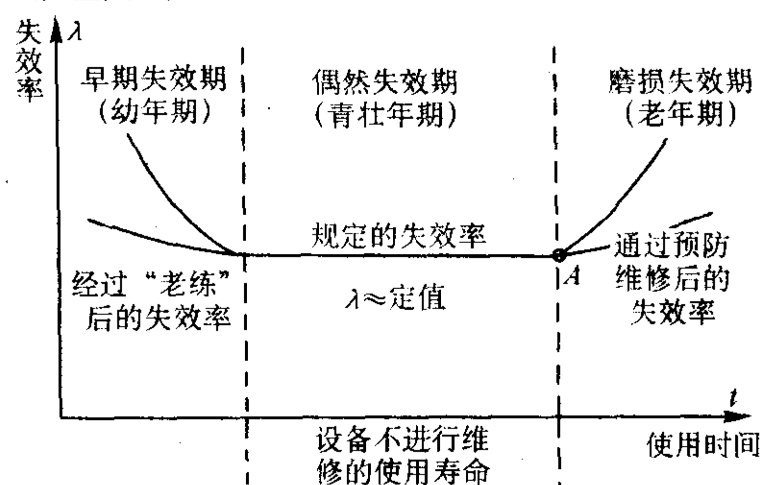


图5.1-2　设备不进行预防维修时的典型失效率曲线——浴盆曲线

按照"浴盆曲线"的形状，即按照产品失效的发展过程，可以将整个失效过程分为三个时期。

① 早期失效期　在产品的使用初期，容易暴露由于设计和制造上的缺陷而导致的失效，因此产品的早期失效率很高。随着使用时间的延长，失效率则很快下降。产品的早期失效期相当于人的"幼年期"。如果在产品出厂之前，进行旨在剔除这类缺陷的"老练"过程，即进行可靠性试验，那么在产品以后的使用时，从一开始便可使失效率大体保持恒定值。

② 偶然失效期　在理想的情况下，产品在发生磨损或老化以前，应是无"失效"的，但是由于环境的偶然变化、操作时的人为偶然差错或者由于管理不善造成的"潜在缺陷"，仍有产品的偶然失效。产品的偶然失效率是随机分布的、很低的和基本上是恒定的，故又称为随机的失效期。偶然失效期相当于人的"青壮年期"，这一时期是产品的最佳工作时期。偶然失效率的倒数即为无失效的平均时间。

③ 磨损失效期　经过偶然失效期后，设备中的元件已到了寿命终止期，于是失效率开始急剧增加，这标志产品已进入"老年期"，这时的失效叫做磨损失效，又称为耗损失效。如果在进入磨损失效期之前，进行必要的预防维修，它的失效率仍可保持在偶然失效率附近，从而延长产品的偶然失效期。

产品失效按其发展过程分类对可靠性工程来说是十分有用的。

2）按产品失效发生的速度分类：①突然失效；②渐进失效；③间歇失效。

3）按产品失效的工程含义分类　按其整体性分为：系统失效和部件失效。按其修复的可能性分为：暂时失效和永久失效。按其相关性分为：独立失效和从属失效；或关连失

表 5.1-1 失效按技术观点分类

失效模式			失效机理
机械失效	断裂失效	韧性断裂	韧性断裂机理
	脆性断裂	低温断裂	低温断裂机理
		辐射脆化	辐射损伤机理
		氢损伤	氢损伤机理
		应力腐蚀	应力腐蚀机理
		液（固）金属脆化	液（固）态金属脆化机理
		液体浸蚀损伤	液体浸蚀机理
		高温应力断裂	蠕变断裂机理
		疲劳断裂	疲劳机理
	非断裂失效 — 磨损	磨粒磨损	磨粒磨损机理
		黏着磨损	黏着磨损机理
	非断裂失效 — 腐蚀	氧化	氧化机理
		电化学腐蚀	电化学腐蚀机理
	非断裂失效 — 变形	弹性	
		塑性	塑变机理

复合失效机理：低周疲劳；应力腐蚀疲劳；高温疲劳，蠕变/疲劳交互作用；疲劳磨损 微振磨损；烧蚀 热蚀；腐蚀疲劳；电腐蚀；腐蚀磨损；冲蚀磨损

断口表面形态：韧窝延伸带；解理（准解理）；沿晶；疲劳条带；磨屑、磨坑；金属（材料）转移；氧化物或硅酸盐；腐蚀减薄及其产物；弹性变形；尺寸、形状变形

复合断口形态：沿晶韧窝；脆性疲劳条带；沿晶疲劳条带；韧性疲劳条带 大应力疲劳特征

效和非关联失效等。

(3) 按经济法观点分类

机械产品的失效会造成一定的经济损失甚至人员的伤亡，往往会引起赔偿和责任的诉讼。失效按“经济法”的观点分类是为了分清和判处失效的法律责任和经济责任。这种分类方法在处理索赔的失效事件时尤为重要。

1) 按失效的责任分类：①产品本质缺陷失效；②误用失效；③正常的磨损（或耗损）失效；④外界影响失效。

2) 按失效的后果分类：①恶性失效；②致命失效（灾难性失效）；③退化失效。

3) 按失效的程度分类：①完全失效；②部分失效。

上述分类中涉及的名词定义和概念详见有关文献。

4 失效分析预测预防的特点和属性

4.1 失效、失效分析的特点和属性

重大的失效事故及其分析工作有如下特点：

1) 发生的突然性和过程的难救援性；

2) 损失的严重性和遇难的悲惨性；

3) 模式的多样性和原因的复杂性；

4) 分析的困难性和研究的探索性；

5) 思路的新颖性和技术的科学性；

6) 影响的广泛性和成果的重要性。

因此失效分析预测预防的属性是：

1) 技术范围有多学科交叉综合的性质——要求参加人员应具有宽广的知识面，较强的组织能力和良好的团队协作精神；

2) 分析过程有探索侦察的性质——要求参加人员应具有科学的实事求是、客观公正的学风和孜孜不倦的钻研探索精神；

3) 最终结果首先要求诊断机电产品失效的模式（病症、性质——他杀、自杀）、原因（病因）和机理（病理），因此失效分析人员相当于机电产品的“法医”。

4.2 失效分析预测预防与相关学科、技术之间的关系

由于机械失效学是一门交叉、边缘、综合的新兴学科，因此，它与多种学科和技术有关。

基础学科与机械失效学相结合，形成不少边缘学科，例如，机械失效学与物理学交叉诞生了失效物理学；与力学的交叉形成了失效力学和损伤力学；数学与机械失效学相结合，促成了可靠性数学的问世。

机械失效学与许多应用学科、技术有密切的联系。“机械”是失效分析的对象，因此失效分析与机械学的专门知识有关；“材料”是失效的载体，这样失效分析就自然地涉及材料科学和工程领域的各种知识；“环境”是失效发生的条件，所以失效分析就与环境科学知识有关；“检测”是失效分析中信息的获取的重要途径和手段，显然失效分析就离不开宏观和微观的检测有关的知识；“分析”是失效分析的核心，因此失效分析必将涉及逻辑学和数理统计有关的专门知识；此外，失效分析不仅是一种“技术活动”，而且还是一种“管理活动”，因此，失效分析过程和它的成果反馈，又与管理科学的专门知识有关。

可以认为，机械失效学是在其他学科基础上，在20世纪70年代形成和发展起来的新的应用性学科，它的发展又对其他学科有反馈促进作用。可以这样说，近代材料科学和工程力学对断裂、腐蚀、磨损及其复合型的失效模式和失效机理研究，为失效学奠定了理论基础。现代的检测仪器仪表学、断口分析术，为失效学奠定了技术基础。数量统计、模糊数学、可靠性工程和电子计算机科学的广泛应用为失效学提供了新的方法途径，以上三个方面的融合贯通，使机械失效学逐渐形成相对独立的，十分活跃的学科。

4.3 失效分析预测预防应注意的事项和对人员素质的要求

(1) 失效分析预测预防应注意的事项

由于失效分析过程具有探索侦察的性质，因此在失效分析的过程中，应特别注意下列事项。

1) 深入进行调查研究，决不要漏掉蛛丝马迹 进行调查研究是失效分析的第一个步骤，必须贯彻到失效分析各个环节和整个失效分析工作的始终。调查研究要亲到现场，深入进行，决不要漏掉蛛丝马迹。调查研究要有提纲、口问手记，随时分析整理。调查研究是失效分析的重要原则和基础。

2) 认真制定失效分析的程序，避免盲目性和片面性 失效分析程序是整个失效分析工作“纲领性”文件，它是整个失效分析成败的关键。“程序”的制定要建筑在深入调查研究的基础上，要经过周密地思考，切不可草草从事。“程序”既要重点突出，避免盲目性，又要考虑到失效原因各种可能，克服片面性。“程序”使整个失效分析工作有章可循和有条不紊地进行，失效分析人员既不能轻易更改“程序”，但也不能一成不变，应根据失效分析的进展和新的情况，作必要的调整。

3) 充分注意失效分析的复杂性和综合性，避免技术上的局限性 失效分析是涉及多种学科的知识和多种测试技术的一门综合性的边缘学科，因此，失效分析人员在专业知识上和经验上的局限性，往往会影响失效分析结果的正确性。为此，不仅要在失效分析小组的组成人员上考虑各有关专业人员的结合，作为每个失效分析人员则要耐心地、虚心地听取不同角度的分析意见，要努力扩大自己的知识面。一般情况下，冶金出身的失效分析工作者，要更多的注意失效零件或部件的结构特点、受力状态、应力分析及其可靠性分析；对于机械和力学出身的失效分析成员则不应忽视失效的机理、材料和工艺因素。只有各个学科的知识、各种测试方法和技术、各类人员的结合和密切合作，才能找到合乎实际的失效原因和补救、预防措施。

4) 尊重客观事实，坚持实事求是，排除人为的干扰 失效是人们主观认识和客观事物不一致的结果，失效后面有至今尚未认识的真理，因此失效分析一定要尊重客观事实。失效分析可能会牵连到某些人或单位的利益，这些人往往会利用自己的地位和影响来改变失效分析的进程和结论，因此失效分析工作者必须要有科学的求实精神，勇于坚持真理，排除各种人为的干扰，这是失效分析能否得出正确结论的关键之一。

(2) 失效分析预测预防对人员素质的要求

鉴于失效分析工作的特点和属性，失效分析工作人员必须具有如下基本的政治和业务的素质。

1) 优良的品德 具有勇挑重担的革命进取精神和牺牲精神；要有求实精神；要有谦虚谨慎的求知欲望和好学精神。

2) 正确的思维方法 要有科学的方法论，例如思维、归纳、演绎、分析、综合、抽象、类比、想象、假说、检验、论证等方面的基本素养。

3) 广泛的知识素养 失效分析工作人员应当在一、两个学科中打好扎实的理论和实践的基础，并不断地学习，吸取新的知识，逐渐扩大自己的知识面，成为“大杂家”。

4) 较强的工作能力 失效分析工作者要有基本的分析能力、实验能力、计算能力、外语能力、总结书写能力、表

达能力和组织工作的能力。

4.4 失效分析预测预防的技术工作

广义地说，失效分析的工作内容应包括失效分析的业务工作（即“门诊”工作）、失效分析的研究工作和失效分析的管理及技术反馈工作。下面就分别给予扼要地讨论。

（1）失效分析的预测预防的业务工作

总的来说，失效分析的业务工作有两个方面，一方面是产品的失效分析，另一方面是产品的安全度评定和剩余寿命的预测。前者是失效事后的分析，而后者是失效事前的分析。

根据失效分析的定义，一个失效事件分析的全过程一般包括侦测（detection）、诊断（diagnosis）和事后处理（prognosis）三个阶段。即利用各种“侦测”手段，调查、侦查、测试和记录有关失效的现场、参数和信息；通过“诊断”鉴别和确定产品失效的模式、过程、原因、影响因素和机理；经过“事后处理”采取补救措施（对服役件）、预防措施（对新生产的产品）和其他技术的、管理的反馈活动，以达到预防、提高和开发的目的。侦测、诊断和事后处理是失效分析工作的三个要素。

产品失效分析的重点无疑是分析产品的早期失效事件、突发性失效事件以及致使的失效事件，因为这些失效事件的分析事关重大或关系到全局。

失效分析的深度应依其分析的目的和要求不同而异。作为法律依据的失效诊断主要的任务是判定失效的程度、后果和责任方面；作为整顿质量管理的线索和获得可靠性工程的数据，一般希望确定失效发生的阶段、对策及其他的工程属性；而为了研究和掌握机械产品失效的规律和它的内在本质，这种失效分析则必须着眼于失效模式、机理、影响因素和控制参量的定性乃至定量的研究。因此，失效分析工作者的任务是根据失效分析的不同目的和要求，作出确切而恰当的诊断及对策。

按照产品的发展的阶段——试制阶段、试生产阶段和定型生产阶段——不同，其失效分析的方面和内容也各不相同。在试制阶段的产品，失效的原因常常与设计因素有关，因此在这一阶段的失效分析应特别强调与设计人员的密切配合，这样不仅有利于失效原因的正确诊断，而且有利于分析结果的迅速反馈；在试生产阶段的产品，其失效原因多半与工艺因素有关，因此在这一阶段的失效分析应注意与工艺人员相结合；而在定型生产阶段的产品，其失效的原因一般则与管理因素有关，因此在这一阶段的失效分析着重考查产品的质量控制网的健全、严密性和散布度，并应将广大的用户和社会上的修理行业纳入到失效分析体系中来。产品的安全度或产品的可靠度 $R(t)$ 是产品在规定的条件下和规定的时间内满意地完成规定功能的概率，它是相对于产品失效率 $F(t)$ 的概念而言的，即 $R(T)=1-F(t)$。因此，产品安全度的评定就是要定量地计算产品工作到某一时刻 t 尚能满足规定功能的概率，或者认为是估算产品工作到某一时刻 t，在尚能满足功能的产品中，在此后的单位时间内发生失效的概率。而产品的剩余寿命的预测实际上是估算产品无故障（失效）的平均工作时间 $\bar{t}$，对偶然失效期（或称为随机失效期）来说，平均工作时间 $\bar{t}$ 即为失效率 λ 的倒数，即 $\bar{t}=1/\lambda$。因此，产品剩余寿命的预测实质上是偶然失效率的估算问题。从这里可以看出，产品的安全度评定和剩余寿命的预测都是失效分析重要的业务工作之一。

产品的安全评定和剩余寿命的预测工作重点是重大的机械产品和带缺陷产品。

进行产品的安全度评定和剩余寿命的预测工作的关键是通过可靠性试验获得产品发生失效的概率有多大，应如何预防等方面的资料和数据。这些资料和数据要经过现场的失效分析来修正。只有这样才能使产品的安全度评定和剩余寿命的预测结果接近于客观实际。

当然产品安全度评定和剩余寿命预测也可以根据人们积累的经验数据或公式进行，但是这些数据或公式必须与产品的实际失效情况相符合。

关于可靠性工程的基本概念和有关的名词术语详见有关文献。

（2）失效分析预测预防的研究工作

失效分析的研究工作主要是指失效分析物理的研究，失效机理的研究，失效诊断的研究和失效预防工程技术方面的研究等四个方面的工作。下面就分别给予简要的介绍。

1）失效诊断的研究工作　失效诊断是分析失效原因的思维学和方法论。具体的说，失效的诊断研究一般应包括失效诊断依据的研究、失效诊断思路的研究和失效诊断技术与方法的研究三个方面。

从总的来说，失效诊断的依据可以是产品的性能参数与工况、可靠度的数理统计资料和失效零、部件的残骸上获得的信息，或者是上述三种的综合判据。依据产品的性能参数与工况进行失效的诊断和其工况分析及状态监测密切相关，因此它既是失效诊断技术又是失效的预防措施之一；依据产品的可靠度统计资料进行失效诊断，实质上就是进行产品的可靠度计算、评定和剩余寿命的估算工作；而依据失效残骸分析进行失效诊断则是一种最常用的失效诊断方法。失效残骸诊断术是通过残骸的断口分析、裂纹分析、应力分析、材质分析、工艺分析、工作环境分析及它们的综合分析对失效的原因做出正确的判断，并对其补救和预防措施做出决策的。失效残骸诊断术是失效诊断研究工作中一个十分重要的课题。

失效诊断思路的研究是侧重于研究失效诊断的逻辑思维或推理判断的方法和程序，它一般应该包括逻辑推理、判断、模糊数学的判别和电子计算机在失效诊断方面广用等，其中失效形式及影响分析法——FMEA 法，失效模式及影响危险度分析法——FMECA 法、失效树法——FTA 法、现象树法——ETA 法、特性因素法等应用较多。失效诊断思路研究的目的和任务就是要完善、发展和推广应用上述各种逻辑诊断思路。

关于失效诊断技术的研究涉及应该采用什么样的仪器设备和技术方法，例如为了进行失效残骸的材质分析，可以采用许多种物理、化学的测试方法——无损探伤法、金相分析法、断口分析法、化学成分分析法、机械性能测试等，在各类分析方法又可分若干种具体方法。失效诊断技术的研究就要在了解各类方法的基本原理、优缺点和应用范围的基础上，指导失效分析工作者选用正确的方法获得必要而充分的信息和资料、数据作为失效诊断的有力依据。可见，失效诊断技术的研究不应把注意力放在各种方法的理论或仪器的具体构造细节上去，而应注重各种方法的优缺点、应用范围以及用于同一目的、不同方法之间横向比较和正确选用上。

2）失效预防的研究工作　失效预防的研究工作是为了探讨失效的补救措施、预防方法和管理途径。失效分析成果的反馈和失效预防，以防止同类失效事故的重复发生是失效分析的重要目的之一。失效预防的研究工作范围很广，它应包括机械系统或设备的性能参数的分析和诊断技术工况的监测方法、技术、装置和系统的开发和采用，产品的安全度评定和剩余寿命的估算，失效预防工程方法（包括表面损伤的预防方法，断裂失效临界状态的评定和预测方法——失效评定图、断裂控制图、变形机制图、蠕变机制图和疲劳机制图等）的研究开发和应用推广，关于机械产品失效维修的原则、原理、技术、方法及其应用方面的研究等。可以看出失

效预防方法的研究工作是艰巨的、十分重要的却大有可为的。

3）失效机理的研究工作 失效机理是研究失效的物理、化学原因、失效过程及其影响因素等。如果认为失效物理的研究内容主要是失效模式（模型）及其应用的话，那么失效机理的研究则主要集中在失效的本质（物理、化学原因）和过程。失效机理和失效模式不同，以人的生病来打比喻，失效机理相当于病理，而失效模式则相当于病症。具体地说，失效机理主要研究各种失效方式，研究材料成分、组织结构和性能等内部因素对失效过程的影响。

失效机理一般可按表5.1-1进行分类。值得指出的是，失效机理的研究不仅需要研究单一的失效机理，而且要特别加强各种复合机理的研究，例如：高温疲劳机理、低周疲劳机理、应力腐蚀疲劳机理、腐蚀疲劳机理、疲劳磨损机理、蠕变疲劳交互作用机理、微振磨损机理、冲蚀磨损机理、烧蚀机理、电腐蚀机理、热蚀机理等等的研究。

十分明显，失效机理的研究是失效分析的又一重要理论基础。只有对失效机理，确切来讲，是深入研究和了解复合的失效机理的基础上，才能进行正确的失效分析工作；反过来，大量地、深入地进行富有成效的实际的失效分析工作，又将更进一步加深对错综复杂的复合失效机理的认识。失效机理研究是基础，失效分析工作是实践，它们是促进整个失效分析工作以及失效理论向前发展的两个相辅相承的方面。

4）失效物理的研究工作 所谓失效物理（或称为可靠性物理）就是从原理上，即从原子和分子的角度出发，来解释元件、材料失效的现象。失效物理学的基础是数理统计方法、可靠性工程和材料科学工程学。失效物理的基本研究内容则是失效的物理模型定性及定量的描述方法、失效物理模式的识别及其应用。基本的失效物理模型有：界限模型与耐久模型、应力-强度模型、反应论模型、失效率模型、最弱环模型与串联模型、“绳子”模型、比例效应模型、退化模型或损伤累积模型等。失效物理不仅要对上述基础及模型有所涉及，而且还涉及到材料失效的热力学问题。由此可见，失效物理的内容十分丰富和重要，对失效物理的研究是失效理论的基础研究之一。无疑，对失效物理的深入研究和广泛应用，必将加强失效分析的广度和深度。失效物理模型的概念详见有关文献。

可以认为，失效分析的研究工作内容就是机械失效学的形成、完善和发展以及广泛应用。

（3）失效分析的预测预防的管理和技术的反馈工作

失效分析的管理和技术反馈应该包括全国性、部门性、地区性的失效分析和管理机构的建立；失效分析的技术指导性文件、规程和标准的颁布实行；失效事件的分析工作的组织和管理；失效研究工作的组织和开展；失效及可靠性数据库和技术反馈系统的建立和运转；各级失效分析人员的培训和提高等。值得指出的是，从失效分析的定义中就已经把“管理活动”当作失效分析不可缺少的内容和环节之一，从以上的讨论中可进一步地看出，失效分析的管理和技术反馈内容和范围十分广泛和重要，因此，应给予足够重视。

5 失效分析预测预防的步骤和程序

失效分析的实施步骤和程序旨在保证失效分析顺利有效地进行，因此其细节的制定应根据失效事件的具体情况（失效设备的类型及其失效的严重性等）、失效分析的目的与要求（为机理研究、技术改进、还是为法律仲裁）、以及有关合同或法规的规定来决定。下面推荐的一般通用的失效分析实施步骤和程序，原则上可供参考和引用。

图5.1-3示出所推荐的失效分析实施步骤和程序的方框图。

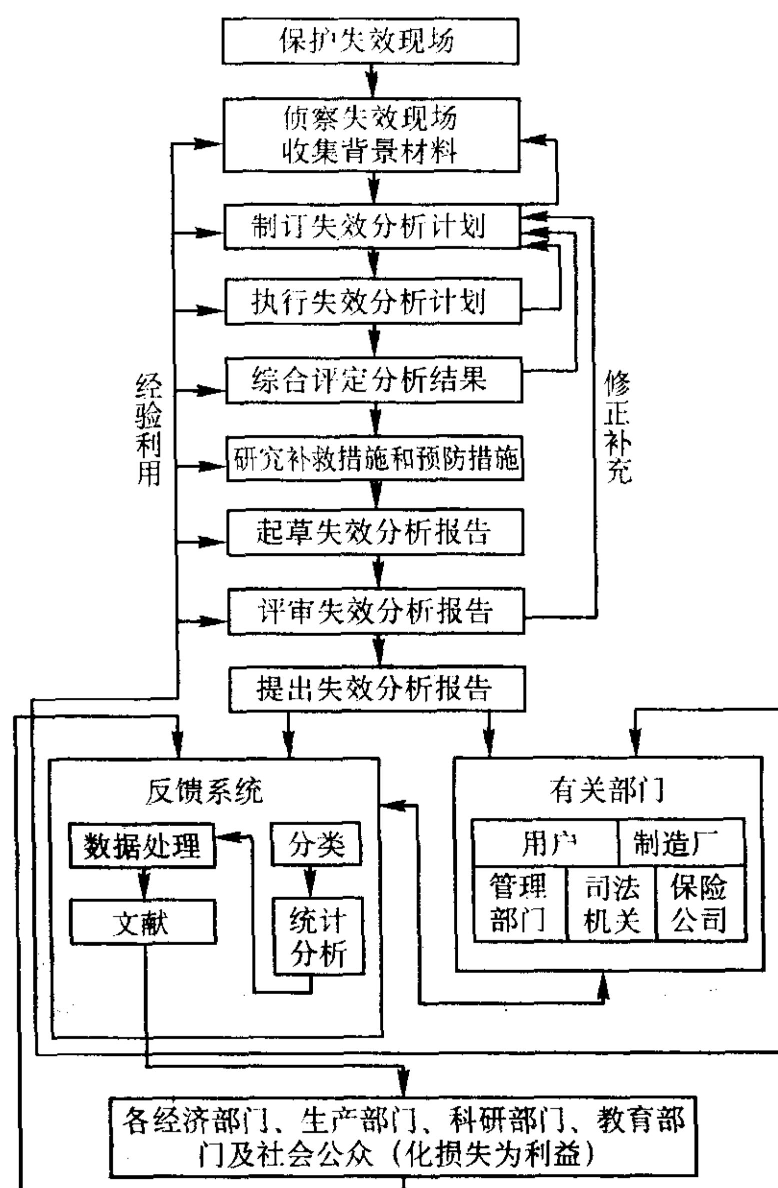


图5.1-3 失效分析实施步骤和程序

（1）保护失效现场

保护失效现场的一切证据维持原状、完整无缺和真实不伪，是供失效分析得以顺利有效地进行的先决条件。失效现场的保护范围视机械设备的类型及其失效发生的范围而定。

（2）侦察失效现场和收集背景材料

失效现场侦察应由授权的失效分析人员执行：并授权收集一切有关的背景材料。

失效现场侦察可用摄影、录像、录音、绘图及文字描述等方式进行记录。

失效现场侦察所应注意观察和记录的项目常有：

1）失效部件及碎片的尺寸大小、形状和散落方位；

2）失效部件周围散落的金属屑和粉末、氧化皮和粉末、润滑残留物及一切可疑的杂物和痕迹；

3）失效部件和碎片的变形、裂纹、断口、腐蚀、磨损的外观、位置和起始点，表面的材料特征：如烧伤色泽、附着物、氧化物、腐蚀生成物等；

4）失效设备或部件的结构和制造特征；

5）环境条件 如失效设备的周围景物、环境温度、湿度、大气、水质；

6）听取操作人员及佐证人介绍事故发生时情况（录音记录）。

所应收集的背景材料通常有：

1）失效设备的类型、制造厂名、制造日期、出厂批号，用户、安装地点、投入运行日期、操作人员、维修人员、运行记录、维修记录、操作规程、安全规程；

2）该设备的设计计算书及图样、材料检验记录、制造工艺记录、质量控制记录、验收记录、质量保证合同及其技术文件，使用说明书；

3）有关的标准、法规、及其他参考文献。

（3）制定失效分析计划

只有极少数的情况下，通过现场和背景材料的分析就能得出失效原因的结论。大多数失效案例都需根据现场侦察和

背景材料的综合分析结果来制定“失效分析计划”，确定进一步分析试验的目的、内容、方法和实施方式。

失效分析计划由授权的分析人员制定，根据具体情况或要求，可由有关方面的代表参加讨论。

对各项试验方案应考虑其必要性、有效性和经济性。一般宜先从简单的试验方法入手，如有必要时才进一步使用费用高的和较复杂的试验方法。如确有必要进行失效模拟试验，其设计应尽可能模拟真实的工况条件，使之具有说服力。

从失效部件上和残留物上制取试件或样品，对于失效分析的成败具有十分重要意义，务必要周密计划取试样的位置、尺寸、数量和取样方法。应当特别强调，失效部件和残留物上具有说服力的位置和尺寸是十分有限的，一旦取样失误，就无法复原而完全丧失说服力，致使整个失效分析计划归于失败，造成不可挽救的后果。

失效分析计划要留有余地，以便在个别试验中发现意外现象时，为了适应新的情况，可中途改变某些方法，或做补充试验。

(4) 执行失效分析计划

失效分析的各项试验应严格遵照计划执行，要有详细记录，随时分析试验结果。

失效分析的试验一般具有几个不同于一般科研试验的特点，应特别予以重视。

1) 一般都要求在很短的时间内取得试验结果，因此既要保证按时完成，又要防止在匆忙中发生疏忽和差错。

2) 许多失效分析工作涉及法律问题，为此，各项试验工作应建立严格的责任制度。试验人员在试验记录和报告上签名。

3) 试件、样品都要直接取自失效实物。一般不能用其他来源的试件样品代替。

4) 失效分析是人们进一步认识未知客观世界的一种科研活动，试验人员切不可在思想上存在先入为主的概念，错误认为失效分析的试验只不过是已知条件的复验，以致放松对试验过程中出现新现象的观察，实际上，失效分析往往含有新发现和技术突破，试验人员更应注意观察这种试验的全过程。

(5) 综合评定分析结果

授权的失效分析人员，要经过充分的讨论，对现场发现、背景材料及各项试验结果做综合分析，确定失效的过程和原因，做出分析结论。

综合分析，特别是在复杂的失效案例情况下，可用故障树（Fault Tree）或找其他形式的逻辑图分析方法。

在大多数情况下，失效原因可能有多种。应努力分清主要原因和次要原因。

综合分析讨论会有详细的发言记录和代表共同意见的会议纪要，由与会人员签名，存入失效案例档案。

(6) 研究补救措施和预防措施

失效分析的目的不仅限于弄清失效原因，更重要的还在于研究提出有效的补救措施和预防措施。从大量同类和相似失效案例分析积累的丰富经验有利于这类措施的研究。

补救措施和预防措施可能涉及设备的设计结构、制造技术、材料技术、运行技术、修补技术以及质量管理的改进，乃至涉及技术规范、标准和法规的修订建议。这类研究工作量往往很大，除个别简单情况可由承担失效分析的人员进行外，一般由失效分析人员提出问题或补救方案，由负责单位责成有关专业部门或单位进行专题研究，提出研究报告，做为改进设备的依据。

(7) 起草失效分析报告

失效分析报告一般可不规定统一的格式，但行文要简练，条目要分明，内容一般应包括下列项目：

1) 题目；

2) 任务来源：包括任务下达者及下达日期，任务内容简述，要求的分析目的；

3) 各项试验过程及结果；

4) 分析结果——失效原因；

5) 补救措施和预防措施或建议；

6) 附件（原始记录、图片等）；

7) 失效分析人员签名及日期。

对于大宗同类产品的失效分析，宜规定一定的报告形式。以便于事后的统计分析工作和计算机辅助失效分析。

(8) 评审失效分析报告

失效分析评审会的组织形式及其参加人员可由有关方面协商决定，一般宜由失效分析工作人员、失效设备的制造厂商代表、用户代表、管理部门代表、司法部门代表和聘请的其他专家组成。各方面代表应本着尊重科学、尊重事实和法律的态度履行其评审职责，不得对失效分析人员以任何形式施加不正当的压力和影响。失效分析人员的客观公正立场应受维护和尊重。

(9) 提出失效分析报告

失效分析报告通过评审后，按评审决议修改并制定成报告正式文本，内容项目除上述第（7）条所示外，还宜增加下列三项：

1) 评审意见：包括评审人员签名及日期；

2) 呈送及抄送单位：包括抄送“反馈系统”；

3) 密级。

(10) 反馈失效分析成果

反馈系统是失效分析成果的管理系统，目的在于充分利用失效分析所获得的宝贵技术信息、推动技术革新、促进科学进步和提高产品质量。

失效分析的反馈系统可采取多种组织形式。例如可与企业的技术开发和情报部门结合。可与国家的质量管理部门、可靠性研究中心、数据中心及数据交换网相结合，把输入的大量失效分析报告和来自数据交换网的其他信息，经过“分类”、“统计分析”、“数据处理”，制成各种形式的文献，例如快报、数据手册、指导性文件、年鉴、书刊等，传递到各个经济部门、生产部门、科研部门、教育部门、司法部门及新闻部门，把失效造成的损失化为巨大的效益。

6 失效分析的预测预防的常用分析思路

常用的失效分析思路有二类，一类是以残骸（零件）为对象；一类是安全系统工程分析法，它以失效系统（设备、装置）为范畴。前者以物理、化学的方法为主，着眼于“微观”，后者则以统计、图表和逻辑的方法为主，立足于“宏观”。考虑到本书的篇幅，这里只介绍残骸分析法、统计图表分析法、文字表格法和逻辑分析法。后三种方法均属于安全系统工程分析法。

6.1 残骸分析法

残骸分析法是从物理、化学的角度对失效零件进行分析的方法。如果认为零件的失效是由于零件广义的“失效抗力”小于广义的“应力”的缘故，而“应力”则与零件的服役条件有关，因此，失效残骸分析法总是以服役条件、断口特征和失效的抗力指标为线索的。

零件的服役条件大致可以划分为静载荷、动载荷和环境“载荷”。以服役条件为线索就是要找到零件的服役条件与失效模式和失效原因之间的内在联系。但是，实践表明，同一服役条件下，可能产生不同的失效模式；同样，同一种失效模式，也可能在不同的服役条件下产生，因此，以服役条件

为线索进行失效残骸的失效分析，只是一种初步的“入门”方法，它只能起到缩小分析范围的作用。

“断口”是断裂失效分析重要的证据，它是残骸分析中断裂“信息”的重要来源之一。但是在一般情况下，断口分析必须辅以残骸“失效抗力”的分析，才能对断裂的原因下确切的结论。

如图5.1-4所示，以失效抗力指标为线索的失效分析思路，主要关键是在搞清楚零件服役条件的基础上，通过残骸的断口分析和其他理化分析，找到造成失效的主要失效抗力指标，并进一步研究这一主要失效抗力指标与材料成分、组织和状态的关系。通过材料工艺变革，提高这一主要的失效抗力指标，最后进行机械的台架的模拟试验或直接进行使用考验，达到预防失效的目的。

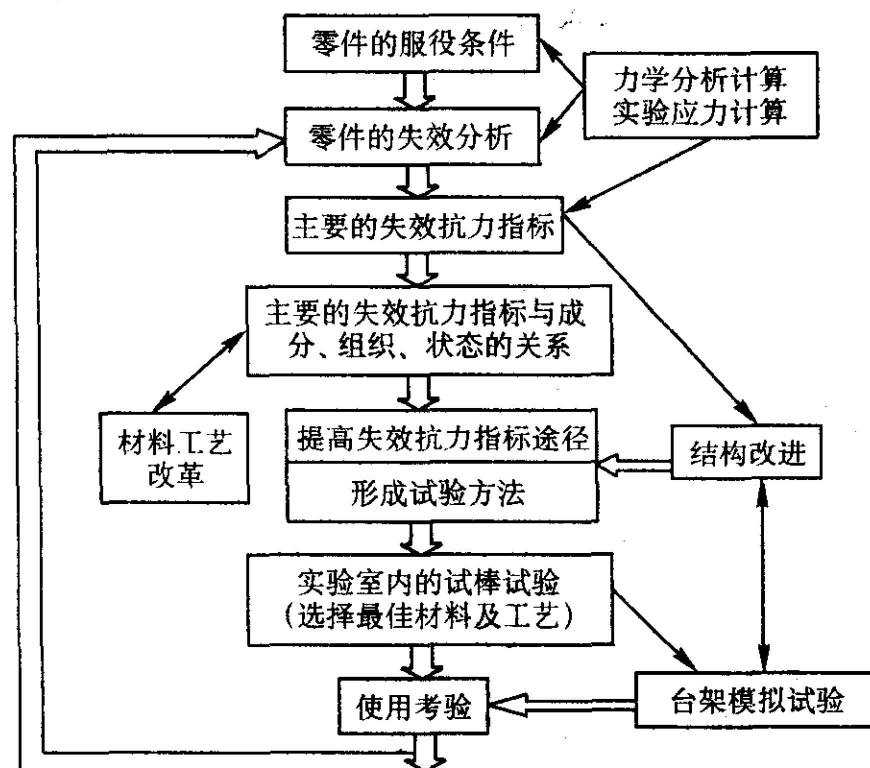


图5.1-4　以失效抗力指标为线索的失效分析思路示意图

十分明显，以失效抗力指标为线索的失效分析思路是一种材料工作者常用的、比较综合的方法。它是工程材料开发、研究和推广使用的有效方法之一。

值得指出的是，在不同的服役条件下，要求零件（或材料）具有不同的“失效抗力”指标的实质是要求其强度与塑性、韧性之间应有合理的配合。因此，研究零件（或材料）的强度、塑性（或韧性）等基本性能及它们之间的合理配合与具体服役条件之间的关系就是这一思路的核心。而进一步研究失效抗力指标与材料（或零件）的成分、组织、状态之间的关系是提高其失效抗力的有效途径（图5.1-5）。

6.2　统计图表分析法

统计图表分析法是安全系统工程事故（或失效）危险识别技术中的实用方法之一。

安全系统工程是安全系统工程学的一个重要分支。它是运用系统工程学的原理和方法，对系统中或生产中的安全问题（或失效问题进行定性和定量的分析、评价及预测，并采取综合安全（或预防）措施予以控制，使系统失效的可能性减少到最低限度，从而达到最佳的安全状态。

安全系统工程是安全技术和安全管理的结合。它的基本内容：事故成因的理论；事故危险的识别技术；事故危险的评价技术和事故的控制技术等。

事故危险识别的主要方法可以分为五种：即统计图表分析法、文字表格法、逻辑分析法、调查实验法和数学解析法。

统计是一种从数量上认识事物的方法。

失效统计是运用科学的统计分析方法，对大量的失效资料的数据进行加工、整理、综合、分析，从而揭示失效产生

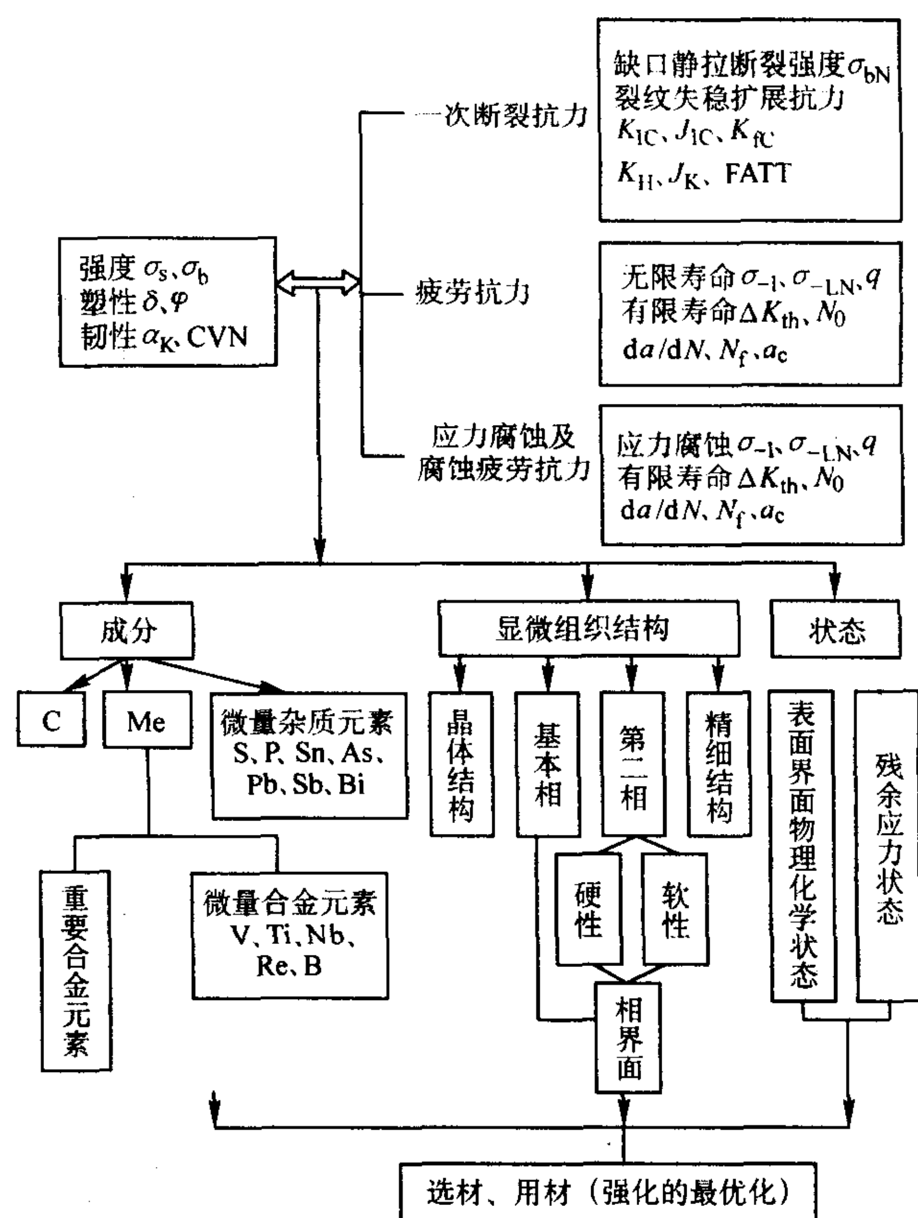


图5.1-5　材料失效抗力指标与成分、组织、状态关系示意图

的规律，为防止失效的发生指明方向。

简单地说，填有统计指标的表格叫做统计表。任何一种统计表，都是统计表格与统计数字的结合体。利用表中的绝对指标、相对指标的平均指标，可以研究各种失效现象的规模、速度和比例关系，因而，它是失效分析的重要工具。

统计图是一种表达统计结果的形式。它用点的位置、线的转向、面积的大小等来表达统计结果，可以形象地直观地研究失效现象的规模、速度、结构和相互关系。

统计图表分析法是利用过去的、现在的资料和数据进行统计，推断未来，并用图表表示。

统计图表分析法可以提供失效发生及发展的一般特点及规律，可供类比，为预测失效准备条件。用于中、短期预测较为有效。

统计图表分析法的优点是：简单易行。但不能考虑失效发生及发展的因果关系，预测精度不高。

使用此法的必要条件是：

1）必须有可靠的历史资料和数据；

2）资料、数据中存在某种规律和趋势；

3）未来的环境和过去相似。

在失效分析中，常用的统计图表有五种：比重图、趋势图、主次图、控制图、分布图。

1）失效比重图　失效比重图是一种表示失效构成情况的平面图形，在平面图上可以形象地反映各种失效构成所占的百分比。一般，用一定弧度对应的面积代表该类失效所占的比重，故称为比重图。

2）失效趋势图　失效趋势图是用连续曲线的升降变化来反映失效的动态变化过程的图形。它可以使我们掌握失效发生的历史过程和趋势。它包括失效趋势的动态曲线图和对数曲线图两种。

动态曲线图按一定的时间间隔统计失效数字，利用曲线的连续变化反映失效的动态变化的图形。

对数曲线图是失效趋势图的一种特殊形式，用于变量变化范围很大的情况。其横坐标表示时距，以等差数列为尺度。纵坐标表示失效数，以对数数列为尺度。

3）失效主次图　主次图是主次排列图的简称。又称为分层排列图，简称排列图。主次图是按数量多少的顺序排列的条形图与累计百分比曲线图相结合的坐标图形。主次图的横坐标可以以失效模式、失效机理、或失效原因为对象。纵坐标的单位为失效的数量或为累计百分比。

分析失效发生的原因时，用主次图可以清楚地、定量地反映出各个因素影响的大小，帮助找出主要原因。

4）失效控制图　失效控制图是在失效趋势图的基础上作出的。它实际是一种标有控制界限的趋势图。

5）失效分布图　失效分布图是为失效地区平面分布图。它可以用于明确失效的多发区域。

6.3 文字表格法

文字表格法是一种常用的失效识别方法之一。它主要包括安全检查表、预计危险分析法（简称PHA）、特性因素图法和失效模式影响和致命度分析法。这里只介绍后两种方法。

（1）特性因素图法

特性因素图法是分析失效原因常用的方法之一。在失效分析中，“特性”是指失效事件或异常现象（结果），“因素”是指引起失效或异常现象的因素（原因），特性因素就是将产品或系统的失效作为“结果”，以导致产品或系统发生失效的诸因素作为“原因”，绘出图形，进而通过图形分析，从错综复杂、多种多样的失效因素中找出造成失效的主要原因的一种图形。因此，特性因素图法又叫因果图法。由于其图形似鱼骨状故又名鱼刺（骨）图。日本质量管理学家石川馨最早使用这种方法，故有时特性因素图又称为石川图。

特性因素图的作法如下。

1）按具体需要选择因果图中的“结果”，放在因果图的最“右边”（相当于“鱼头”）。

2）用带箭头的粗实线或用⇨表示直通“结果”的主干线。

3）通过调查分析，判别出影响“结果”的所有原因。先画出“大原因”，用直线与主干线相连，并在直线尾端常用长方形（或圆圈）框（或圈）起来，在框（或圈）内填入“大原因”的内容；进而依次细分所属的全部原因，直至能采取解决问题的措施为止。各大、小原因之间用不同的直线表示因果之间的关系。

4）对主要的或关键的原因常用▭框起来，以表示醒目。根据实际需要，对这些关键或主要的原因还可以作单独的特性因素图，以便进一步重点地深入分析。

综合上述各点，一般的特性因素图模式见图5.1-6所示。

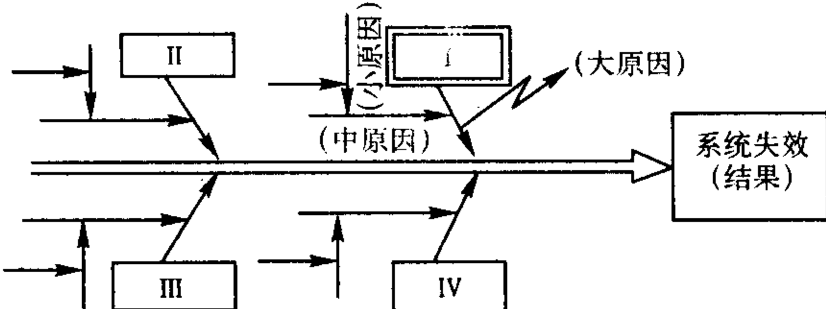


图5.1-6　一般的特性因素图模式示意图

图5.1-7给出压力容器破裂失效分析的特性因素图，可以看到，压力容器的破裂是由于焊接技术不成熟造成的焊接缺陷引起的疲劳断裂。

特性因素图是从错综复杂、多种多样的失效因素中，采取逐渐缩小“包围圈”的办法，最后找出造成失效的主要因素（原因）。这种方法一般多应用于质量管理或用于失效分析的规划，使整个分析工作不遗漏任何因素。但是这种方法本身并不包含“判据”——即根据什么判为疲劳断裂的。此外，这种方法一般只能用作定性分析，而不能用作定量的计算。

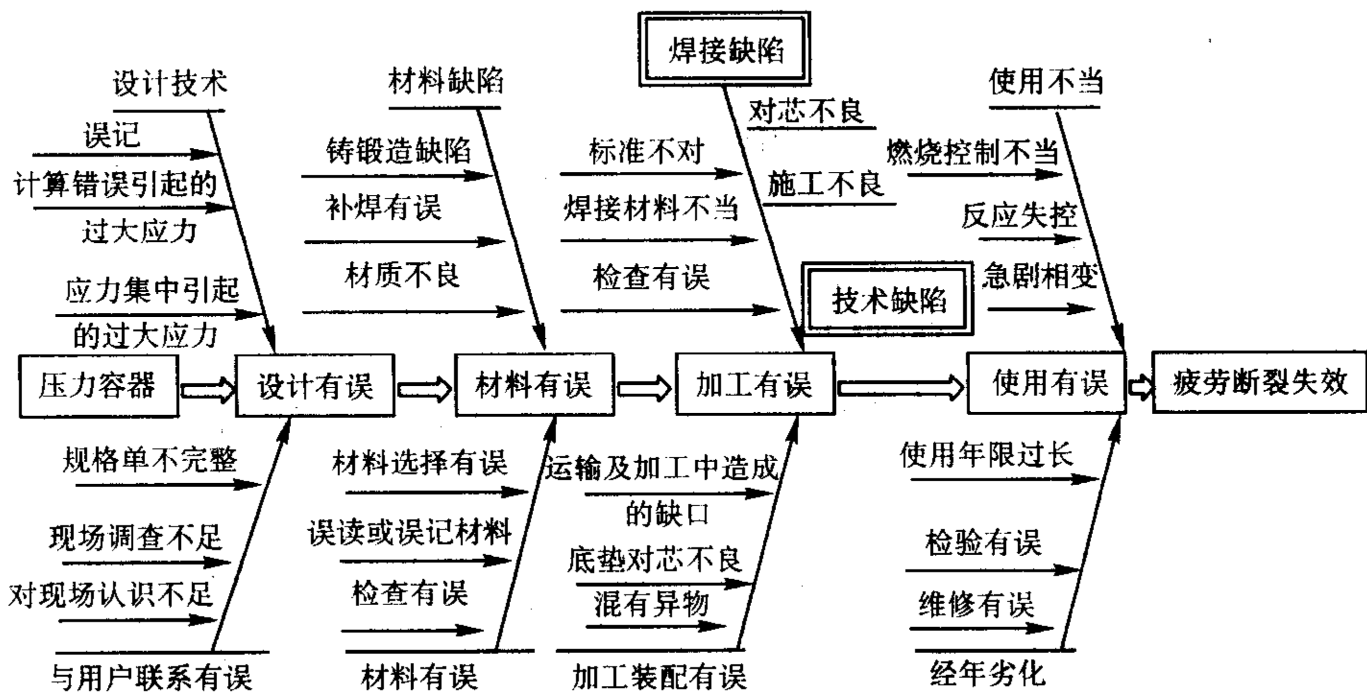


图5.1-7　压力容器疲劳断裂失效特性因素图

（2）失效模式影响和致命度分析法

失效模式影响和致命分析法（Failure Mode and Effects and Criticality Analysis，FMECA）是失效模式分析（FMA）、失效影响分析（REA）以及失效后果分析（FCA）三种分析方法的组合或总称，它是在FMEA（Failure Mode and Effects Analysis）即失效模式影响分析的基础上再加一层任务，即加上判断这种失效模式影响的致命性（或名“致命度”、“危害度”、“严重性”）有多大。也就是说FMECA是FMEA的一种扩展，其扩展的内容就是考虑致命度和失效模式出现的概率。因此，对既定失效模式的致命度分析就称为FMECA。FMECA的基础是FMEA。为了掌握FMECA必先掌握FMEA。

按《可靠性设计指南》规定FMEA是“通过研究可能发生的失效及其系统的影响来鉴定设计产品可靠性的一种技术；是确定失效原因的一种系统方法；是规定为把失效影响减小到最低限度应采取措施的一种系统方法。这种技术和方法可用在整个系统到零件的任何一级”。

1）FMEA（FMECA）的任务

① 在系统的可靠性逻辑关系已知的情况下，列出系统全部元器件的全部失效模式；

② 用归纳推理方法，分析列出各种失效模式对系统的

所有功能级上所造成的影响和后果；

③ 判断每种失效模式对系统各功能级所造成的致命度（严重度、或危害度）大小；

④ 必要时应估计造成各种致命度的概率大小；

⑤ 根据分析结果，提出相应对策。对于那些致命度大的元器件进行更换，或增加冗余度，或修改产品设计等，力求把潜在的、危害度大的失效消灭在设计阶段。

为此，为了进行（FMECA）分析，必须掌握下列原始资料：

① 有关系统结构方面的资料；

② 有关系统使用维护方面的资料；

③ 有关系统所处的环境方面的资料。

为了比较全面了解FMEA（FMECA）分析思路，对其基本步骤、应用案例、优缺点及其发展等方面给以扼要地介绍。

2) FMEA（FMECA）的基本步骤　由于系统在设计和使用中具有多变而又复杂的性质，在进行FMEA（FMECA）分析时，采用标准化的格式是很有必要的。其基本步骤见图5.1-8。

下面就按图5.1-8所示的基本步骤分别加以叙述。

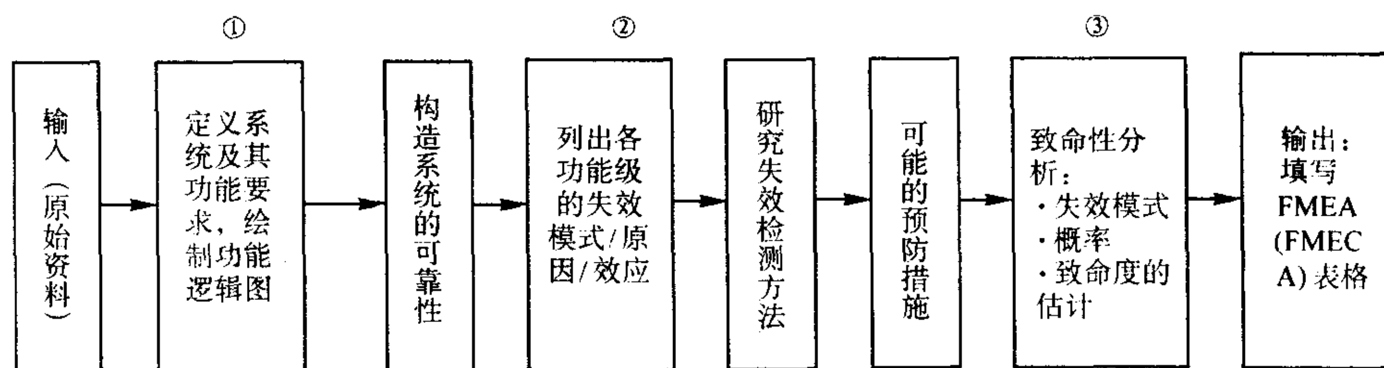


图5.1-8　FMEA（FMECA）的基本步骤流程图

① 定义系统及其功能要求，绘制功能逻辑图　失效分析（或可靠性分析）是针对某一具体对象（即系统或产品、或零件）进行的。FMEA（FMECA）分析思路与其他方法一样，对所分析的对象应有深入的了解。即对系统的任务（功能）给予明确的定义；对系统的运输或工作程序及方式有明确的叙述；对系统使用的环境条件应有规定；对系统失效要有明确定义。进而通过认真的分析，给出系统的功能框图（又称功能逻辑图）。以上就是进行FMEA（或FMECA）分析的第一步。

绘制功能逻辑图是开展FMEA（或FMECA）的基础。所谓功能是指人或物所必须完成的事项，对物而言，功能就是指它的功用、用途。对于某种零部件或产品，一般具有多种功能，这些功能的性质、重要程度往往是不同的。产品的基本功能相对于系统的其他功能来说是主要的、必不可少的功能；上位功能是起目的作用的功能，故又名目的功能；下位功能是起手段作用的功能，又名手段功能。上、下位功能是相对的，一个功能对它的上位功能是手段，即为下位功能，而对它的下位功能来说又是目的，即为上位功能。一般对一个功能追问"它的目的是什么?"时，就可以找到它的上位功能；当追问"实现它的手段是什么?"时，就可以找到它的下位功能。图5.1-9给出小型电筒功能要素间的关系。

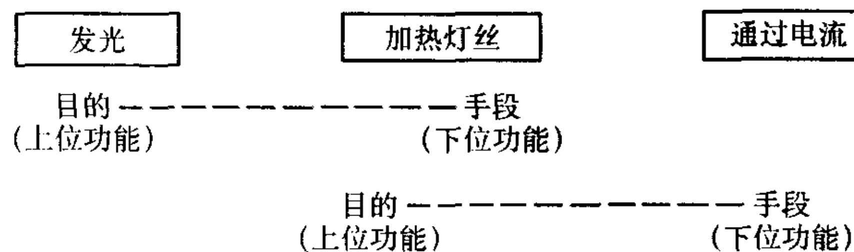


图5.1-9　小型电筒的功能的相互关系

要绘制功能逻辑图，首先要对产品的多种功能进行整理。所谓功能的整理，就是把构成功能的各要素按一定的逻辑体系，相互连接起来，从局部功能和整体功能的相互关系上分析问题，达到真正把握基本功能的目的。

根据功能间的关系，将上位（目的）功能排在最左边，下位功能（手段）排在右边，以此类推，就可得出"功能系统图"了。如图5.1-10所示，F_0 为系统的最上位功能，也可称为最终功能或一级功能；$F_1F_2F_3$ 是 F_0 的下位功能（手段），是三个并列的二级功能；F_{11}、F_{12}、F_{13}是并列的三级功能，余类推。同时 F_1、F_2、F_3 与各自的下位功能组成自己的"功能领域"。

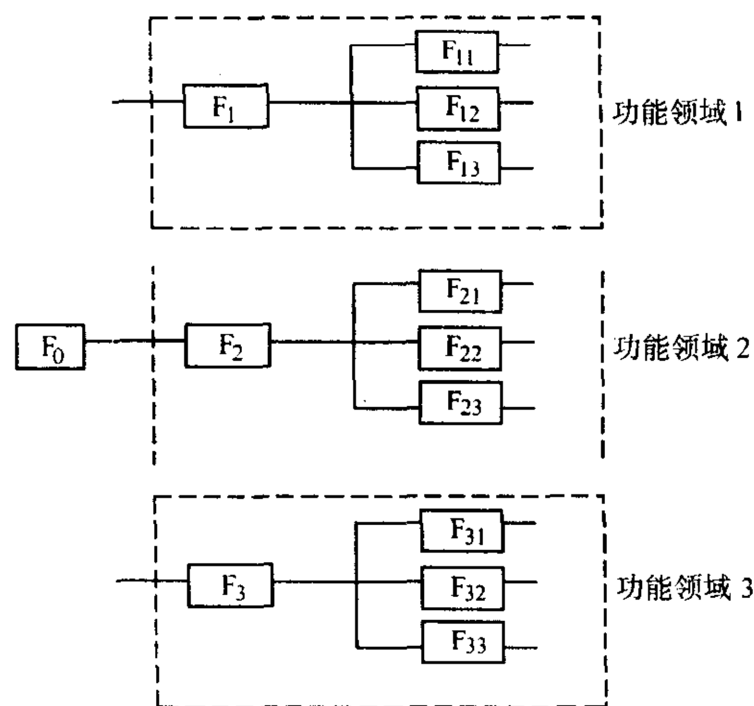


图5.1-10　功能系统图（功能逻辑图）的形式

值得特别指出的是，功能逻辑图不是系统的原理图，也不是系统的结构图、信号流程图，而是一种模型。它表明整个系统各零、部件间的功能逻辑关系，即表示该系统各功能级间的相互依从关系。

功能逻辑图的作用有两方面：一方面它可以作为创造、选定结构可行性方案的依据；另一方面它可作为构造可靠性框图的基础。

② 构造系统的可靠性框图　可靠性框图是"从产品可靠与否的角度去研究系统与组成系统各部件间的逻辑关系"。它是分析各种失效模式所能造成影响的逻辑依据。因此，可靠性框图实际上是一种"可靠性模型"。有了可靠性框图才能算出可靠性的特征值。所以它在可靠性分析、设计和失效分析中，它是用得最广的一种方法。

可靠性框图是在已建立功能逻辑框图的基础上，经过仔细的工程分析和判断，按照"框图的建立原则"建立起来的。

"框图的建立准则"主要有"串联准则"、"并联准则"和"框图中的方块类似电路中开关"等。串联准则认为，当一组部件中任何一个失效都会影响规定任务（功能）的实现时，则这些部件可用串联框图（图5.1-11a）表示；而当一组部件中只有任意一个部件正常，则系统正常，或"只有当全部件失效时，系统才失效"，其可靠性框图可用并联框图

(图5.1-11b) 型式表示。在可靠性框图中，每一个方块看成为电路中的一个“开关”，即当元件正常工作时，“开关”就闭合，“电流”导通；当元件损坏或失效时，“开关”就断开，“电流”被切断。因此在可靠性框图中，在起点S和终点T之间，只要存在任何一条通路，系统工作正常，否则，整个系统就失效了。

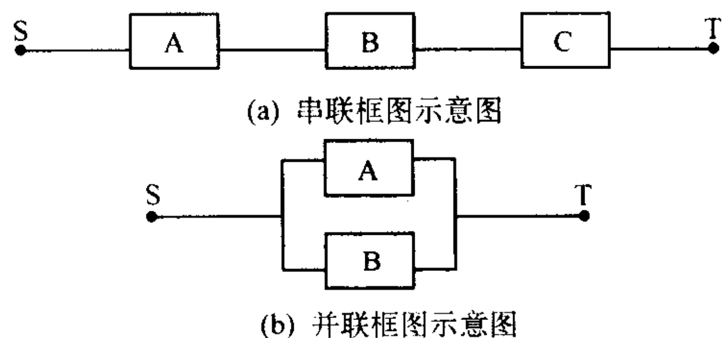


图5.1-11 可靠性框图建立准则

根据以上“准则”，就可以构成较复杂的可靠性框图，并可对其进行可靠性的分析。

从形式上看，可靠性框图与功能逻辑图基本上是相同的。实际上，两者既有联系，又有区别，其联系是可靠性框图是在功能逻辑图的基础上画出来的；其区别是，功能逻辑图表示分系统及其元部件可靠性对系统总的任务（功能）的影响如何，图中的方块是可以在一定范围内移动的（如图5.1-11a中串联系统中的可靠性框图的各方块的位置顺序可以任意排列）。因此决不能把功能逻辑图简单地移植成为可靠性框图。

在建立系统的可靠性框图之后，最好预先准备一张以系统运行为线索，其中包含特殊元部件、运行方式、指标、时间限制和环境在内的预想的失效模式清单，这样可以加快找到那些关键元器件的失效模式，原因和效应的过程。因为一个系统的正常运行与否，往往是以系统的某些关键元部件的性能或失效与否为转换的。下面着重讨论FMEA（FMECA）基本步骤中致命性（或后果）分析。

③ 致命性（或后果）分析　致命性分析（Criticality Analysis）仅适用于FMECA，它是进行失效分析定量化的关键一步。

所谓致命度就是“各种失效模式对系统功能（任务）影响的严重程度”。致命性分析，就是针对某一确定的系统，进行每种失效模式的效应（影响）分析，进而确定每一个元件（部件、系统）的致命度（危害度、严重度）的分析过程。

估计致命度的目的是为了定量计算失效模式对系统损失的概率，在进行致命性分析时需要估计失效模式的概率和致命度计算这样两方面的工作。

失效模式的概率 α_{ij} 系指产品 i 以失效模式 j 而引起产品失效和频数比，即所考虑的次数与该产品全部失效次数之比。每种具体产品的特定的失效模式出现的概率可通过试验、分析、推导、估计来获得其定量的数据，也可以类似产品的失效模式概率统计数据为基础，经过修正后获得。

3) 系统致命度CR的计算方法

① 用网络图分析法求系统致命度CR　网络图是以失效的严重程度的等级（如分为Ⅰ、Ⅱ、Ⅲ、Ⅳ等）作为纵坐标、失效模式的概率作为横坐标构成的图。每一种失效模式在上述图上就有一点。如图5.1-12上的 a 点或 b 点，失效模式的致命度 CR 可用 a 点或 b 点连线的长短来表示，其连线越长，致命度也就愈高。

从图5.1-12中可以看出，b 点的致命度比 a 点的高，而 A 点的致命度为最大。

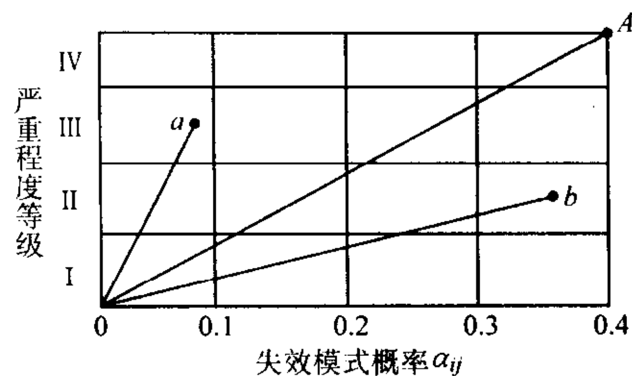


图5.1-12 求致命度的网络图

网络图分析法使用方便，它可很快定出失效模式的致命度大小，但是它只有各种失效模式的致命度相对大小而难以严格的定量化。

② 解析法求系统致命度 CR　其具体的求法有以下二种：

第一种，求 CR_s 的基本公式

元部件 i 以失效模式 j 发生失效致使该元部件 i 发生失效的致命度 CR_{ij} 可按下式计算，

$$CR_{ij} = \alpha_{ij}\beta_{ij}\gamma_i \tag{5.1-1}$$

式中，α_{ij} 为元部件 i 以失效模式 j 而引起该元部件失效的失效模式的频数比；β_{ij} 为元部件 i 以失效模式 j 发生失效造成该元部件损伤（或失效）的概率；γ_i 为元部件 i 的基本失效率（查有关手册或通过试验得到）。

国际上将 β_{ij} 称为丧失功能的条件概率。此概率可采用下述值：

$\beta_{ij}=1$，表示产品肯定发生损伤或失效

$\beta_{ij}=0.5$，表示产品可能发生损伤或失效

$\beta_{ij}=0.1$，表示产品很少可能发生失效

$\beta_{ij}=0$，表示对产品无影响

系统的致命度 CR_s 为

$$CR_s = \Sigma_i\Sigma_j CR_{ij} \tag{5.1-2}$$

式中，Σ_i 为元部件 i 全部失效模式使系统发生失效的失效概率之和；Σ_j 为全部元件使系统发生失效的失效概率之和。

第二种，考虑应力变化时，系统的致命度可按下式计算：

$$CR_s = \sum_{n=1}^{j}\beta\alpha K_A K_F \gamma_G t \times 10^6 \quad (n=1,\ 2,\ \cdots,\ j) \tag{5.1-3}$$

式中，CR_s 为每当系统完成 10^6 次使用时，由元部件失效致使这类系统失效的致命度；n 为导致系统发生失效的元部件的失效模式数；j 为导致系统发生失效的元部件的最后失效模式数；t 为系统完成一次任务时，其元部件的工作时间或工作次数；γ_G 为每小时或每一工作周期内，元部件的基本失效率；K_A 为当测量 γ_G 时的工作应力同元件部件使用时的应力有差异时，用以修正基本失效率 γ_G 的系数；K_F 为当测量 γ_G 的环境应力同元部件使用时的环境应力有差异时，用以修正基本失效率 γ_G 的系数；α 和 β 的含义同前。

4) 以“风险”大小计算致命度 CR_s　当各种失效模式引起系统失效的概率和损失的大小已知时，系统的致命度 CR_s。可用“风险”大小来计算，其表达式为

$$CR_s = \sum_{i=1}^{n} P_i \times C_i \quad (i=1,2,\cdots n) \tag{5.1-4}$$

式中，P_i 为元部件各种失效模式引起系统失效的概率；C_i 为元部件各种失效模式引起系统失效的损失大小。

式（5.1-4）反映了一个直观的基本概念，即系统的致命度随它的发生概率及其导致的损失增加而增加。

以上三种方法，第一种方法一般用于产品设计的初期，第二种方法多半用于已经实验室或现场使用的产品；而第三种方法常用于需要考虑失效带来的损失的情况下。

5）FMEA（FMECA）的优缺点　FMEA（FMECA）有如下优点：

① 它适用于产品研制的全过程，适用于研制中的各个阶段，适用于电气、机械、民用、宇航等专业；

② 它可以帮助研制人员把失效及影响减少到最小，从而提高产品或系统的可靠性水平；

③ FMEA的原理简单，方法简便，基本是定性分析，也可进行定量分析；

④ 应用FMEA（FMECA）的实际效果较大，国外早就列入产品研制，我国亦即把FMECA定为国标；

⑤ 它可以在一定程度上反映人的因素（如操作上）所引起的失误等；

⑥ 它是其他失效分析的基础之一，它既可以独立使用，也可作为可靠性定量分析方法的补充和保证。若与其他失效分析综合使用，其收获会更大。

FMEA（FMECA）存在的主要不足之处如下：

① 在国际草案里指出："此方法用于分析元件导致整个系统的失效非常有效"，但对多功能和具有大量元件的复杂系统，实施起来就较为困难、繁琐；

② 工作量大，很费时间，工作显得比较枯燥；

③ 它是一种单因素的分析方法，FMEA本质上是单因素的失效分析。对于多因素同时起作用，或相互作用而导致一种后果情况就难以分析；

④ 在进行失效分析时，在环境效应（影响）事关重大的情况下，FMEA法有其局限性。

为了发挥FMEA法之长处而避其所短，一方面FMEA本身在不断的发展中，比如FMEA的矩阵分析法就是它的新的发展，这对扩大FMEA的用途具有重要的意义；另一方面它最好与其他失效分析方法（如FTA、ETA）综合使用。关于FMEA的矩阵分析法，可进一步参考有关的专业文献。

6.4　失效树分析法

失效树分析法是一种逻辑分析方法。

逻辑分析法包括事件树分析法（简称ETA）、管理失误和风险树分析法（简称MORT）和失效树分析法（简称FTA）等。这里只介绍失效树分析法。

失效树分析（Fault Tree Analysis）早在20世纪60年代初就由美国贝尔研究所首先用于民兵导弹的控制系统设计上，为预测导弹发射的随机失效概率作出了贡献。此后许多人对失效树分析的理论和应用进行了研究。1974年美国原子能管理委员会主要采用失效树分析商用原子反应堆安全性的Wash-1400报告发表，进一步推动了对失效树的研究和应用。迄今FTA法在国外已被公认为当前对复杂安全性、可靠性分析的一种好方法。

失效树分析法是：在系统设计过程中，通过对可能造成系统失效的各种因素（包括软件、硬件、环境、人为因素等）进行分析，画出逻辑框图（即失效树），从而确定系统失效原因的各种可能的组合方式或发生概率，以计算系统失效概率，采取相应的纠正措施，以提高系统可靠性的一种设计分析方法。

FTA法具有很大的灵活性，即不是局限于对系统可靠性作一般的分析，而且可以分析系统的各种失效状态。不仅可以分析某些元部件失效对系统的影响，还可以对导致这些元部件失效的特殊原因进行分析。

FTA法是一种图形演绎方法，是失效事件在一定条件下的逻辑推理方法。它可以围绕某些特定的失效状态作层层深入的分析。因而在清晰的失效树图形下，表达了系统的内在联系，并指出元部件失效与系统之间的逻辑关系，找出系统的薄弱环节。

FTA法不仅可以进行定性的逻辑推导分析，而且可以定时地计算复杂系统的失效概率及其他的可靠性参数，为改善和评估系统的可靠性提供定量的数据。

FTA法的步骤，因评价对象、分析目的、精细程度等而不同，但一般可按如下的步骤进行：

① 失效树的建造；

② 失效树的定性分析；

③ 失效树的定量分析；

④ 基本事件的重要度分析。

下面就分别加以扼要的介绍。

（1）失效树的建造

失效树的建造是FTA法的关键，因为失效树建造的完善程度将直接影响定性分析和定量计算结果的准确性。失效树的建造工作十分庞大烦杂，所以要求建树者必须十分慎重、仔细、并要广泛地掌握设计、使用等各方面的经验和知识。

建树一般可按以下步骤进行：

① 广泛收集并分析有关的技术资料；

② 选择顶事件；

③ 建树；

④ 失效树的简化。

建树的过程事实上是对系统仔细、透彻分析的过程。不同的人从不同的角度所建的失效树是不相同的，目前还没有一种有效的统一的建树方法。一般的建树方法可分为两大类：演绎法和计算机辅助建树的合成或决策表法。这里先介绍演绎法。

演绎法是这样进行的：先写出顶事件（即系统不希望发生的失效事件）作为第一行，在它的下面并列地写出导致顶事件发生的直接原因——包括硬件失效、软件失效、环境因素和人为的因素等——作为第二行，在第一行与第二行之间，用相应的符号（逻辑门）表示出它们之间的逻辑关系。如果还要进一步分析造成第二行事件的原因，则把导致第二行的那些失效事件（称为中间事件）发生的直接原因作为第三行，同样应用适当的逻辑门将第二行与第三行联系起来。按照这样的方法步步深入，一直追溯到不需要（或不可能）继续分析的原因（称为底事件）为止。这样，就建成了一棵以顶事件为"根"、中间事件为"节"、底事件为"树叶"的、具有几级的、倒置的失效树。

从以上建树过程可以看出，要建一棵树首先需要解决的问题是失效树中所使用的符号。

失效树中使用的符号如表5.1-2所示。

失效树的建造是一件十分复杂和仔细的工作，要求注意以下几点：

① 失效分析人员在建树前必须对所分析的系统有深刻的了解；

② 失效事件的定义要明确，否则树中可能出现逻辑混乱乃至矛盾、错误；

③ 选好顶事件，若顶事件选择不当就有可能无法分析和

表 5.1-2 失效树中使用的主要逻辑符号

符 号	事 件	说 明
	事件	在失效树中指出的由其他事件引起的某一事件
C（输出）（输入）AB	或门	A 或 B 事件可导致 C 事件时使用的逻辑门
C（输出）（输入）AB	与门	A 和 B 事件必须都出现才能使 C 事件发生时使用的逻辑门
C B A	禁门	只有当 B 规定的条件存在时，事件 A 才能导致事件 C 时用的逻辑门
	不发展事件	不再发展的事件
	基本事件	为了更完整地使用失效树，必须指出的不进一步发展的事件
A	转换	转到失效树的其他部分

计算。对同一个系统，选取不同的顶事件，其结果是不同的。在一般情况下，顶事件可以通过初步的失效分析，可从各种失效模式中找出该系统最可能发生的失效模式作为顶事件；

④ 合理确定系统的边界条件——规定所建立的失效树的状况。有了边界条件就明确了失效树建到何处为止。边界条件一般包括确定顶事件、确定初始条件和确定不许可的事件等；

⑤ 对系统中各事件之间的逻辑关系及条件必须分析清楚，不能有逻辑上的紊乱及条件上的矛盾。

为了使失效树有一个感性的认识，下面举例讨论失效树是如何进行失效分析的。

例如，低合金超高强度钢一般在低温回火或等温（马氏体等温或贝氏体等温）淬火状态下使用。在服役期间，低合金超高强度钢也常发生断裂失效（破坏）。失效树的顶事件就是构件的破坏。这种破坏可由不同的事件——疲劳、过载、应力腐蚀开裂及具有最大可能性的氢脆等等——造成的。这些事件，每一个都通过“或”门与顶事件相连（图 5.1-13）。断口分析表明，失效残骸的断口形态不同于过载和疲劳。因此，过载和疲劳是“不发展事件”，并分别用棱形表示在图 5.1-13 中。当然如果断口分析不能排除这些事件时，那么仍有必要进一步地发展。对于氢脆来说，它是在临界应力强度和临界含氢量共同作用下发生的，因此临界应力强度（图 5.1-13 中的事件 15）和临界含氢量（图 5.1-13 中的事件 14）应采用“与门”与氢脆（图 5.1-13 中的事件 4）相连，其中临界含氢量为不发展事件。应力腐蚀开裂（事件 3）则是临界应力强度（事件 6）和造成开裂元素的临界浓度可以是临界氢浓度（事件 10），也可以是除氢以外的其他物质的临界含量（事件 11），这样事件 10 和事件 11 应

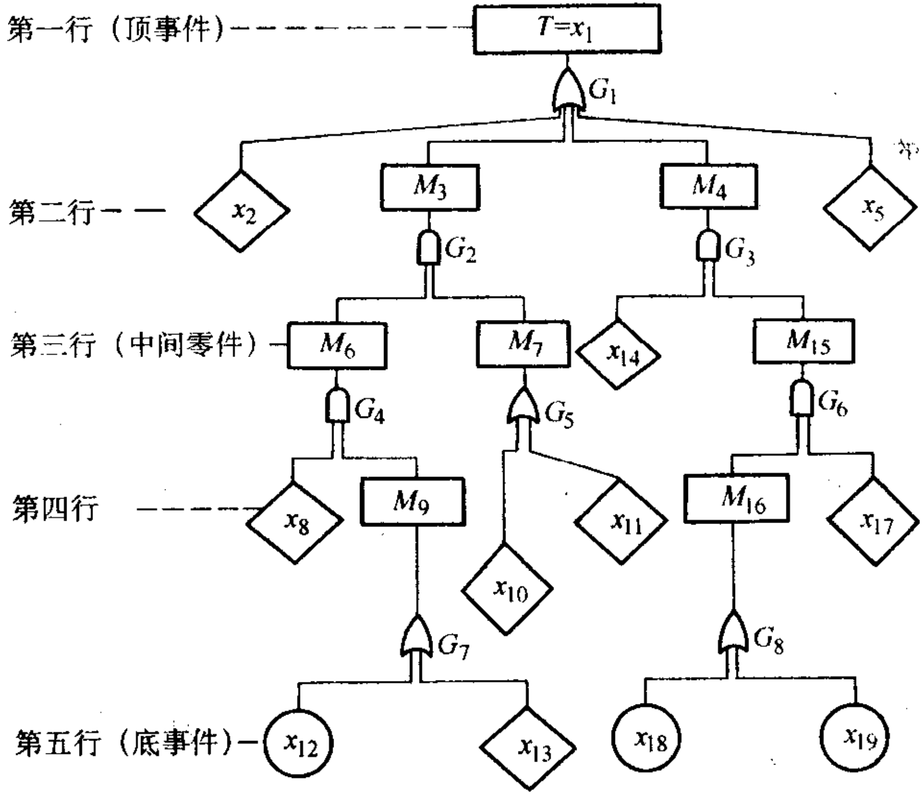


图 5.1-13 某超高强度钢构件破坏的失效树

1—构件破坏；2—过载；3—应力腐蚀；4—氢脆；5—疲劳；6—临界应力强度；7—造成开裂元素的临界浓度；8—构件上的负载；9—流变应力 > σ_i；10—临界氢含量；11—除氢以外，其他物质的临界含量；12—加工缺陷 > σ_i；13—使用过程中的发展；14—临界氢含量；15—临界应力强度；16—流变应力 > σ_i；17—构件上的负载；18—施加负载 > P_i；19—残余应力 > S_i

用“或门”与事件 7 相连。事件 10 和事件 11 均为不发展事件，故均用棱形框起来。可以看出，如果认为应力腐蚀开裂与氢脆都是由于临界应力强度上临界氢浓度引起的；那么，在失效树的这一行不能区分应力腐蚀开裂和氢脆，不过，应力腐蚀开裂和氢脆应该在断裂源的起始位置上找到差别。应力腐蚀开裂的临界氢浓度应在暴露表面上显示出来，因此它的断裂源一般在“暴露表面上”，而氢脆的临界氢浓度可能在电镀表面或次表面先达到，因此它的断裂源应在电镀表面上或次表面上。所以到底是应力腐蚀开裂还是氢脆在失效树的第二行就可以初步确定了。虽然应力腐蚀开裂和氢脆的条件之一都是临界应力强度，并且它们临界应力强度都取决于构件上的负载（事件 8 和事件 16）和材料的流变应力大于材料的临界门槛应力 σ_i（当然，应力腐蚀的门槛应力数值与氢脆的门槛应力数值不同），但是由于应力腐蚀开裂一般起始于暴露表面，构件的表面流变应力对构件的平均负载不敏感，而对表面的加工缺陷等原因所造成应力集中或应变集中则十分敏感，因而在应力腐蚀系统中，加工缺陷处的流变应力大于材料的应力腐蚀门槛应力用“或门”与事件 9 相连；在氢脆系统中，由于氢脆一般起源于电镀层的次表面，构件上的负载（事件 16）可以是施加的负载（事件 18）也可以是构件内部的残余应力（事件 19），故事件 18 和事件 19 用“或门”与事件 16 相连。材料的氢脆门槛应力受表面加工缺陷的影响较小，不需要进一步的展开分析（事件 17 为不发展事件）了。

从以上 FTA 法在构件断裂失效分析中的具体应用情况中可以看出，FTA 法可以对特定的失效事件作层层深入地逻辑推理分析，在清晰的失效树的帮助下，最后找到这一特定失效事件的失效原因或该构件的薄弱环节，因此，FTA 法是进行残骸失效分析的好方法之一。

(2) 失效树的定性分析

失效树的定性分析的目的是为了寻找系统的最薄弱的环节，即发现系统最容易发生失效的环节，以便集中力量解决这些薄弱环节，提高系统的可靠性。

1）最小割集及其求法　所谓最小割集是指包含在这个集合里的全部基本事件发生是造成顶事件发生的必要充分条件，它表示系统的危险性，每一个最小割集都是顶事件发生的一种可能，即最小割集表示了哪些故障和差错同时发生时，顶事件就发生。失效树中有几个最小割集，顶事件的发生就有几种可能。最小割集越多，系统越危险。

最小割集的求法很多，常用的方法有行列法，又叫福塞尔（Fussell）法及布尔代数化简法。

行列法是以“与门”使割集容量增大、“或门’使割集数量增加为基本原则，从顶事件开始，依次把上一行事件代换成下一行事件，代换的方法为把“与门”事件横向排列起来，“或门”事件纵向并列起来，这样一直代换下去，直到全部是基本事件为止。再用布尔代数化简，先去掉每行割集中的重复元素，消去重复的行（割集），就得到失效树的最小割集了。用布尔代数化简法求最小割集法，就是用布尔代数直接对失效树进行化简的逻辑运算，最后化简得到的若干交集，每一个交集就是一个最小割集。虽然用布尔代数化简法求系统的最小割集比起行列法要简单得多，但行列法仍得到国际上普通承认，其重要的原因是它便于计算机编程求其最小割集。

① 用行列法简法求失效树的最小割集。

行列法的计算过程列表（表5.1-3）如下。

表5.1-3　行列法求失效树的最小割集的过程

步骤	1	2	3	4
过程	x_2	x_2	x_2	x_2
	M_3	M_6，M_7	x_6，M_9，x_{10}	x_8，x_{12}，x_{10}
	M_4	x_{14}，M_{15}	x_8，M_9，x_{11}	x_8，x_{13}，x_{10}
	x_5	x_5	x_{14}，M_{16}，x_{17}	x_8，x_{12}，x_{11}
			x_5	x_8，x_{13}，x_{11}
				x_{14}，x_{18}，x_{17}
				x_{14}，x_{19}，x_{17}
				x_5

从第一行（顶事件）至第二行为行列法中的步骤1，由于 G_1 是“或门”，因此，x_2、M_3、M_4 和 x_5 等纵向并列起来。从第二行至第三行为步骤2，由于 G_2、G_3 均为“与门”，因此 M_6、M_7 横向排列起来，并取代 M_3，x_{14}、M_{15} 横向排列起来取代 M_4。第三行至第四行为步骤3，由于 G_4 是“与门”，故 x_8、M_9 横向取代 M_6，G_5 是“或门”，故 x_{10}、x_{11} 纵向并列，并取代 M_7。从第四行至第五行为步骤4，由于 G_7、G_8 均为“或门”，因此 x_{12}，x_{13} 和 x_{18}、x_{19} 纵向并列起来分别取代 M_9、M_{16}，结果最后得到最小割集，即：

$\{x_2\}$，$\{x_8、x_{12}、x_{10}\}$，$\{x_8、x_{13}、x_{10}\}$，$\{x_8、x_{12}、x_{11}\}$，$\{x_8、x_{13}、x_{11}\}$，$\{x_{14}、x_{18}、x_{17}\}$，$\{x_{14}、x_{19}、x_{17}\}$ 和 $\{x_5\}$

这就是说，只要上述8个最小割集里的全部基本事件发生，就会造成顶事件的发生——构件断裂。

② 用布尔代数化简法求失效树的最小割集：

$$\begin{aligned}T &= G_1 = x_2 + G_2 + G_3 + x_5 \\ &= x_2 + G_4 G_5 + x_{14} G_6 + x_5 \\ &= x_2 + x_8 G_7 (x_{10} + x_{11}) + x_{14}(G_8 \cdot x_{17}) + x_5 \\ &= x_2 + x_8 (x_{12} + x_{13})(x_{10} + x_{11}) + x_{14}(x_{18} + x_{19}) \cdot x_{17} + x_5 \\ &= x_2 + (x_8 x_{12} + x_8 x_{13})(x_{10} + x_{11}) + x_{14} x_{18} x_{17} + x_{14} x_{19} x_{17} + x_5 \\ &= x_2 + x_8 x_{12} x_{10} + x_8 x_{12} x_{11} + x_8 x_{13} x_{10} + x_8 x_{13} x_{11} + x_{14} x_{18} x_{17} + \\ &\quad x_{14} x_{19} x_{17} + x_5 \end{aligned} \tag{5.1-5}$$

其结果也得到与行列法相同的8个最小割集，即：

$\{x_2\}$，$\{x_8、x_{12}、x_{10}\}$，$\{x_8、x_{13}、x_{10}\}$，$\{x_8、x_{12}、x_{11}\}$，$\{x_8、x_{13}、x_{11}\}$，$\{x_{14}、x_{18}、x_{17}\}$，$\{x_{14}、x_{19}、x_{17}\}$，$\{x_5\}$

2）最小径集及其求法　与最小割集的意义相反的是最小径集。最小径集是指包含在这个集合内部的基本事件不发生，就保证顶事件不发生的基本事体的最小组合，它表示系统的安全性。即有几个最小径集就有几种使顶事件不发生的可能方案，它为控制事故提供依据。失效树中最小径集越多，系统越安全。

最小径集的求法是利用它与最小割集的对偶性的原理进行的。先将失效树画成与之对偶的“成功树”（即把失效树中的“与门”变成“或门”，“或门”变成“与门”，其他事件不变，这样就得了成功树），然后再按求最小割集的方法来求成功树的最小割集，而成功树的最小割集便是原失效树的最小径集。

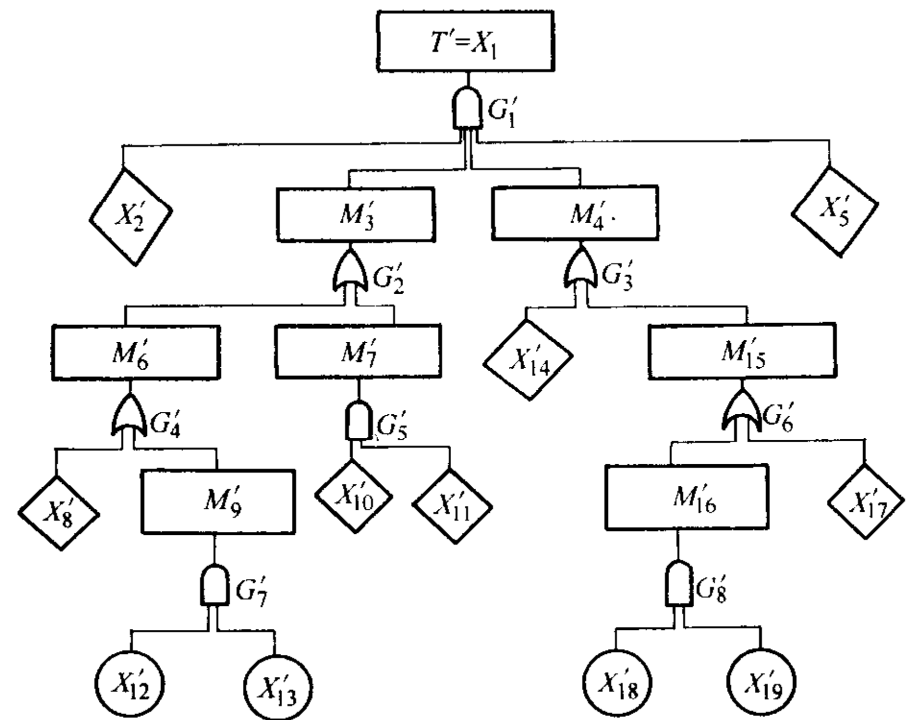


图5.1-14　与图5.1-13为对偶的成功树

为了与最小割集进行对比，下面仍以图5.1-13的故障树为例，求它的最小径集。

以图5.1-13为对偶的成功树如图5.1-14所示。用布尔代数化简法求图5.1-14的最小割集如下：

$$\begin{aligned}T' &= G'_1 = X'_2 M'_3 M'_4 X'_5 = X'_2 (M'_6 + M'_7)(X'_{14} + M'_{15}) X'_5 \\ &= X'_2 M'_6 X'_{14} X'_5 + X'_2 M'_6 M'_{15} X'_5 + X'_2 M'_7 X'_{14} X'_5 + \\ &\quad X'_2 M'_7 M'_{15} X'_5 \\ &= X'_2 (X'_8 + M'_9) X'_{14} \cdot X'_5 + X'_2 (X'_8 + M'_9) M'_{15} X'_5 + \\ &\quad X'_2 X'_{10} \cdot X'_{11} X'_{14} X'_5 + X'_2 X'_{10} X'_{11} (M'_{15}) X'_5 \\ &= X'_2 X'_8 X'_{14} X'_5 + X'_2 X'_9 X'_{14} X'_5 + X'_2 X'_8 M'_{15} X'_5 + \\ &\quad X'_2 X'_9 M'_{15} X'_5 + X'_2 X'_{10} X'_{11} X'_{14} X'_5 + X'_2 X'_{10} X'_{11} X'_{16} X'_5 \\ &\quad + X'_2 X'_{10} X'_{11} X'_{17} X'_5 \\ &= X'_2 X'_8 X'_{14} X'_5 + X'_2 X'_{12} X'_{13} X'_{14} X'_5 + X'_2 X'_8 X'_{18} X'_{19} \\ &\quad X'_5 + X'_2 X'_8 X'_{17} X'_5 + X'_2 X'_{12} X'_{13} X'_{18} X'_{19} X'_5 + X'_2 X'_{12} \\ &\quad X'_{13} X'_{17} X'_5 + X'_2 X'_{10} X'_{11} X'_{14} X'_5 + X'_2 X'_{10} X'_{11} X'_{18} X'_{19} \\ &\quad X'_5 + X'_2 X'_{10} X'_{11} X'_{17} X'_5 \end{aligned} \tag{5.1-6}$$

结果得到9个交集，这9个交集是图5.1-14成功树的9个最小割集，也是图5.1-13失效树的9个最小径集，即

$\{X'_2，X'_8，X'_{14}，X'_5\}$，$\{X'_2，X'_{12}，X'_{13}，X'_{14}，X'_5\}$，$\{X'_2，X'_8，X'_{18}，X'_{19}，X'_5\}$，$\{X'_2，X'_8，X'_{17}，X'_5\}$，$\{X'_2，X'_{12}，X'_{13}，X'_{18}，X'_{19}，X'_5\}$，$\{X'_2，X'_{12}，X'_{13}，X'_{17}，X'_5\}$，$\{X'_2，X'_{10}，X'_{11}，X'_{14}，X'_5\}$，$\{X'_2，X'_{10}，X'_{11}，X'_{18}，X'_{19}，X'_5\}$ 和 $\{X'_2，X'_{10}，X'_{11}，X'_{17}，X'_5\}$

这就是说，图5.1-14系统不发生失效的方案有以上9个。

(3) 失效树的定量分析

失效树的定量分析的任务就是要计算或估计系统顶事件发生的概率及系统的一些可靠性指标。一般来说，多部件复杂系统的失效树定量分析是十分困难的。有时无法用解析法求其精确结果，而只能用一些简化的方法进行估算。

1) 利用失效树的最小割集计算顶事件的发生概率 如果各最小割集中没有重复的基本事件，顶事件的发生概率按下式计算：

$$g = \sum_{r=1}^{k} \prod_{x_i \in k_r} q_i \quad (5.1\text{-}7)$$

式中，g 为顶事件的发生概率；i 为基本事件的序数；q 为基本事件 i 的发生概率；r 为最小割集的序数；k 为最小割集的个数；$x_i \in k_i$ 第 i 个基本事件属于第 r 个最小割集；$\prod$ 为求概率积；$\sum$ 为求概率和。

如果各最小割集中彼此重复的基本事件，求顶事件发生概率时，必须将上式展开，消去每个概率积中的重复因子。即：

$$g = \sum_{r=1}^{k} q_i - \sum_{1 \leq r \leq s \leq k} \prod_{x_i \in k_r \cup k_s} q_i + \cdots + (-1)^{k-1} \prod_{\substack{r=1 \\ x_i \in k_r}}^{k} q_i \quad (5.1\text{-}8)$$

式中，r、s 为最小割集的序数；$\sum_{r=1}^{k}$ 为求 k 项代数和；$\sum_{1 \leq r \leq s \leq k} \prod_{x_i \in k_r \cup k_s}$ 为属于任意两个不同最小割集的基本事件概率积的代数和；$x_i \in k_r \cup k_s$ 为第 i 个基本事件，或属于第 r 个最小割集，或属于第 s 个最小割集；$1 \leq r < s \leq k$ 为任意两个最小割集的组合顺序。

由上可知，计算是十分繁琐的。在精确度要求不高的情况下，可以仅用上式的首项作近似的计算，即用下式计算

$$g \doteq \sum_{r=1}^{k} \prod_{x_i \in k_r} q_i \quad (5.1\text{-}9)$$

2) 对图 5.1-13 失效树进行定量的计算 首先对树中各基本事件（X_i）的发生概率进行半定量的取值（见表 5.1-4），按上式计算

$$g \doteq \sum_{r=1}^{k} \prod_{x_i \in k_r} q_i = q_2 + q_8 q_{12} q_{10} + q_8 q_{12} q_{11} + q_8 q_{13} q_{10} + q_8 q_{13} q_{11} + q_{14} q_{18} q_{17} + q_{14} q_{19} q_{17} + q_5 = 4.22 \times 10^{-6}$$

表 5.1-4 基本事件发生概率取值表

代号	基本事件名称	发生概率 q_i	$1-q_i$
x_2	过载断裂	10^{-8}	0.000 000 01
x_5	疲劳断裂	10^{-8}	0.000 000 01
x_8	构件上的负载	10^{-3}	0.001
x_{10}	临界氢含量	10^{-1}	0.1
x_{11}	除氢外其他物质的临界含量	10^{-1}	0.1
x_{12}	加工缺陷 > σ_i	10^{-2}	0.01
x_{13}	使用过程中的发展	10^{-3}	0.001
x_{14}	临界氢含量	10^{-1}	0.1
x_{17}	构件负载	10^{-3}	0.001
x_{18}	施加负载 > P_i	10^{-2}	0.01
x_{19}	残余应力 > S_i	10^{-2}	0.01

(4) 基本事件的重要度分析

基本事件在整个系统中的重要程度分析包括三个方面的内容，即结构重要度分析、概率重要度分析和临界重要度分析。

1) 结构重要度分析 结构重要度分析是从失效树的结构上分析各基本事件的重要程度，其分析方法常用的有用真值表计算结构重要系数、利用最小割集或最小径集判断其重要度及用概率重要系数公式来估计结构重要系数。

这里采用一种简便方法来分析图 5.1-13 失效树中各基本事件的结构重要度，即给每一个最小割集都赋予值 1，而最小割集中的每个基本事件都得到相等的一份，然后计算每个基本事件的累积得分，根据得分多少，排列出其结构重要度顺序。

经过计算得各基本事件的累积分数如下：

$$x_8 = 4 \times \frac{1}{3}$$

$$x_2 = x_5 = 1$$

$$x_{10} = x_{11} = x_{12} = x_{13} = x_{14} = x_{17} = 2 \times \frac{1}{3}$$

$$x_{18} = x_{19} = 1 \times \frac{1}{3}$$

故其结构重要度顺序为：

$I_\phi(8) > I_\phi(2) = I_\phi(5) > I_\phi(10) = I_\phi(11) = I_\phi(12) = I_\phi(13) = I_\phi(14) = I_\phi(17) > I_\phi(18) = I_\phi(19)$

2) 概率重要度分析 概率重要度分析是考查各基本事件发生概率的发生概率的变化对顶事件概率的影响程度。即对顶事件发生概率 g 函数求各基本事件发生概率 q_i 的一阶偏导，从而得到各基本事件的概率重要系数，即

$$I_g(i) = \partial g / \partial q_i \quad (5.1\text{-}10)$$

如果顶事件的发生概率 g 函数用其首项近似表示，即

$$g \doteq \sum_{r=1}^{k} \prod_{x_i \in k_r} q_i = q_2 + q_8 q_{12} q_{10} + q_8 q_{12} q_{11} + q_8 q_{13} q_{10} + q_8 q_{13} q_{11} + q_{14} q_{18} q_{17} + q_{14} q_{19} q_{17} + q_5 \quad (5.1\text{-}11)$$

那么，各基本事件的概率重要度系数分析为：

$I_g(8) = 2.2 \times 10^{-6}$

$I_g(12) = I_g(13) = 2 \times 10^{-4}$

$I_g(18) = I_g(19) = 1 \times 10^{-4}$

$I_g(14) = I_g(17) = 2 \times 10^{-5}$

$I_g(10) = I_g(11) = 11 \times 10^{-5}$

$I_g(2) = I_g(5) = 1 \times 10^{-8}$

概率重要度顺序为

$I_g(8) > I_g(12) = I_g(13) > I_g(18) = I_g(19) > I_g(14) = I_g(17) > I_g(10) = I_g(11) > I_g(2) = I_g(5)$

3) 临界重要度分析 临界重要度是从基本事件发生概率的变化对顶事件发生概率的影响程度（即概率重要度）和基本事件自身发生概率的双重角度来衡量各基本事件的重要性，其计算公式如下：

$$CI_g(i) = q_i I_g(i) / g \quad (5.1\text{-}12)$$

式中，$CI_g(i)$ 为 i 基本事件的临界重要系数；g 为顶事件的发生概率；q_i 为 i 基本事件的发生概率；$I_g(i)$ 为 i 基本的概率重要度系数。

根据上式求出图 5.1-14 故障树中各基本事件的临界重要度系数如下：

$CI_g(8) = 5.21 \times 10^{-1}$

$CI_g(12) = CI_g(14) = 5.21 \times 10^{-1}$

$CI_g(10) = CI_g(11) = 2.6 \times 10^{-1}$

$CI_g(18) = CI_g(19) = 2.37 \times 10^{-1}$

$CI_g(13) = 4.74 \times 10^{-2}$

$CI_g(17) = 4.74 \times 10^{-3}$

$CI_g(2) = CI_g(5) = 2.37 \times 10^{-11}$

因此，各基本事件的临界重要度顺序为：

$CI_g(8) > CI_g(12) = CI_g(14) > CI_g(10) = CI_g(11) > CI_g(18) = CI_g(19) > CI_g(13) > CI_g(17) > CI_g(2) = CI_g(5)$

应用于以下几个方面：

① 改善系统设计；

② 确定系统需要监测的部件或元部件；

③ 制订系统故障（或失效）诊断时的参考，在进行系统的失效分析时，应该把分析的重点放在重要度大的零、部件（或基本事件）上。

6.5 失效模拟和加速失效模式模拟试验方法

失效模拟（故障再现）是事故检查中经常采用的一种分析和验证方法。故障再现可以验证现场调查和残骸分析中所得出的事故直接原因，可以在残骸不全、证据不充分的情况下，提供事故的可能原因；可以解决残骸分析中的某些疑点，排除某些现象；还可以显示失效（故障）的发展过程、残骸的破坏顺序等。因此，为了查清失效的直接原因，在整个失效分析中，往往需要进行多次失效模拟工作。

失效模拟的方法有模拟试验和调查同类机械设备或系统两个方面。

模拟试验是根据现场调查和残骸分析的情况，在机械设备发生失效的实际工况条件下，使其再次发生同样的失效模式，然后根据试验的结果分析其失效原因。

由于设计、制造、使用、维护、修理等方面的问题而造成的失效，都有一定的规律，某些同型号、同批次的机械设备，在基本相同的条件下使用，可能发生同类的失效模式，因此，用调查一定数量的同类机械设备的办法，也可以间接地达到失效模拟（故障再现）的目的。

失效模拟（故障再现）有时对失效原因的判断起着十分关键的作用，例如1954年1月一架“慧星”喷气旅客机起飞20~30 min后，爬高到10 000 m，突然失去联系。飞机在空中爆炸、坠入海中，旅客、乘务员全部死亡。事故发生后全部“慧星”飞机停飞。但是由于残骸还未全部捞起来，事故的原因并未找到，只是对飞机上50多处有怀疑的地方进行了改装。同年4月刚恢复飞行不久，又一架“慧星”在同样的情况下，发生了同样的事故。检查员分析两次事故可能是同一原因，因为两次都是在起飞20~30 min后，正好爬高到航线上的最大高度时在空中爆炸的。第一次事故后飞机上已进行改装了的弱点可以排除在事故原因之外，于是，剩下一个最大的可能原因，即增压座舱发生疲劳破坏。考虑到这两架飞机的使用寿命都不长，是不是可能发生疲劳破坏还不能肯定，于是，采取失效模拟的办法来证实事故发生的原因。用一个大型水柜把整个机身都装在里面，只有两个机翼伸在外面，机翼上加重复负载，使机翼产生的上下方向弹性变形，从而带动机身产生向内向外的弹性变形。试验一直进行到机身上一面积达19 m^{2}的构件被压破为止，这种破裂如在空中发生，必然导致飞机的解体，观察疲劳裂纹是从窗户型材的角部开始发展的。试验中发生疲劳破坏时所加负载的次数大致相当于发生事故的飞机的使用时间。与此同时，第一架飞机的残骸分析还在继续进行。从一块尾翼外蒙皮上发现印有报纸上的文字和一个硬币的印迹。这也证明座舱爆炸的设想是正确的，因为只有座舱首先在空中爆炸，客舱里的东西才有可能被气流吸出来，撞在尾翼上。以后，在座舱顶部安装无线电罗盘的窗户型材上发现了与试验相类似的疲劳裂纹。从此所有喷气旅客机都在设计上采取了防止增压座舱发生疲劳破坏的措施，1958年以后，再没有发生喷气客机的此类事故。

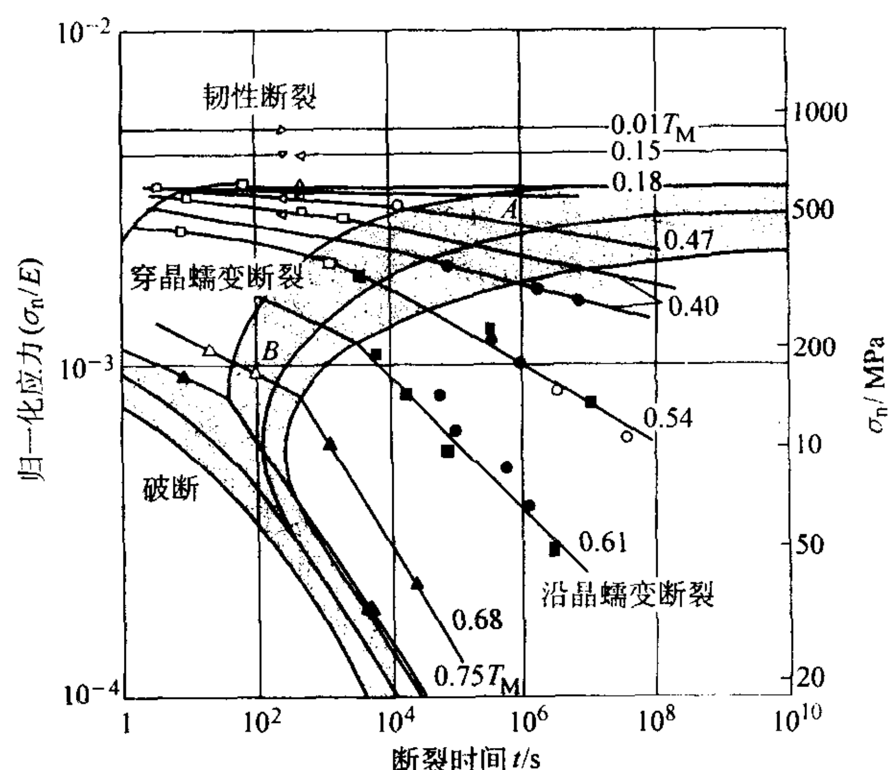


图5.1-15 蒙乃尔-400（Monel 400）合金的蠕变机制图

为了节省失效模拟的时间，有时可以进行加速失效模拟试验。加速失效模拟试验与加速寿命试验不同，前者的目的是为了验证失效的过程、影响因素和机理，而后者的目的则是为了解决可靠性试验中的时间和数值的障碍问题。因此，对加速失效模拟试验，必须在失效机理不变的情况下，试验才有价值，而对加速寿命试验，只要保证退化或失效分布模型具有同一性或规律性即可。例如，对于蠕变断裂的加速失效模拟试验，只能在同一机制区域内进行。

如图5.1-15从A点到B点，可使断裂时间从10^6 s缩短到10^2 s，而对蠕变断裂的加速寿命试验，可以跨跃不同的机制区域，只要在加速状态和非加速状态（标准状态或实际使用状态）的断裂时间之间有规律可循即可。

6.6 失效的事后处理及失效预防

失效的事后处理及失效预防是失效分析的目的和归宿，它涉及到各种技术方法、管理制度和法规标准。在技术方法方面，它包括失效补救、状况监测和控制、剩余寿命估算、维修更换及各种强化技术等等，下面就分别进行扼要地讨论。

(1) 失效补救

失效补救是指正在运行中的同类机械（系统）的处理问题，以防止同类失效事故的再次重复发生。失效补救措施要强调稳、准和及时。例如××发动机涡轮轴断裂，涉及近万台发动机的检修问题，排除了××起事故隐患。失效补救措施首先是向使用部门反馈，但是向设计部门或制造部门反馈并采取必要的根治措施也是必不可少的，因为只有这样才能杜绝同类失效事故的不断发生。

对正在运行中的同类机械（系统）进行检修，一般要求采用无损检测方法，并且最好在生产现场甚至在运行过程当中“动态”的进行。要求针对不同的失效原因、部位，采取不同的、行之有效的无损检测方法或专门设计特殊的工具夹。

(2) 状况监测和控制

状况监测和控制是对正在运行中设备或系统的工作状况进行监视、测试和控制，以便定量地掌握设备的状态（设备所受的应力、设备的强度和性能、以及设备功能的退化和故障等等）、预测设备的可靠性和剩余寿命。如果存在异常，则对其原因、部位、危险程度等进行识别和评价，并决定采用应急的修正措施等等。

状况监测和控制的参量应该选择设备（或系统）的退化特性值。特性值应该是那些对设备或系统退化敏感的、对失

效（或退化）有预测能力和容易观测的参数。它可以是应力、温度、气压、电参量、振动参量、声参量、污染参量、性能参量……，等等。近年来，又采用了电视监控系统。多参量监控系统和电子计算机监控系统，进一步提高了状况监控的直观性和可靠性。

为了加深对状况监测和控制的认识，下面简要介绍声发射监控技术、振动监控技术和断裂的控制技术。

1）声发射监控技术　声发射技术是根据结构内部发出的应力波判断结构内部损伤程度的一种新的无损检测方法。声发射技术与超声、X射线等常规无损检测方法主要区别在于前者是一种动态的方法。它能连续监控结构内部损伤的全过程。由于这一主要特点，声发射技术已在原子能、航空、航天、冶金材料、地震、地质、石油、化学、电力、矿业和建筑等工业领域得到了广泛的应用。

众所周知，材料或结构受外力或内力作用产生变形或断裂，以弹性波形释放出应变能的现象称为声发射。除极少数材料外，金属和非金属材料在一定条件下都有声发射现象，所以，声发射检测几乎不受材料的限制。进一步试验证实，产生突发型声发射信号，不仅与材料的断理韧性值有关，还与材料的屈服强度、断裂机制和裂纹尺寸等因素有关。

由于材料变形、裂纹扩展等的不可逆性质，声发射也有不可逆性。因此，要进行声发射检测，必须知道材料的受力历史，或者在构件第一次受力时就进行检测。

利用多通道声发射装置，可以确定缺陷所在位置。声发射检测的这一特点对大型结构如锅炉等的检测特别方便。

但是，声发射检测到的是一些电信号，根据这些电信号来解释结构内部的缺陷变化往往比较复杂，不仅需要丰富的经验，有时还需要用其他无损检测方法加以验证。另外，声发射检测常常受环境的噪声干扰，虽然在声发射技术中，已有多种排除噪声的方法，但在某些情况下，还会使声发射技术的推广应用受到限制。

由于声发射技术具有上述一系列特点，近年来有许多科学家和工程技术人员致力于它的发展和应用。特别是声发射技术在监视疲劳、应力腐蚀、氢脆和断裂分析及在监控压力容器安全性方面的应用。

2）振动监控技术　当机械的内部发生异常时，一般情况下都会随之出现振动大小及其性质的变化。所以根据对振动的测量和分析，不用停机和解体分析，就可对机械的退化和失效的特征有所了解。这就是为什么在状况监控和现场设备诊断中，广泛利用机械振动法的原因。

在进行监控或诊断时，虽然可以在振动计上对测出的振动波形直接读取，但是，为了正确了解设备的状态，一般要有数据记录器等波形记录下来，然后用各种信号处理技术加以分析，得出需要的、具有典型性的离散值。

振动监控或设备诊断的振动法一般分简易诊断和精密诊断两种。设备的简易诊断主要是为判断设备的老化状态、危险程度、重要程度等；而设备的精密诊断则是经简易诊断判断存在异常而进一步查明原因时所用的方法。

简易的判断标准一般有三种：

① 绝对判断标准——在同一部位（如在轴承上）测定的值与“判断标准”相比较，判断结果为良好/注意/不良；

② 相对判断标准——对同一部位定期测定，按时间先后进行比较，将正常情况的值定为初始值，根据实测值达到初始值的倍数进行判断；

③ 类比判断标准——有数台机型相同的机械时，按相同条件逐台进行测定，经过相互比较作出判断。

一般情况下，在现场最便于使用的是绝对判断标准。

精密的诊断要对振动的频率、相位、振幅……，时间等参数进行分析，必须最大限度地测定异常振动的各种性质，才有可能对异常的原因作出正确的判断。

(3) 断裂分析（评定、控制）图及其应用

通过断裂（失效）分析，确定断裂（失效）的特征参量（参数），建立其断裂分析（评定、控制）图，达到控制断裂（失效）的目的，这是工程上常用的有效方法之一。

根据不同的断裂（失效）特征参数，可以将其分为断裂分析图、比值分析图（RAD）、失效评定图、失效区域图、蠕变断裂机制图和疲劳机制图等。下面就几种断裂（失效）评定图给以扼要的介绍：

1）比值分析图（RAD）　金属材料的断裂与其屈服强度之间的联系可以通过比值分析图RAD方法来建立，如图5.1-16所示。比值分析图是以动态撕裂能量DTE（采用能保证产生完全平面应变状态所需厚度的较深缺口的试样冲断所得的能量）为纵坐标、以其屈服强度为横坐标、以质量最佳的钢为“上工艺极限曲线”、以质量最劣的钢为“下工艺极限曲线”所组成的图形。这一图形被不同的K_{Ic}/σ_{ys}，（K_{Ic}为材料的断裂韧性；σ_{ys}为材料的屈服强度）直线分割成不同的区域。当$K_{Ic}/\sigma_{ys}\geq 1.0$时，为塑性断裂状态；$0.65<K_{Ic}/\sigma_{ys}\leq 1.0$时，为弹-塑性断裂；$K_{Ic}/\sigma_{ys}<0.63$时，为弹性断裂（平面应变断裂）。由此可以看出，为了保证材料为塑性断裂必须要使材料强度、韧性、冶金质量及板材厚度之间合理的配合。

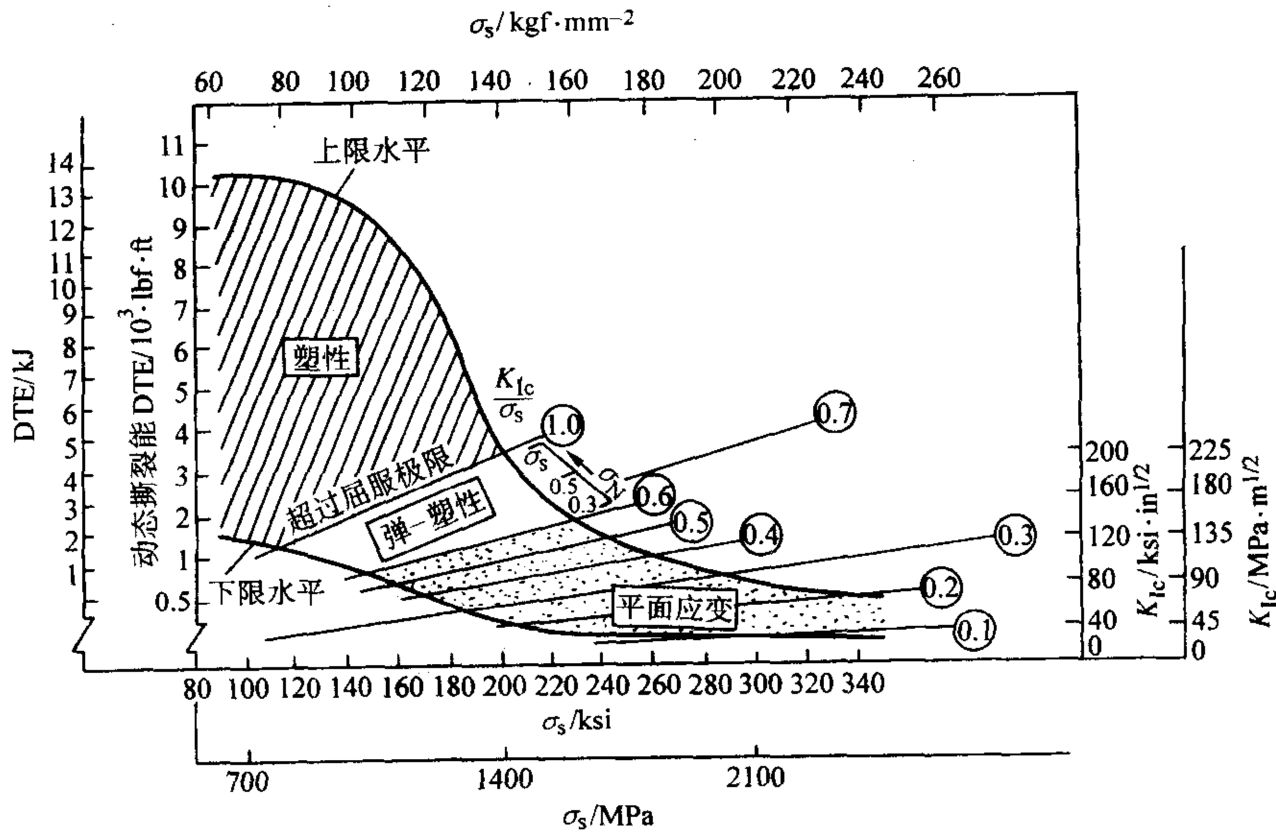


图 5.1-16　比值分析图

比值分析图表明断裂状态和屈服强度（金属质量的函数）之间的联系。截面为 25.4 mm（1 in）的断裂状态，用适当的比值线加以确定。应该指出的是，对于不同的合金和不同的厚度的材料，虽然仍有塑性、弹－塑性和弹性三个区域，但是有时它们的区域界线的 K_{Ic}/σ_{ys} 比值会有变化。

比值分析图不仅可用于防断裂设计、冶金质量的控制，而且可以用于断裂参数的控制和分析。

2）失效评定图（R6图）　失效评定图是20世纪80年代评估含裂纹构件安全裕度的一种新方法。它是以"双判据"法为基础发展起来的。失效定图称为R6图。

众所周知，在塑性状态，失稳的控制参量是极限载荷 P_0，而在脆性（弹性）状态，断裂控制参量则是断裂韧性 K_{Ic}，这是两种截然不同的失效机制。那么，在弹－塑性状态，其失效的控制参量应该是一种弹－塑性的综合参量。失效评定图就是表示上述两种截然不同的失效机制间的转换曲线或内插曲线，因此它适用于完全弹性脆断到完全塑性失稳失效的整个范围。失效评定图合理地和综合地反映裂纹体几何、载荷类型和材料塑性流动性能的影响，迅速判断含裂纹体几何、载荷类型和材料塑性流动性能的影响，迅速判断含裂纹构件是否安全的一种方便的分析方法。

众所周知，在塑性或弹塑性状态，其断裂的控制参量为 J 积分。在平衡的条件下，根据裂纹"扩展力"和"阻力"相等的原则，经过推导最后可以得到

$$\frac{L_r^2}{\frac{8}{x^2}\ln\left[\sec\left(\frac{x}{2}\times L_r\right)\right]}=K_r^2 \qquad (5.1\text{-}13)$$

式中，$K_r=K_I/K_{Ic}$（K_I 为裂纹附近的应力强度因子，K_{Ic} 为材料的断裂韧性）；$L_r=P/P_0$（P 为外加载荷，P_0 为塑性失效、屈服、断裂或失稳的极限载荷）。

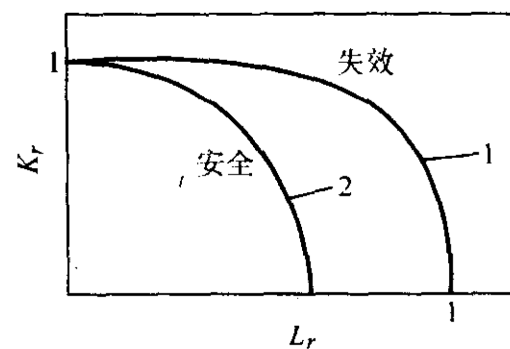


图 5.1-17　失效评定图

将上述方程按 K_r 为纵坐标，L_r 为横坐标作图，即得失效评定图（图 5.1-17）图中曲线为失效评定曲线（曲线1）。由坐标轴 K_r、L_r 及曲线围成的区域为完全区，失效评定曲线之外的区域则为不安全区。因此，如果要对具体构件进行安全评定，只要对给定的裂纹构件、载荷类型和材料性质，求出 K_r 和 L_r 即可。

从图 5.1-17 可以看出，R6 曲线是一条不规则的曲线，为了使评定更为简便和安全，可以将失效评定曲线取为四分之一圆（如图 5.1-17 中的曲线2）。

同样，失效评定图也可以用于质量控制、安全设计和失效分析。

3）蠕变断裂机制图　可以通过解各种变形机制的本构方程，并且分别分析各个变形机制相互依赖或相互独立的关系，确定在哪一个试验范围内，支配材料性能的蠕变和断裂的机制。Ashby 和他的合作者，在以归一化应力为纵坐标、相对温度为横坐标，标出各种蠕变和断裂机制的区域，这就是蠕变断裂机制图。

研究表明，由于不同结构材料的性能不同，它们的蠕变断裂机制图也各异。Ashby 提出，共有五类蠕变断裂机制图（图 5.1-18）和如下的断裂机制：解理1或沿晶断裂1（预裂

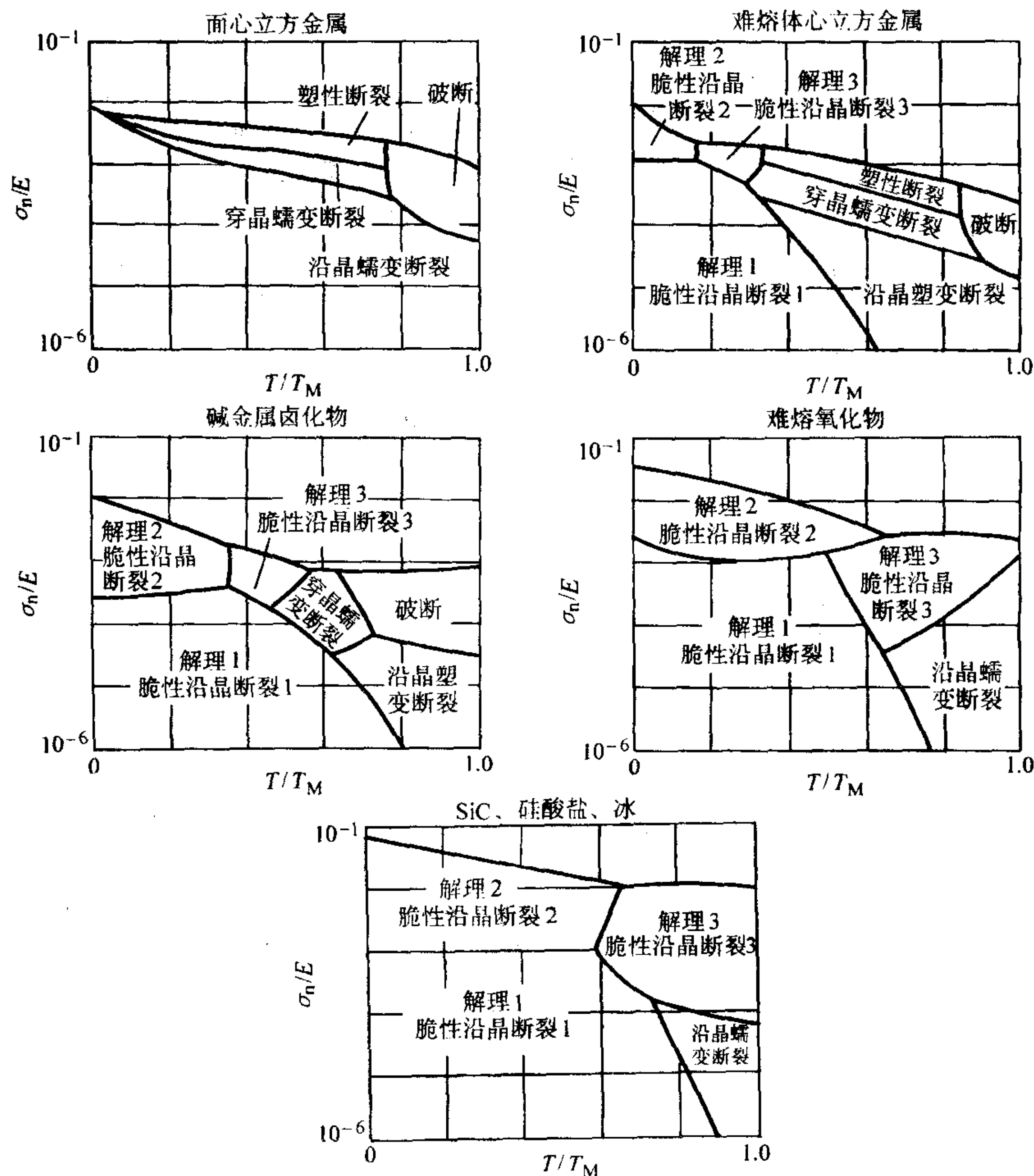


图 5.1-18　五种蠕变断裂机制示意图

纹扩展)、解理2或沿晶断裂2（滑移或孪晶诱生裂纹扩展)、解理3或沿晶断裂3（大塑性变形诱发的脆性断裂)、沿晶的蠕变断裂、穿晶的蠕变断裂和破断（动态再结晶）等。

蠕变断裂机制图是很有实际价值的，只要知道三个主要参数——归一化应力、相对温度和断裂机制——中的任何两个，就可以从这些图中求得第三个参数。它有助于蠕变断裂的安全设计和失效分析。

4）疲劳机制图 疲劳断裂机制图如图5.1-19所示，其纵坐标为σ_s/K_{Ic}，它反映了材料的相对塑性，横坐标为应力强度因子K_{max}，它反映疲劳裂纹扩展过程中裂纹顶端的应力强度的大小。图5.1-19被几条曲线分成若干区域，每个区域都为一种主要的断裂机制所控制。例如，在高的相对塑性（σ_s/K_{Ic}的值小)、应力强度小（K_{max}值小）的情况下，即图的左下角的区域，疲劳断裂的机制是穿晶条带机制（$T+S$)，随着K_{max}的提高，其断裂机制依此转为“纯”的韧窝机制（D)、条带+韧窝机制（$S+D$)；在中等塑性的情况下，随着K_{max}的提高，其断裂机制由沿晶和穿晶的疲劳条带（$I+T+S$）向条带+韧窝（$S+D$）转变，最后也转变为韧窝（D)；而在低塑性的情况下，则先由沿晶断裂，转变为沿晶+解理+韧窝（$I+C+D$)，最后转变为韧窝断裂（D）等。

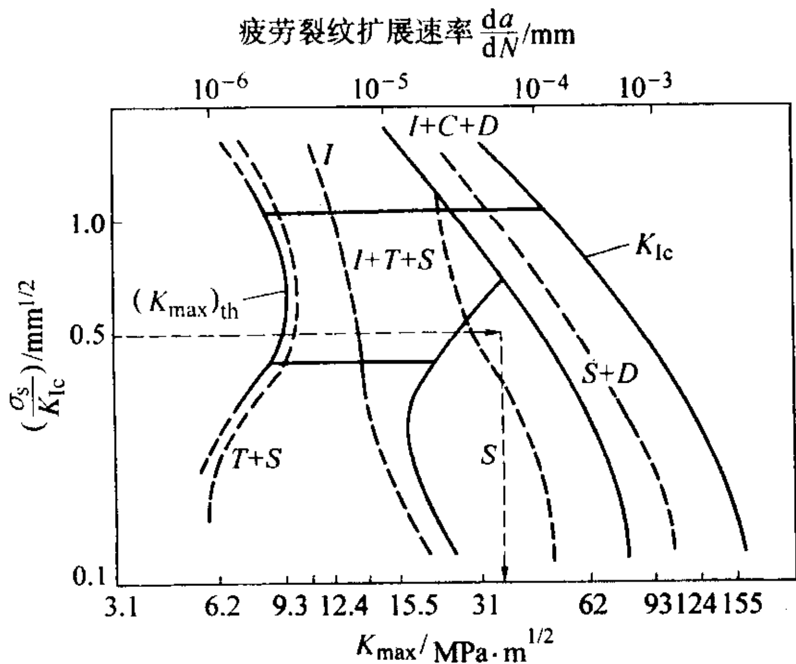


图5.1-19 低合金在室温大气条件下的疲劳机制图
T—穿晶断裂；I—沿晶断裂；D—韧窝断裂；
C—解理或准解理断裂；S—疲劳条带断裂

应该指出的是，图5.1-19是在低合金结构钢、室温、大气条件下得到的，随着不同合金疲劳环境条件的变化，其疲劳机制图的形状也会发生变化。

在疲劳机制图上叠加等扩展速度da/dN曲线，就成为完整的疲劳机制图了。

和蠕变断裂机制图一样，只要知道三个主要参数σ_s/K_{Ic}、K_{max}和疲劳断裂机制中的任何两个，就可以从图中求得第三个参数。例如，如果确定了疲劳断口的电子显微特征和受力情况就可以通过疲劳机制图分析其材质的情况，并为改进材料和工艺提供线索。

5）失效区域图 为了判断机械零件在什么条件下易产生什么样的失效模式，不少资料提出了各种失效区域图。图5.1-20为一种轮齿失效区域图。图中在相应的载荷和节圆速度下，表示出了轮齿失效形式的区域。由于这种图随具体的工作条件和齿轮的本身的情况而异，因此，它不能直接用于设计或判断齿轮的工作能力，但它能表明失效的出现条件和彼此之间的关系。

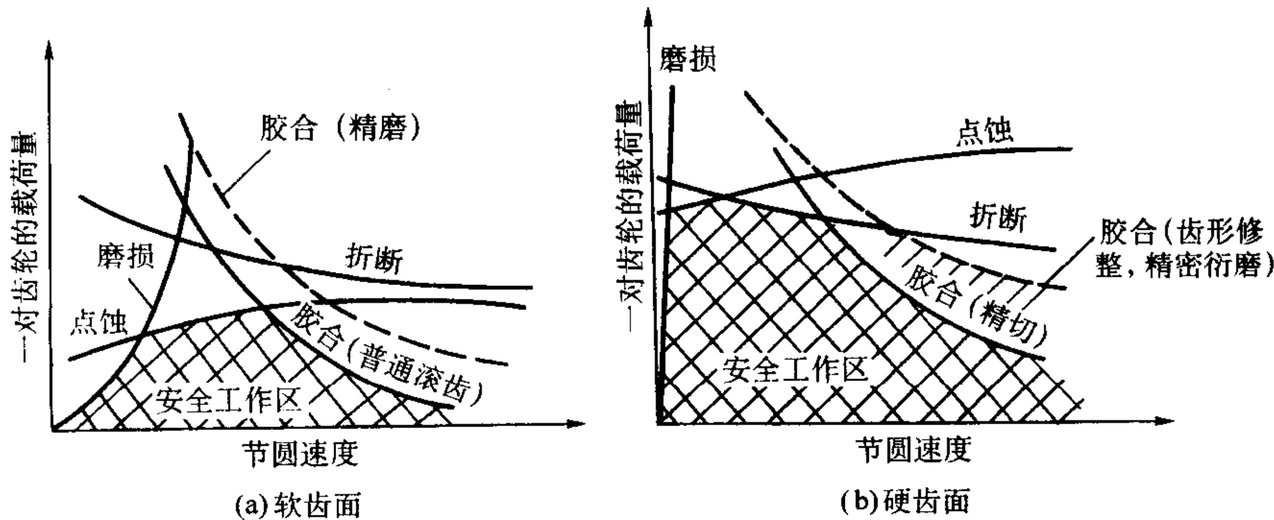


图5.1-20 轮齿的失效区域图

(4) 剩余寿命估算

安全剩余寿命估算是失效事后处理及失效预防中的重要工作之一。按估算的对象不同可分为零件的寿命估算和系统的寿命估算；按估算的方法可分为以失效机理为基础的寿命估算、以失效模式为基础的寿命估算和以可靠性工程为基础的寿命估算等。

1）以失效机理为基础的寿命估算 关键零部件的失效能代表整个系统失效，已知该零部件的失效机理的条件下，从该失效机理的动力学特性来估算其剩余寿命。这种方法是材料工作者和力学工作者常用的方法之一。

当零件的失效是单一的失效机制，或由一种失效机制起主要控制作用时，其剩余寿命估算方法、应力腐蚀寿命估算方法、蠕变寿命估算方法等，已在有关专著中涉及，这里不再赘述。

2）以失效模式为基础的寿命估算 根据当前的归纳分析，失效模式有多种，如应力-强度模型、反应论模型、最弱环模型、串联模型、绳子模型或并联模型等估算方法简述如下。

① 应力-强度模型及其寿命的估算方法 当应力超过产品的强度时，失效就会发生。这是一个材料力学的模型。如果我们引进“退化”和“概率”的概念，就成为一个完整的应力-强度模型了。

例如最初应力与强度是留有充分的安全余量的，经过一定时间后，应力分布和强度就会交叠，这就发生一定几率的失效（图5.1-21)。

设应力和强度均服从于正态分布，即应力的概率密度为

$$f(x_L)=\frac{A}{\sigma_L\sqrt{2\pi}}e^{-\frac{(x_L-\mu_L)^2}{2\sigma_L^2}} \tag{5.1-14}$$

强度的概率密度为

$$f(x_S)=\frac{A}{\sigma_S\sqrt{2\pi}}e^{-\frac{(x_S-\mu_S)^2}{2\sigma_S^2}} \tag{5.1-15}$$

式中，μ_L和μ_S分别为应力强度的数学期望，其中μ_S是时间t的函数，即$\mu_S=\mu_S(t)$，μ_L和μ_S分别为应力和强度

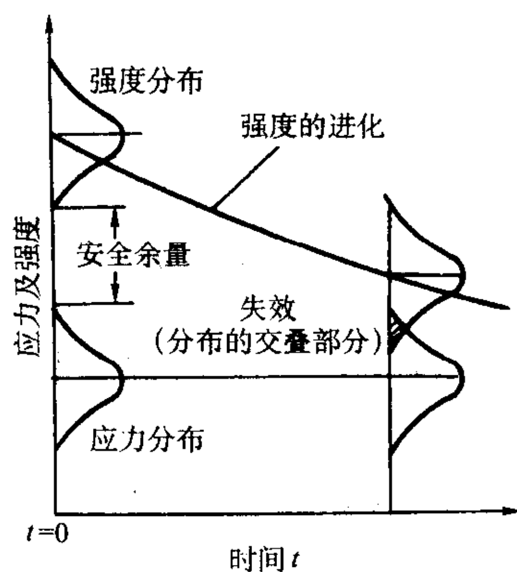


图5.1-21 应力-强度模型

分布的均方差。

并设 $x_S - x_L = x$ 的分布也是正态分布的，则其失效密度函数相当于以 $\mu = \mu_S - \mu_L$ 为均值、以 σ 为其标准偏差的正态分布，即

$$f(x,t) = \frac{\Lambda}{\sqrt{2\pi(\sigma_s^2 + \sigma_L^2)}} e^{-\frac{[x_S - (\mu_S - \mu_L)]^2}{2(\sigma_S^2 + \sigma_L^2)}} \quad (5.1\text{-}16)$$

则其失效概率 ϕ（x）可由上式正态分布积分求得。

如果失效的定义是其累积失效概率 ϕ（x）= 定值，那么就可以将不同时间的 μ_S（t）值代入上式，查标准正态分布表，求得 $\phi_S - t$ 的数值对应关系，最后确定其寿命。

② 反应论模型及其寿命的估算法　失效的过程需要超过一个能垒，越过此能垒（称为激活能 ΔE）的频数是按一定概率发生时，即服从于所谓玻耳兹曼分布，此方程称为阿列尼斯关系式［式（5.1-17)］。

$$\frac{\partial x}{\partial t} = K = \Lambda e^{-\Delta ERT} \quad (5.1\text{-}17)$$

式中，x 为特征值或退化量；$\partial x/\partial t$ 为退化速度，以 K 表示；Λ 为频数因子；ΔE 为激活能；R 为气体常数；T 为热力学温度。

对于除温度以外还包括有其他应力 σ 的普遍情况，有

$$\frac{dx}{dt} = k = \Lambda e^{-\Delta ERT} \cdot \sigma^a \quad (5.1\text{-}18)$$

式中，k 为玻耳兹曼常数；a 为材料和温度的常数。

如果认为当特征值 x 一旦达到 x_{sp} 时，则其寿命终止，那么便有

$$\int_{x_0}^{x_{sp}} dx = \int_0^L K dt \quad (5.1\text{-}19)$$

设 K 与时间无关，则寿命 L 为

$$L = (x_{sp} - x_0)/K \quad (5.1\text{-}20)$$

当 $x_s = 0$ 时，并只有“温度”应力的情况下，

$$L = \frac{x_{sp}}{\Lambda} e^{+\Delta ERT}$$

或

$$\ln L = \Lambda + \frac{B}{T} \quad (5.1\text{-}21)$$

式中 $\Lambda = \ln \frac{x_{sp}}{\Lambda}$，$B = \Delta E/K$。

当还有其他应力的情况

$$L = \frac{x_{sp}}{\Lambda} e^{\Delta ERT} \cdot \frac{1}{\sigma^a}$$

或

$$\ln L = \Lambda + \frac{B}{T} - a\ln\sigma \quad (5.1\text{-}22)$$

当 T = 常数时，有

$$\ln L = C - a\ln\sigma \quad (5.1\text{-}23)$$

式中　$C = \Lambda + \frac{B}{T}$

通过上述各式，就可以估算不同前提条件下的寿命。

③ 最弱环模型（或串联模型）及其寿命估算方法　可以认为，最弱环模型是建立在失效率模型基础上的。可以认为，当零件或系统的失效率（或可靠度）达到一定数值时，零件或系统就失效了。在一般情况下，失效率 λ（t）与可靠度 R（t）之间有如下关系

$$R(t) = e^{-\int_0^t \lambda(t)dt} \quad (5.1\text{-}24)$$

在偶然失效期，其失效率 λ 与 t 无关，故有

$$R(t) = e^{-\lambda t} \quad (5.1\text{-}25)$$

如果认为 R（t）= R_{sp} 时，零件或系统就失效了，那么，其寿命 L 为

$$L = \frac{1}{\lambda}\ln\left(\frac{1}{R_{sp}}\right) \quad (5.1\text{-}26)$$

对一个串联系统来说，系统的失效率等于构成该系统的各独立部件失效率之和，即

$$\lambda_s(t) = \lambda_1(t) + \lambda_2(t) + \cdots + \lambda_n(t) = \sum_{i=1}^{n} \lambda_i(t) \quad (5.1\text{-}27)$$

如果在这些独立的部件中，第 i 个部件是最弱环，即它的失效率 λ_i（t）最大，因此，系统的失效率主要受该部件的失效率 λ_i（t）控制，即有

$$\lambda_s(t) \approx \lambda_i(t) \quad (5.1\text{-}28)$$

这就是最弱环模型（或称串联模型）。

串联模型不仅适用于系统的分析，而且也适用于失效机理的分析。因为，元件、材料若存在若干相互独立的失效机理，而其中任何一个机理的进行都可能导致元件、材料的失效，所以，元件、材料的失效率可以用各个机理的失效率之和来表示；即

$$\lambda_s = \sum_{i=1}^{n} \lambda_{机理i} = \lambda_{机理1} + \lambda_{机理2} + \cdots + \lambda_{机理n} \quad (5.1\text{-}29)$$

如果认为寿命 L 等于平均失效间隔 MTBF（Mean Time Between Failure)，即

$$L = MTBF = \frac{1}{\lambda} \quad (5.1\text{-}30)$$

那么，对串联模型就有

$$L = \frac{1}{\lambda_s(t)} = \frac{1}{\sum_{i=1}^{n} \lambda_i(t)} \quad (5.1\text{-}31)$$

或

$$L = \frac{1}{\lambda_s} = \frac{1}{\sum_{i=1}^{n} \lambda_{机理i}}$$

$$\lambda_1(t) = \lambda_2(t) = \cdots = \lambda_n(t) = \lambda(t) \quad (5.1\text{-}32)$$

或

$$\lambda_{机理1} = \lambda_{机理2} = \cdots = \lambda_{机理}$$

$$L = (MTBF)_s = \frac{1}{n} \cdot \frac{1}{\lambda(t)} = \frac{MTBF}{n} \quad (5.1\text{-}33)$$

式中，$(MTBF)_s$ 为整个系统的平均失效间隔；$MTBF$ 为构成该系统的各独立部件的平均失效间隔。

④ 并联模型及其寿命估算方法　设备、装置、系统的组成要素中，只有当所有的要素同时失效时，该设备、装置或系统才发生失效，这叫做并联模型。这就是说，并联系统的寿命取决于系统中最可靠的、寿命最长的结构要素，即

$$L_s \approx L_{最长} = \frac{1}{\lambda_{最小}} \quad (5.1\text{-}34)$$

式中，L_s 为系统的寿命；$L_{最长}$ 为所有要素中寿命最长要素的寿命；$\lambda_{最小}$ 为该寿命最长要素的失效率。

如果构成关联系统的所有结构要素的失效率均相等，即其独立结构要素的寿命均相同（以 L 结构表示），那么该系统寿命 L_s 为

$$L_s = L_{结构}\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}\right) \quad (5.1\text{-}35)$$

3）以可靠性工程为基础的寿命估算　可靠性寿命估算是可靠性工程的重要组成部分，它是根据产品的试验数据和现场数据，利用数理统计的方法，并结合实际使用及各种相关因素给出产品的寿命的过程。做好这项工作的关键是保证样品数据的来源，选取合适的寿命指标及正确、实用的统计方法。

可靠性寿命评估程序一般可以用如下框图表示（图5.1-22）。

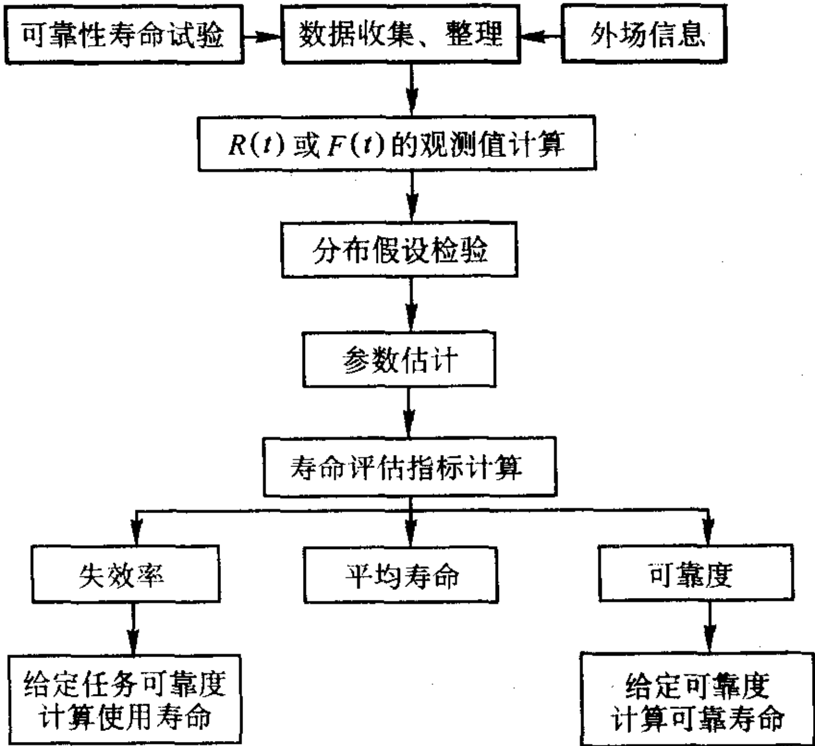


图 5.1-22　可靠性寿命评估的程序框图

下面结合案例具体地说明可靠性寿命的评估程序和方法。

由外场统计了某设备 247 件样品，得到数据如表 5.1-5 所示，求该设备要求可靠度为 0.951 的可靠寿命。

表 5.1-5　某设备的可靠性试验数据

$t_{i-1} \sim t$ 使用时间 t/h	Δr (t_i) 失效数	Δk (t_i) 删除数
约 700	0	27
700 ~ 800	0	18
800 ~ 900	0	19
900 ~ 1 000	5	29
1 000 ~ 1 100	6	44
1 100 ~ 1 200	5	18
1 200 ~ 1 300	5	27
1 300 ~ 1 400	6	20
1 400 ~ 1 500	6	12

解：① 计算设备的可靠度函数

参加试验的样品数 $n=247$，按残存比率法计算该设备的可靠度函数如表 5.1-6 所示。

表 5.1-6　该设备的可靠度函数

t_i	$\Delta r(t_i)$	$\Delta k(t_i)$	$n_s(t_i)$	$X(t_i)$	$S(t_i)$	$R(t_i)$
700	0	27	220	0	1	1
800	0	18	202	0	1	1
900	0	19	183	0	1	1
1 000	5	29	149	0.027 3	0.972 7	0.972 7
1 100	6	44	99	0.040 27	0.959 7	0.933 5
1 200	5	13	76	0.050 5	0.949 5	0.886 4
1 300	5	27	44	0.065 79	0.934 2	0.828 0
1 400	6	20	18	0.136 4	0.863 6	0.715 1
1 500	6	12	0	0.333 3	0.666 7	0.476 8

计算公式

$$X(t_i) = \frac{\Delta r(t_i)}{n_s(t_i-1)}$$

$$n_s(t_i) = n - \sum_{i=1}^{i}\Delta k(t_i) - \sum_{i=1}^{i}\Delta r(t_i) \quad (5.1\text{-}36)$$

$$S(t_i) = 1 - X(t_i)$$

$$R(t_i) = \prod_{i=1}^{i} S(t_i)$$

根据表 5.1-6 的数据，计算该设备时间区段的平均失效率。

$$\bar{\lambda}(t_{i-1},t_i) = \frac{\Delta r(t_i)}{n_s \cdot t_{i-1} \cdot \Delta t_i} \quad (5.1\text{-}37)$$

设备在 $t_{i-1} \sim t_1$ 区间的平均失效率可以算出。计算结果见表 5.1-7。

表 5.1-7　计算结果列表

$t_{i-1} \sim t_i$/h	Δr (t_i)	n_s (t_i)	$\bar{\lambda}$ (t_{i-1}, t_i)
900 ~ 1 000	5	183	2.732×10^{-4}
1 000 ~ 1 100	6	149	4.027×10^{-4}
1 100 ~ 1 200	5	99	5.05×10^{-4}
1 300 ~ 1 400	5	76	6.579×10^{-4}
1 400 ~ 1 500	6	44	13.6×10^{-4}
1 500 ~ 1 600	6	18	33.33×10^{-4}

将表 5.1-6、表 5.1-7 的计算结果绘成曲线，如图 5.1-23、图 5.1-24。

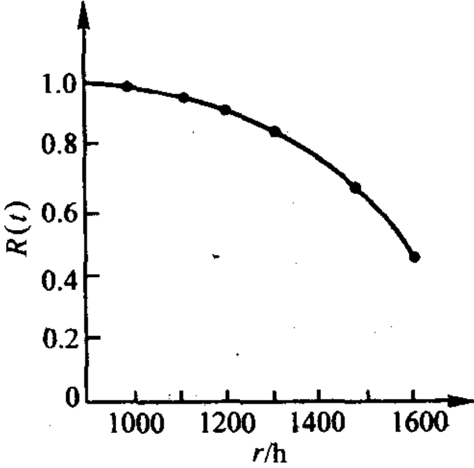


图 5.1-23　设备的可靠度曲线

② 分布的假设检验及参数估计　由图 5.1-24 设备的失效率随时间的增加而增加，属 IFR 型。初步假设其寿命为正态分布。

用最小二乘法估计正态分布的参数 σ 和 μ。根据表 5.1-6 的数据计算

$$F(t) = 1 - R(t) \quad (5.1\text{-}38)$$

$$F(t) = \Phi\left(\frac{t-\mu}{\sigma}\right) = \Phi(Z) \quad (5.1\text{-}39)$$

计算结果见表 5.1-8。

将表中数据 Z_i 和 t_i 输入计算器，进行回归计算，得正态分布之参数：

平均寿命 $\mu = 16\ 143$ h

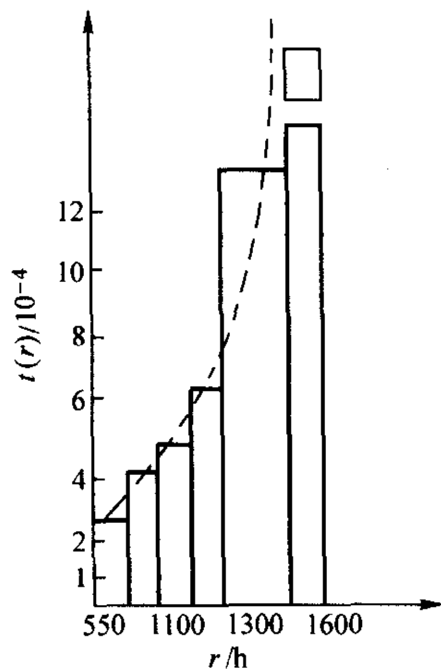


图 5.1-24 设备的失效率曲线

寿命均方差 $\sigma=326.2$ h

相关系数 $r=0.986$

根据所估得的参数计算的可靠度函数的理论值也列入表5.1-8中。

表 5.1-8 可靠度函数的理论值

时间 t_i/h	$R(t_i)$ 实测	$F(t_i)$ 实测	Z_i	$\frac{t_i-\mu}{\sigma}$	$F(t_i)_{理论}$	$R(t_i)_{理论}$
100	1			-4.64	0.000 001 742	0.999 998 2
200	1			-4.34	0.000 007 124	0.999 992 9
300	1			-4.03	0.000 027 89	0.999 972 11
400	1			-3.72	0.000 099 51	0.999 900 39
500	1			-3.42	0.000 313 1	0.999 686 9
600	1			-3.11	0.000 935 4	0.999 064 6
700	1			-2.80	0.002 555	0.997 445
800	1			-2.50	0.006 21	0.993 79
900	1			-2.10	0.014 26	0.985 74
1 000	0.972 7	0.027 3	-1.922	-1.88	0.030 05	0.969 95
1 100	0.933 5	0.066 5	-1.501	-1.58	0.057 05	0.942 95
1 200	0.886 4	0.113 6	-1.207 5	-1.27	0.102	0.898
1 300	0.828	0.172	0.946 4	-0.96	0.168 5	0.813 5
1 400	0.715 1	0.284 9	-0.568 2	-0.35	0.363 2	0.636 8
1 500	0.476 8	0.523 2	0.058 25	-0.044	0.484	0.516

由最小二乘法计算得到的相关系数 $r=0.986$，说明实测的可靠度曲线和正态分布的可靠度曲线拟合得很好，这也相当于检验了最初的假设——设备失效为正态分布。

③ 设备可靠寿命的计算

该设备要求的可靠度 $R'=0.95$，即

$$1-\Phi\left(\frac{t_r-\mu}{\sigma}\right)=0.95$$

$$\Phi\left(\frac{t_r-1\ 614.3}{326.2}\right)=0.05$$

查正态分布表得

$$\frac{t_r-1\ 614.3}{326.2}=-1.64$$

所以当要求设备的可靠度为0.95时，其可靠寿命为1 079.3 h。

(5) 维修、更换及有关的修复技术

为了保持可修复系统、设备、元器件的可靠性而进行的工作叫做维修。表示可修复系统、设备、元器件维修的难易程度或特性叫做维修性。而其在规定的条件下和在规定的时间内完成维修的概率则称之为维修度。

维修方式分为产品失效后进行维修的事后维修和在失效前就进行维修的预防维修。因为事后维修（又叫失效的事后处理）是在发生失效后再进行的维修，当系统庞大时，这种被动情况对成本和安全方面十分不利，因而这种预防维修更受到重视。

为了缩短事后维修的维修时间，为了设计出最适用的预防维修的检查方式，就必须充分利用有关的失效信息。在一般情况下，事后维修应在失效事故发生后立即进行，而预防维修的对象是其失效率曲线为增加型的产品，并且应该在观察对象的恶化趋势的同时进行有针对性的维修，这里就有失效判断和检测问题，因此，预防维修又叫预知维修，也叫状态监测维修。图5.1-25表示定期的状态监测与状态监测维修之间的关系。

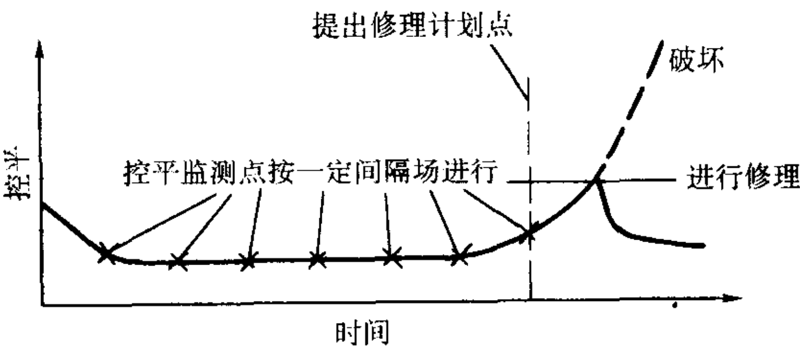


图 5.1-25 定期状态监测与状态监测维修之间的关系

可想而知，系统越大、自动化程度越高，检测故障就越显得重要。为此，必须掌握故障的发生趋势，并研制能判断各种故障趋势的、能处理大量数据的辅助系统，也就是说，应该研究包括非破坏性检查技术在内的失效分析方法，进一步采用微小型元件，研制良好的检测元件（传感器），研究失效模式分析与综合判断的方法，以及研制有效的显示方式等，并把这样的诊断装置放在大系统内，起到动态监控的作用。

维修替换（或维修更换）的方法可为修复、个别替换和全部替换（对工作了一定的时间的、但尚未发生故障的零件进行全部替换）等几种。

修复技术大致有如下几种：

① 机械修复法　包括机加工修复、锉修和喷丸等。

② 修复电镀法　包括一般电镀和刷镀等。

③ 焊接或堆焊修复法。

④ 其他方法，例如热喷涂、烧结金属粉末层等。

根据不同的情况，可以采用不同的修复方法。

7 失效分析预测预防的发展趋势

失效分析预测预防技术的发展趋势可以用图5.1-26来简要表述。

机电装备失效分析预测预防实践和学术方面的重大进展之一，是促进和带动了一门交叉综合分支新兴学科——失效学体系——的形成和发展，从而使失效分析完成了从一门技术逐步提高到一个分支学科的飞跃。这是当代科学技术发展的结果，是我国几代科学工作者为之毕生奋斗的目标，它将对机电装备失效分析预测预防工作产生深远的影响和作用。

失效学是研究机电装备（系统、设备和元器件）的失效分析诊断（简称失效诊断）、失效预测和失效预防的理论、技术和方法及其工程应用的分支学科。它的产生是有其近代科学技术进步的深刻背景的。可以认为，近代材料科学和工

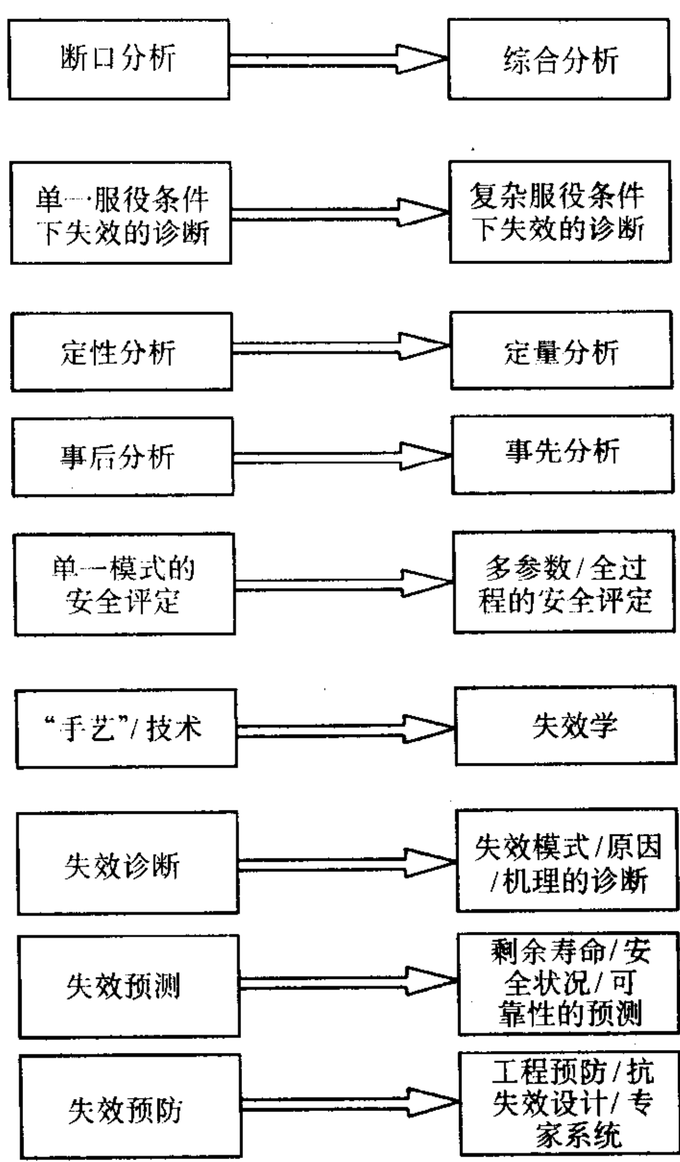


图 5.1-26 失效分析预测预防的发展趋势简图

程、工程力学、疲劳等学科对断裂、腐蚀、磨损及其复合型（或混合型）的失效模式和失效机理的深入研究，积累了相当丰富的创新观点、见解和物理模型，为失效学的建立奠定了理论基础；现代的检测仪器、仪表科学的迅猛发展，以及检测技术的不断提高，特别是断口、裂纹、痕迹分析技术体系的建立、发展和完善，为失效学的发展奠定了技术基础；数理统计学简报完善、模糊数学的突起、可靠性工程的发展应用和电子计算机的广泛普及，为失效学的完善奠定了方法基础。上述这三者的融会贯通，使失效学逐渐建立、发展和完善成为一门相对独立的、综合的新兴学科成为可能。虽然关于失效学的"基本内容"、"内涵和外延"的雏形早在1985年已有人提出，但是就其体系的系统性和完整性，就其内容的深度和广度而言，近年来，又有了很大的发展。我们认为，失效学学术体系应由失效诊断学、失效预测学和失效预防学三个分支体系组成。现在就按这三部分简要叙述如下：

7.1 失效诊断的理论、技术和方法的进展

失效诊断是失效分析的主要任务之一。失效诊断的理论、技术和方法的核心是其思维学、推理法则和方法论。失效诊断中的五条基本原则（即整体观念原则、立体性原则、从现象到本质的原则、动态性原则和两分法原则）和五种具体方法（即相关性方法、抓关键问题的方法、对比的方法、历史的方法和逻辑方法），是对失效诊断过程的原则和方法的概括描述。关于失效诊断思路方法在有关文献中均有所叙述或论述。

失效诊断依其目的要求和内容深度可分为三个层次的诊断：即失效模式的诊断、失效原因的诊断和失效机理的诊断。现分别简述它们的进展。

(1) 失效模式诊断的进展

1) 失效模式是指失效的表现形式，一般可理解为失效的类型。如断裂的一级失效模式有：韧性断裂失效、脆性断裂失效和疲劳断裂失效；非断裂失效的一级失效模式有：腐蚀失效、磨损失效和变形失效。失效模式诊断是失效分析中首当其冲的重要问题，它对整个失效分析和预防决策来说，具有定向的作用。

2) 失效模式常可分为一级失效模式、二级模式、甚至三级失效模式等。在上述一级疲劳断裂模式的基础上，按频率高低可以有高频疲劳失效、低频疲劳失效和混频疲劳失效等二级疲劳断裂失效模式；按应力大小可以有高周疲劳失效、低周疲劳失效和高低周复合疲劳失效等二级疲劳断裂失效模式；按应力来源可以有机械疲劳失效和热疲劳失效等二级疲劳断裂失效模式，在机械疲劳失效中又可以有弯曲疲劳失效、扭转疲劳失效、拉—拉疲劳失效、接触疲劳失效等三级疲劳断裂失效模式；按环境介质可以有疲劳失效、腐蚀疲劳失效、应力腐蚀疲劳失效等二级疲劳断裂失效模式；按宏观疲劳裂纹走向与主应力的关系可以有正断型疲劳失效（如疲劳的第二阶段）、剪切型疲劳失效，如疲劳的第一阶段等二级疲劳断裂失效模式；按微观疲劳裂纹走向与显微组织的关系可以有穿晶疲劳失效、沿晶疲劳失效等二级疲劳断裂失效模式，在穿晶疲劳失效中，又可以有晶格型的疲劳失效和非晶格型的疲劳失效等三级疲劳断裂失效模式等。失效模式的诊断一般要求诊断到二级失效模式，甚至要求诊断到三级失效模式，一般情况下，二级、三级的失效模式的诊断比一级失效模式的诊断要难得多。但二级、三级的失效模式的诊断对失效原因的诊断以及失效机理的诊断是不可缺少的。

3) 断裂失效模式的诊断的依据或判据如下。

① 失效残骸　根据残骸的轨迹、断口的宏观性质、断口的变形顺序等首先寻找首先破坏件（肇事件），然后对首先破坏件（肇事件）的断口性质、裂纹走向、变形情况、痕迹来源、力学性能、显微组织、工艺过程、热处理状态等进行单项和综合的分析，对失效模式进行判断。

② 应力分析　包括应力的来源、性质和大小的估算，特别是对首先破坏件（肇事件）的结构和受力的分析和估算、工况和环境分析，并与断口的定性和定量分析相对照，以及与仍在服役件的调查相结合进行综合分析，对失效模式进行判断。

③ 失效模拟　对主要的失效模式和主要的控制或影响参量在实验室或现场进行模拟试验，并对模拟失效的断口与实际肇事件断口进行对比分析，对其失效模式进行判断。

4) 磨损失效的二级模式一般可分为磨粒磨损失效、黏着磨损失效、疲劳磨损失效、腐蚀磨损失效和微动磨损失效五类。每类又可分为若干个三级磨损失效模式。磨损失效模式的诊断一般是依据①被磨损表面的形貌和亚表面层组织与性能的变化；②磨屑形貌、磨屑成分组织结构的变化；③磨损系统中各参量的关系和变化等。但是由于磨损本身是一种伴随着摩擦的存在而存在的摩擦面材料的逐步损失、迁移或变形的过程，虽然一般认为有摩擦就有磨损，但磨损并不等于磨损失效。由磨损到磨损失效实际存在一个由量变到质变的转化，对于不同的机械，这种转化的分界点是不一样的。因此磨损失效模式的诊断与断裂失效的模式诊断之间一个明显的不同点，也是磨损失效模式诊断的难题之一，是磨损失效的失效点的判定问题。这是磨损失效模式诊断中首先需要解决的问题之一。

5) 腐蚀失效按形态一般可以分为八大类：即均匀腐蚀、电偶腐蚀或双金属腐蚀、缝隙腐蚀、孔蚀、晶间腐蚀、选择性浸出、冲蚀或气蚀、应力腐蚀和氢损伤等。腐蚀失效模式的诊断依据一般是：①腐蚀表面形貌；②被腐蚀材料的成分、组织和性能的变化；③腐蚀产物的成分、组织和结构分析；④腐蚀环境和参量分析，包括气氛、介质、温度、应力、电极电位等电化学性能等。由于腐蚀失效过程的影响因

素很多、随机性较强、实验室研究结果又与工程实际还有较大差距，至今，对于腐蚀失效模式、原因和机理的诊断，基本上是定性的、经验的和统计的方法占有支配的地位。

应该特别强调的是，虽然失效模式的诊断是整个失效分析工作首当其冲的问题，但是它又是一个贯彻始终的，应该根据各方面的实验结果或因素进行综合的、反复的分析判断的问题。特别是对二级或三级的失效模式的诊断更是如此。切不可认为这是一蹴而就的事情。

6）失效模式的诊断可分为定性诊断和定量诊断。一般来说，定性诊断是模式诊断的基础和前提，定量诊断则是模式诊断的深入和发展。

定性诊断的技术和方法主要用于一级失效模式的诊断，而定量诊断技术和方法则主要用于二级或三级失效模式的诊断；定性诊断的技术和方法主要是依据失效残骸的分析，特别是肇事件的宏观断口分析的结果进行判断，而定量诊断的技术和方法则需要根据应力分析和失效模拟的结果进行判断。可以看出，如何根据失效事故的具体情况，合理地选用定性和定量的技术和方法，并且根据单一的和综合的判据，对失效的一级、二级甚至三级模式进行正确的、适时诊断，是一个失效工作者应该认真总结和提高的问题。

7）一般地说，对各种典型的失效模式的诊断，特别是对各种典型的一级失效模式的诊断并不困难。一般只要根据单一的诊断判据（如宏观断口）就可以得出一级失效模式的诊断结论。失效模式诊断的难点在于对过渡的、多因素作用的、非典型的二级、三级失效模式的正确和适时的诊断。例如，单就疲劳断裂失效而言，介于韧性断裂和疲劳断裂之间的、脆性断裂（特别是沿晶脆性断裂）和疲劳断裂之间的低周疲劳断裂失效的诊断；超高强度钢（特别是宏观断口疲劳特征不明显时）的疲劳断裂失效的诊断；铸造合金的疲劳断裂失效的诊断、腐蚀疲劳断裂和疲劳断裂断口之间的区分诊断；腐蚀扩展，包括应力腐蚀扩展和腐蚀疲劳扩展的区分诊断；板材（或板式构件），特别是薄板的低周疲劳诊断等。由于多种原因引起的20多种二级脆性断裂失效模式的区别和正确诊断，还很少有人进行系统的研究。因此，失效分析工作者，如果要不断提高自己的诊断水平，就应该对过渡的、多因素作用的、非典型的二级或三级失效模式的正确和适时诊断的技术和方法做深入和系统的研究。

在失效模式的诊断中，人们常依据自己的经验或根据已有的断口、裂纹、金相图谱，但是已有的多种图谱，特别是断口图谱和案例集，多数是“特征诊断”式的，它们只能说明什么材料、什么状态，在什么环境（包括应力、温度、气氛）下断裂断口的特征形貌。到目前为止，还很少有图谱给出当条件系统变化时特征形貌的规律性变化的知识，例如当材料的力学性能系统变化时、或应力系统变化时、或应力系统变化时、或温度系统变化时、或零件形状系统变化时、或介质参数系统变化时，特征断口形貌有规律的变化的情况。这说明，当前断口分析技术和方法，虽然已经有一些定量分析的研究工作，例如对金属疲劳断口的物理数学模型和它的定量反推分析做了一些有益的探索，这是一个好的开端，但是，从总体上讲，还处在定性分析阶段。因此，努力把断口分析的技术和方法提高到定量分析的理性阶段，这更是失效分析工作者面前的重要任务之一。

在失效模式诊断中综合诊断技术和方法的应用，特别是应力分析和失效模拟技术和方法的综合应用，虽然对一般的失效分析来说，有时它们是辅助的、引证性的，并不是必不可少的、必须进行的，但是对重要的失效事故分析和预防而言，有时却是十分必要的，特别是对失效结论有争议的情况下。当前工作，已有越来越多的失效分析工作者在具体的失效分析案例研究中，重视应用综合诊断技术和方法的问题，但是，这方面的实践和研究还不够系统和深入。

对于失效模拟实验中的力学模型的设计、介质环境条件的模拟、力学参数的加速、断裂数据的当量关系以及模拟断口和实际断口（或对腐蚀、磨损失效来说，模拟表面和实际表面）的对比分析技术和失效模拟实验的有效性和可靠性分析等方面的问题还有待针对不同类型的失效模式进行认真和深入的研究，才能真正发挥失效模拟技术和方法在失效模式诊断中的作用。

8）值得特别提出的是，失效模式的概率诊断、模糊诊断和计算机人工智能诊断系统的问题。

由于影响或控制失效的各因素的不确定性，包括构件受力的不确定性、构件尺寸的随机分布、环境因素的不确定性、材料力学性能的不确定性、内部缺陷的随机分布等，因此，构件的失效也将由偶然性和必然性的协同作用所控制。必然性反映规律，而偶然性反映差异，基于诸多不确定因素的失效模式也具有不确定性，或随机性，而不是确定型的。所以失效模式的诊断也应采用概率诊断或模糊诊断的技术和方法，特别是计算机人工智能诊断系统。在不久的将来，有望建立一个可供查询、诊断、预测的多媒体工具智能系统，并成为我国参加尤里卡项目“工业装置腐蚀失效分析方法”的主要研究内容。这是一个可喜的有发展潜力的开端。

总之，失效模式的诊断理论、技术、方法和实践等方面已取得了长足的进展，但从学术体系的完整性、从技术的有效性和方法的严谨性的角度来看，还有很多工作可做，这也是失效分析工作者努力奋斗的方向和目标之一。

（2）失效原因诊断的进展

1）失效原因，有时又称事故原因，是指酿成失效事故（或事件）的直接关键因素。失效原因也可以分为一级失效原因、二级失效原因、甚至三级失效原因。一级失效原因，一般指酿成该失效事故（或事件）的首先失效件（肇事件）失效的主要直接关键因素处于投付使用过程中的哪个阶段或工序，可以分为设计原因、制造原因、使用原因、环境原因、老化原因等。二级失效原因是指一级失效原因中的直接关键环节，如设计原因中又可以分为设计原则、设计思路和方案、结构形状和受力计算、选材和力学性能等次级原因。失效原因的诊断是失效分析和预防的核心和关键，它不仅是失效预防的针对性和有效性的重要前提和基础，而且它常与酿成失效事件的责任部门和人员相联系。因此，对失效原因诊断应该特别强调其科学性和公正性。

2）失效原因诊断的思路和方法有不同的分类方法：按诊断采用的方法可以分为理化诊断方法和系统工程诊断方法。理化诊断方法又可分为以失效件（肇事件）强度为主线的诊断方法、以失效件（肇事件）的生产工艺为主线的诊断方法、以失效件（肇事件）服役条件为主线的诊断方法和以失效件（肇事件）或设备类别为主线的诊断方法等。一种故障树分析法（FTA）、特征因素分析法、失效模式及后果分析法（FMEA或FMECA、时序树分析法（ETA）等。按事故起因链关系，可以分为因果关系的诊断方法和非因果关系的诊断方法。因果关系诊断方法又可分为多米诺骨牌法、多重原因和分支事件链法、轨迹交叉论法、扰动论法、意外释放能量论法等，非因果关系的事故起因链模型方法主要是管理

失控的事故致因理论及其方法等。从目前的情况来看，失效原因的理化诊断方法采用较多，而系统工程诊断方法采用较少；因果关系的逻辑推理诊断方法采用得多，而非因果关系的事故起因链模型方法采用得少。这可能与我国失效分析工作者的知识构成有关。因此，我们失效分析工作者面临着扩大知识面、不断提高失效原因诊断理论、技术和方法综合能力和水平的任务。

3）导致失效的原因往往与该失效产品所处的开发阶段有关。一个定型的、大批量生产产品的小概率或长寿命产品的偶然失效事故的失效原因往往与工艺或环境中的偶然差错或因素有关；而一个试制品或重复出现的大概率的失效事故的失效原因，常与设计结构因素有关。因此，在失效原因诊断中首先应该确认该失效产品所处的开发阶段和注意失效概率的大小。

4）不少人常采用断口分析技术作为失效原因诊断的主要依据，甚至只依靠断口的形貌就得出失效原因的结论，由此造成了误判或错判。一般认为，断口分析技术对于断裂失效的模式、特别是对于典型的一级断裂失效模式的诊断是比较准确的，但是断口分析不是“万能”的依据，将它应用于断裂失效原因的诊断就应该格外慎重。有的人一旦发现断口上有夹杂、或有气孔、或有毛刺，不分析断裂的走向、不观察夹杂或气孔是否在断裂源处，或堆在断裂源处，不估算夹杂（或气孔、毛刺）的大小是否与临界裂纹的尺寸相当，不定量或半定量地分析该夹杂（或气孔、毛刺）对这种断裂模式的影响有多大，也不分析导致断裂的应力的来源和大小，就把次要原因当作主要原因，从而使整个失效事故分析和预防工作走了弯路。这些历史的教训是值得我们认真吸取的。

5）在失效原因的诊断中，还有一个需要特别强调的问题是：要多从具体的失效事故的事实出发，少犯经验主义的错误。因为原则上来说，失效是绝对的，是当时人们的科学技术认识与客观事物发展规律不一致的结果。失效事故中蕴藏着“新鲜事物”。压力容器的“爆沸”失效分析和诊断就是一个典型的例子。压力容器的“爆沸”是一种危险的爆炸失效方式，它是由于在压力容器爆炸前的瞬间，内部压力突然下降，使容器处于过热状态的介质在大气压下迅速沸腾蒸发，体积激烈膨胀而引起的一种爆炸现象。曾对“爆沸”的必要条件和充分条件进行了定量分析，力图揭示“爆沸”的规律性。这是一种大胆的探索。因此，对压力容器的“爆沸”爆破失效事故原因的分析诊断，首先要通过定性分析确定。“爆沸”事故的分析中，有人根据从爆破断口上找到的缺陷的性质，得出其“爆沸”的原因，这往往是错误的。

6）还应该特别指出的是：鉴于失效原因或影响失效原因的因素很多，它们的不确定性更加突出，特别是现代装备使用的条件越来越苛刻、安全储备系数越来越小、工作能力也越来越复杂，因此，单因素引起的失效原因已基本不存在，大多是多因素的非线性耦合交互作用的结果。这不仅将促使我们去研究多因素失效作用的非线性叠加理论、损伤累积过程的物理数学模式，而且将逐步建立多层（重）原因和分支事件链的诊断理论、技术和方法体系和失效原因诊断中概率诊断、模糊诊断和计算机人工智能诊断技术和方法的兴趣和应用。由此可以看出，在失效原因诊断的理论、技术和方法的实践和学术方面还任重道远，还有很多工作要做。

(3) 失效机理诊断的进展

1）失效机理的诊断是指对失效的内在本质、必然性和规律性的研究，它是人们对失效性质认识的理论升华和提高。一般来说，一个失效事故是外因和内因共同作用的结果。常把应力、温度、气氛介质等作为影响失效的外因；而把材料的成分、组织、缺陷、性能和它们的表现当作影响失效的内因。失效的机理学就是内因和外因共同作用而最终导致失效事件发生的热力学、动力学和机构学，即失效内在的必然性和固有的规律性。拿生病来作比喻，失效模式相当于病症，失效原因是生病的原因，而失效机理则相当于病理。一个失效事件，在失效机理尚未揭示的情况下就得出失效模式和原因的诊断结论意见，很可能是不牢靠的，或者是不科学的，并有可能造成误判。因此，失效机理的诊断或研究是十分重要的，而且，只有揭示失效的必然性和规律性，才能真正做到对同类失效事件的有效预防，做到“举一反三”和亡羊补牢。

2）失效机理诊断的基础是对失效机理的研究。失效机理的研究可以分为宏观失效机理的研究（即失效的热力学和动力学研究）和微观失效机理的研究（即失效的微观机制理论和模型的研究）。按照研究尺度的大小，微观失效机理的研究又可分为细观尺度、分子尺度和原子尺度的研究。失效机理研究采用的方法可以分为定性的方法和定量的方法，定量的方法双可分为确定型方法和概率型（或模糊型）方法。失效机理研究所采用的模型按所属学科分为理化模型和物理数学模型，并可进一步细分为力学模型、物理模型、化学模型、材料学模型、数学模型以及它们联合作用的复合模型，具体的有：界限模型与耐久模型、应力强度干涉模型、反应论模型、失效率模型、最弱环模型与串联模型、绳子模型与并联模型、比例效应模型、退化模型或损伤积累模型等，可以看出，失效机理的定量研究的内容和范围十分广泛和充实。

3）近几年来，在失效机理的研究方面做了大量的工作，并取得了长足的进展和提高，但是深入程度和系统性仍显得不足，对很多失效的机理仍是一知半解。例如，有文献指出，人们对疲劳短裂纹的萌生机制和扩展规律还很不清楚，并引用英国 R.A.Smith 教授指出的“反抗疲劳的设计仍然是一个挑战性的问题。我们对材料中局部小区域内的缺陷和对复杂环境中作用载荷的大小都缺乏精确的了解，因此，材料的疲劳失效还会出现。虽然火车轴发生疲劳断裂的事故今天已经很少了，但 1986 年日本航空公司波音 747 大型宽体客机的失事警告我们，我们防止疲劳失效的知识和能力仍然是很有限的。”因此认为，未来的材料疲劳研究的重点问题之一是疲劳的基本过程、规律和机制。研究者着重指出失效分析工作者应重视的几点新方向之一是“微电子装备的可靠性”。可惜我们对微电子器件和装备的失效机理、检测技术、预测预防方面的研究甚少。这是我们今后应该加强研究的课题之一。与此同时，对失效的物理数学模型的研究也应该有所进展和提高。

总之，失效诊断的理论、技术和方法研究已有雏形，但仍需加强。我们的总体目标是对每起失效事故准确、客观、科学和适时的诊断。为此，应注意积累和总结经验，努力揭示其规律和本质，不断提高失效诊断的能力和水平，并逐步建立和发展一整套关于失效诊断的理论、思路、技术和方法的学术体系。

7.2 失效预测的理论、技术和方法的进展

失效预测研究是失效诊断研究的继续和发展。失效预测研究是变事后的、被动的失效诊断为事前的、主动的失效控

制的重要研究。失效预测研究可以分为安全状况预测、剩余寿命预测和累积失效概率（可靠度）预测等三个层次的二级方向和内容。

（1）安全状况预测的进展

1）安全状况预测又称为安全评估（或安全评定）。安全评定就是按“合于使用”原则对含缺陷零件进行安全与否的评价。安全评定的三要素是应力、缺陷和材料力学性能。根据失效模式的不同——是“静态”的失效模式还是“动态”的失效模式——大致可以将安全评定技术划分为断裂和塑性破坏安全评定以及疲劳安全评定两种。其他失效模式的安全评定基本上是这两种失效模式安全评定技术和方法的扩展和延伸。

2）断裂安全评定的技术和方法是通过国际通用的失效评定图技术实现的。失效评定图是英国中央电力局（CEGB）的“双判据”评定法（通常称为R6评定法）为基础，结合美国通用电气公司（GE）得到的形变塑性理论解，扩充了R6失效评定图的使用范围，使之能用于分析应变硬化材料、并考虑二次应力（残余应力或热应力），随后完善发展起来的，其评定依据是双参数的失效评定图。失效评定图是一个以 K_r（$K_r=K_I/K_{Ic}$，K_I 为裂尖附近的应力强度因子，K_{IC} 为材料的断裂韧度值）为纵坐标、以 L_r（$L_r=P/P$，P 为向零件施加的一次载荷，P_0 为材料的塑性屈服极限载荷）为横坐标、以下列国际通用的失效评定曲线为界线：

$$K_r=(1-0.14L_r^2)[0.3+0.7\exp(-0.65L_r^6)]$$

并以不同数值的 L_r^{max} 为截止线围成的一个图形。实际零件的评定点落在图形的范围之内为安全，落在范围之外为不安全；评定点落在纵坐标上或靠近纵坐标为脆性断裂，落在横坐标上或靠近横坐标为塑性破坏或准塑性破坏。断裂安全评定研究的主要任务是如何计算各种缺陷类型、大小、分布时，各种材料力学性能条件下，零件的 K_r、L_r 值的失效评定图中评定点位置的确定。

3）塑性破坏评定也适用于凹坑缺陷。压力容器的凹坑缺陷的塑性破坏评定是指采用塑性极限准则对凹坑缺陷的安全与否进行评定。它是根据凹坑缺陷的极限载荷图进行的。凹坑缺陷的极限载荷图是以带凹坑的压力容器极限载荷 P_1 与无凹坑的压力容器的极限载荷 P_{10} 的比值 P_1/P_{10} 为纵坐标、以凹坑的量纲为1的几何参量 G_0 ［$G_0=CA/(T\sqrt{RT})$，A 为凹坑缺陷的半长轴，C/T 表示凹坑缺陷的相对深度，$A/\sqrt{RT}$ 表示凹坑缺陷的相对长度］为横坐标、以下列直线作为界线的图形：

对球形容器 $G_0=(1-2.0P_1/P_{10})/0.6$

对圆筒形容器 $G_0=(1-2.0P_1/P_{10})/0.09$

实际的评定点落在图形的范围内为安全，落在范围外为不安全。

凹坑缺陷的塑性破坏评定研究的主要任务是如何计算存在凹坑情况下，各种形状的压力容器或其他零件）的塑性极限载荷的问题。

4）疲劳安全评定方法基本上按照二级疲劳模式——是应力疲劳还是应变疲劳——来确定的。如果是应力疲劳，它的疲劳安全评定可按疲劳图来进行；如果是应变疲劳，则应按Paris公式计算其在极限寿命内的裂纹长度 $a_{疲劳}$ 与依据断裂的临界裂纹长度 a_c 比较来进行的。如果，$a_{疲劳}<a_c$ 为安全，否则为不安全。

疲劳图分为两类：第一类疲劳图和第二类疲劳图。第二类疲劳图是以应力振幅值 σ_a 为纵坐标，以平均应力 σ_m 为横坐标，以 $\sigma_{-1}-\sigma_b$ 之间的关系曲线为界线组成的图形。实际受力的 σ_a 和 σ_m 点在图形范围内为安全，在图形范围外为不安全。

可见，疲劳安全评定技术的关键是应力幅、平均应力和裂纹长度的计算方法问题。

5）我国在制定和实践“压力容器缺陷评定规范”（CVDA-84）的基础上，经过“八五”国家科技重点项目“在役锅炉压力容器安全评估与爆炸预防技术研究”的研究，于1995年提出了“在役含缺陷压力容器安全评定规程”（SAPV-95）（第四稿）。这个“规程”是以近代科学技术发展为其理论基础，充分吸收内外压力容器安全评定技术和规范最新研究成就和发展趋势、积极继承我国的CVDA规范的精华、密切吸取十多年来在压力容器安全评定工程实践经验和系统综合攻关科研成果的基础上提出来的。这个“规程”对体积型缺陷的安全评定采用我国首创的凹坑缺陷压力容器塑性极限载荷分析法及其成果，在平面缺陷的断裂评定和塑性破坏评定中，采用了三级评定和失效评定图技术、我国独有的处理二次应力工程分析方法和更符合实际的塑性修正因子，以及采用了我国研究的裂纹（缺陷）间的弹塑性干涉效应分析法和提出其弹塑性干涉效应系数G等，使得这个“规程”具有20世纪90年代的国际先进水平。这是我国压力容器安全评定技术和方法的一大进展，对我国压力容器和其他构件的安全可靠使用具有重要的作用。

值得提出的是，结构完整性评定技术在我国逐渐得到了成功的应用，并创造了大量的经济效益。

6）虽然我国在结构完整性评定技术和方法方面取得了很大的进步和提高，但是在下面诸多方面还需要进一步深入的研究：

① 深入研究各类缺陷的起裂、扩展、撕裂、失稳等断裂全过程的安全评定技术体系、有机衔接和合理应用，包括深入研究与裂尖三维约束程度有关的多参量弹塑性断裂、非匹配焊缝、缺陷等高应变区的断裂分析工程方法和裂纹的非 J 积分控制扩展的静载断裂评定技术等；

② 深入研究除室温静载断裂失效模式以外的其他断裂失效模式，如蠕变断裂、应力腐蚀断裂、疲劳断裂、腐蚀疲劳断裂等与环境、介质、时间有关的断裂失效模式的安全评定技术和方法；

③ 建立我国独有的、涉及压力容器和压力管道的、包括安全评定和安全等级划分的安全预测体系、技术、方法和软件。

（2）剩余寿命预测的进展

1）顾名思义，剩余寿命是指在线（役）零部件仍可使用或工作的时间（或循环）；而剩余寿命预测则是由实验室（或现场）收集的数据（包括载荷、环境条件、强度试验、寿命试验等）来预测零部件（或系统）在现场实际使用条件下仍能使用（或工作）的时间（或循环）的技术和方法。它在工程技术方面是一项十分重要的技术领域，对零部件的定寿、延寿来说是一项关键的基础研究。

2）原则上说，只有与时间有关的失效模式（如疲劳、蠕变、应力腐蚀、腐蚀等）才有寿命预测问题，而与时间无关的失效模式（如脆性断裂、韧性断裂等）则只有安全（或断裂）预测或安全评定问题。

从原理上看，寿命预测技术和方法分为互有联系的三类：一是基于物理模型的寿命预测技术和方法；二是基于力学（包括断裂力学）的寿命预测技术和方法；三是基于系统工程的寿命预测技术和方法。它们又分别大致可以分为确定型和概率型（或模糊数学型）的技术和方法。

3）疲劳寿命预测可以分为总寿命法和损伤容限法，在恒幅应力下疲劳寿命的预测主要采用以 Paris 定律关系的寿命估算方法；而在变幅应力（或应变）条件下，则采用 Palmger-Miner 线性损伤法则为基础的逐个循环估算方法（包括屈服区模型、裂纹闭合的数值模型和工程方法等）和特征方法［认为裂纹尖端的随机变化可以用应力强度因子范围的均方根 ΔK_{rms} 来描述，于是变幅裂纹扩展速率可以用 Paris 型关系 $da/dn = C\ (\Delta K_{rms})^m$ 来表示］。不管什么寿命预测方法，其最终目的是追求安全性和经济性的结合，并提出安全－寿命概念和失效－寿命概念等。

4）材料或构件腐蚀寿命的预测是十分重要的。大型装备尤其是军用装备，如飞机、军舰、坦克等装备的总寿命是由使用次数、工作小时数和日历寿命三大额定指标串联构成的。在和平时期，其总寿命基本上是由日历寿命决定的，而日历寿命则取决于腐蚀寿命。由于影响腐蚀的因素很多且复杂，因此，至今仍没有一种可以预测在复杂腐蚀环境介质下，金属构件腐蚀损伤日历寿命的计算模型和确定方法，直接影响我军装备的使用安全和战斗力。国内外公开发表的32种金属材料的腐蚀试验数据分析，成功地提出和验证了腐蚀温度 T 和腐蚀时间 H 之间存在着类似疲劳 $S-N$ 曲线形状的 $T-H$ 曲线和金属腐蚀损伤性累积理论公式，进而研究和提出了一种计算在使用环境温度-时间腐蚀谱条件下的金属构件腐蚀损伤的日历持续时间（寿命）计算模型和确定方法。这是很有价值的研究工作。但是正如作者所指出的，如同 $S-N$ 曲线中有 σ_{-1} 一样，腐蚀 $T-H$ 曲线是否也存在 T_C 渐近线？各种腐蚀损伤机制中是否都存在 $T-H$ 曲线？不同腐蚀量时，$T-H$ 曲线的形状如何？它们腐蚀累积公式应如何修正？自然大气腐蚀和加速试验的机制模拟和当量关系如何？这些问题还有待于深入、系统地进行研究。

5）虽然剩余寿命预测研究方面做了很多实践和研究，对一些关键的、不可替换的零部件（如核反应堆结构材料或构件）的寿命预测研究也有可喜的进展，但是对蠕变寿命、应力腐蚀寿命，在各种工况条件下，特别复杂、苛刻条件下的材料或构件的退化规律和寿命、以及上述各种情况下的损伤微观机制、表征参量、退化方程和客观寿命预测方法的可靠性等还需要进行全面、系统和深入的研究。

（3）累积失效概率（可靠度）预测的进展

1）众所周知，失效率是产品可靠性的一个重要指标。为了求得产品的可靠度 $R\ (t)$，可观测 $t=0$ 时 $n\ (0)\ = n_0$ 个样品的功能，然后求出到某一具体时刻 t 为止，剩下的满足功能的样品的比率，即求残留率 $n\ (t)\ /n_0$。这就是由样品估计出来的产品的可靠度。累积失效概率 $F\ (t)$ 就是不可靠度 $[1-R\ (t)]$。因此，累积失效概率的预测就是可靠度的预测，其目的是：①在设计阶段通过预测了解可靠性是否满足了要求的指标，确定可靠性设计方法；②通过预测找出提高可靠性、可维修性的关键问题或关键元器件。可靠度预测的方法主要有：应力-强度干涉法、失效树分析法（FTA）、失效模式影响因素和致命性分析法（FMECA）、蒙特卡洛法（monte carlo）等。其中，应力-强度法主要用于元器件，其他几种方法主要用于系统或设备。由于不确定型的模型及其结果不够“确定”或不够可靠和精确，因此，可靠度预测及其分析方法就越来越显得重要。可靠度预测的目标，就是要追求和达到系统、设备、元器件的安全和经济之间的合理匹配。

2）可靠度预测的依据是使用工况条件或在加速试验条件下的大量数据，据此数据对失效概率进行统计分析。因此，它的数学基础是概率、统计的基础理论及其广泛的应用领域。近年来，应用数学的发展大大得益于运筹学和电子计算机科学的进步，因此，可靠度预测的技术和方法的领域又进一步扩大了。

3）总的来看，可靠度预测是一种技术和方法。根据它应用的方面或问题的不同，可产生不同的效果。将它用于失效诊断，则称为可靠性失效诊断；用于安全评估，则称为可靠性安全评估；用于寿命预测，则是可靠性寿命预测。因此，从失效学的角度来看可靠性，就是要用可靠性工程和技术的成果，来解决失效诊断、预测和预防中的方法问题，以便达到安全和经济的双重目的。

4）可靠性技术在失效分析预测预防方面的应用，已有了较好的开端。但是无论从数据量（特别是实际的失效数据量）、数据处理技术（包括概率分布分析和计算机处理），还是从实际应用效果等方面来看，都有待于加强。

7.3 失效预防的理论、技术和方法的进展

1）失效预防是失效诊断和失效预测的最终目的和成果。一个失效事故（或事件）的分析和研究，没有提出失效预防措施是不完整的。失效预防技术就其所属学科而言，可以分为力学的、化学的、物理的、材料的和管理的技术；就其过程来说，一般可以分为工程技术、安全法规或标准的制定或修改、失效分析和预测预防数据库的建立、发展、完善和应用。不管采取主要属于哪一种的失效预防技术、方法和措施，失效预防或“预防事故应当以安全管理为中心”，“要求我们从总体上采取预防措施，也就必然要涉及从全局的管理上采取安全措施，例如各种安全法规、安全标准、安全文化等方面”“并且在系统涉及的所有子系统的全过程采取安全措施”，“才能从根本上预防（减少）事故的发生”。此外，为了用现代科学技术成就分析、整理和总结已有的失效预测预防方面的经验和教训，逐步建立失效预测预防人工智能系统已势在必行。

2）在诸多的失效预防技术和方法中，采用最多的还是工程技术和方法，下面将着重介绍我国在失效预防中采用较多的表面防腐和表面强化技术和方法方面的进展。

表面防腐和表面强化技术和方法是表面工程的一部分。表面工程是近十多年来迅速发展起来的新兴学科。表面工程是经表面预处理后，通过表面涂敷、表面改性或多种表现技术复合处理，改变固体金属表面或非金属表面的形态、化学成分和组织结构，以获得所需要的表面性能的系统工程。由于材料的失效（例如磨损、腐蚀、高温氧化、疲劳断裂等）往往自表面开始，因此，根据需要采用表面防腐和表面强化技术，改善材料（构件）的表面性能，将有效地延长其使用寿命，节约资源，提高生产力，减少环境污染。

表面防腐和表面强化技术和方法主要有：表面喷涂技术，包括氧－乙炔火焰喷涂、等离子喷涂、塑料粉末喷涂和其他特种喷涂等；表面电沉积技术，包括电镀、电刷镀、化学镀、转化膜技术等；表面黏涂技术，包括表面涂层、化学

黏涂、胶黏等；表面强化技术，包括表面真空熔结、表面渗层（渗碳、渗氮等）、激光表面强化、电火花表面强化、气相沉积、离子注入、表面喷丸强化等。我国在上述表面技术的诸多方面进行了深入的研究和大量的工程应用，并取得了很大的效益。

3）失效预防的理论、技术和方法要达到科学性、有效性、可靠性和经济性等目标还要进行更多的、更系统和更加深入的研究，并通过其他科学的知识和成果的移植，逐步建立其学术体系，促进机电装备失效分析预测预防进行良性循环，进而发挥其杠杆作用，从一个侧面来推动我国科学技术的进步和国民经济的可持续发展。

8　结束语

可以认为，机电装备的失效分析预测预防是从失效入手，着眼于成功和进步的科学，是从过去入手，着眼与未来和发展的科学，越来越显示出它的生命力和潜力，已取得了比较完整的工作经验和显著的社会经济效益。当前，它正从各行业工程化研究发展到跨行业、跨学科的理论化、基础化和综合化的研究；从考虑材料在简单服役条件下的状态发展到分析材料（构件）在复杂服役条件下损伤演化机制；从失效机理的定性研究发展到利用各种数学物理模型和现代实验技术为基础的定量分析；从一般的断口分析发展到无损检测、失效诊断、安全评估、寿命预测、事故预防、寿命的控制、定寿、延寿的全过程的研究、从一种简单实用的事故分析技术向一门分支学科——失效学方向发展，正处于感性向理性转变的重要时刻。总之，我国失效分析工作者应继续努力，为我国机电装备失效分析预测预防事业的飞跃共同奋斗。

编写：钟群鹏（北京航空航天大学）

第2章　失效诊断技术和方法

失效诊断包括失效模式诊断、失效原因诊断和失效机理诊断三方面的内容。失效诊断是整个失效分析预测预防工作的前提和基础，对整个失效研究工作具有“定向”的作用。失效诊断得越准确、越细致，对相应的预测预防工作的指导意义就越大。目前，失效诊断工作已引起了工程界的广泛关注和高度重视。

失效诊断是一门系统工程，其理论、技术和方法的核心是其推理规则和方法论。在实际的失效诊断中，主要从残骸、应力和环境等方面进行分别诊断和综合诊断。其中残骸分析又包括失效件本身的断口分析、裂纹分析、痕迹特征分析和变形分析。

在失效诊断过程中，要求广大失效分析工作者实事求是，不预设观点，以客观性作为根本原则；在观察过程中，遵循先宏观、后微观，先低倍、后高倍，先全局后局部，典型区域重点观察的诊断方法和原则。

1　断口诊断技术和方法

断口是断裂失效中，两断裂分离面的简称。由于断口真实地记录了裂纹由萌生、到扩展直至失稳断裂全过程的各种与断裂有关的信息，因此，断口上的各种断裂信息是断裂力学、断裂化学、断裂物理等诸多方面内外因素综合作用的结果，对断口由定性到定量的精确分析，不但可为断裂失效模式的确定提供有力的依据，也可为断裂失效原因的诊断提供线索，对断裂机理的深入研究还可推动断裂物理、断裂化学和断裂力学的发展。同时，断口分析对新材料的研制也是一个十分重要的手段，通过断口分析可以提供有关合金的相组成、组织结构、杂质含量对断裂特性的影响等，为进一步改进材料质量提供方向和可能。基于以上原因，长期以来，在断裂失效中，断口诊断和裂纹诊断总是占据着极其重要的位置，甚至有人把机械失效分析等同于断口分析，虽然这是十分错误的，但也说明了断口分析在失效分析中的地位和作用。

1.1　断口准备

断口准备的目的是为下一步的断口分析提供适于分析的断口。这就要求断口保存得尽量完整、特征原始。由于断口分析主要是在实验室内完成的，因此，从事故（失效）现场到实验室的运输过程中，应尽量使断口得到较好的保护，不会产生二次、甚至三次损伤。有时由于断口上附着有腐蚀介质、污染物，还需要对断口进行适当的清理；有时失效件体积太大，无法装运，还必须进行分解、切割。总之，在断口准备过程中，所有这些问题均应考虑到，总的要求是尽量保证断口（特别是关键断口、起始区断口）的原始特征不被破坏、污染。

1.1.1　断口的清理

零件在使用过程中和失效之后，断裂表面在断裂过程中和断裂之后，表面均不可避免地会受到其他零件和（或）环境的、机械的、化学的损伤与污染，为了能够观察到零件和断口的真实形貌与失效过程，都需要对零件和断口进行清理和清洗，以将覆盖在零件和断口上的尘土、油污、腐蚀产物及氧化膜等可能造成假象的多余物清除掉，以露出需要观察表面的原始形貌。这就是对断口进行清理和清洗的目的。

对断口的清理应遵循以下基本原则：先判断后清理，先表面后内层，尽量采用物理方法清理而少用化学方法。所谓“先判断后清理”，是为了不将有用的信息清除掉。在对断口进行清理和清洗之前，应该对断口进行仔细的观察与检查，先对表面的损伤与附着物等进行判断，必要时，还需进行相应的分析测定。如对腐蚀产物的分析测定，可判定腐蚀介质的种类和腐蚀产物的形成时期，从而有助于分析失效的原因。“先表面后内层”是指在清理过程中，应由表及里地进行，切不可一下就将所有附着物全部清理掉，以免丢失有关失效过程的详细信息。在清理过程中，物理方法，如用软毛刷清理表面浮尘等，只要使用得当，一般不会对断口造成二次损伤、不会在断口上形成假象，应优先考虑使用；而化学方法由于其过程的难以控制，往往容易对断口再次造成损伤，在断口上产生假象，因此，应该尽量不采用。

断口的清洗方法很多，可根据断口材料特性、附着物的种类等因素加以选定。表5.2-1中列出了对金属材料断口常用的清洗方法，在实际失效分析工作中可根据情况加以选用。对于断口表面只有尘埃或油渍污染者，推荐使用丙酮与超声波清洗；对于遭受轻微腐蚀氧化的断口，推荐使用醋酸纤维（AC）纸反复复型剥离法加以清洗；对于遭受较重腐蚀氧化的钢制零件断口，推荐在10% H_2SO_4 水溶液＋缓蚀剂（1%卵磷脂）中超声波法清洗；对于高温合金的高温氧化断口，可使用“NaOH＋$KMnO_4$”热煮法予以清洗。需要注意的是，不管使用何种方法清洗，都要以既要除去断口表面的附着物，又不损伤断口的形貌特征为原则。

表5.2-1　清洗断口表面常用方法

类型	程序与内容	应用范围
机械剥离法	醋酸纤维（AC）纸反复复型剥离	可清除较轻微的锈蚀
化学腐蚀法	① 在10% H_2SO_4 水溶液＋1%卵磷脂缓蚀剂中超声清洗	碳钢、合金钢、不锈钢、耐热钢及铝合金
	② 在0.5%的乙二胺四醋酸钠水溶液（EDTA）中	铝合金、钛合金、合金钢、不锈钢、耐热钢
	③ 6 mol/L HCl 酒精溶液＋2 g/L 六亚甲基四胺，浸渍1～30 min或超声清洗5～30 min	清除钢断口上的锈层
	④ 在3% HCl 酒精溶液＋2 g/L 六亚甲基四胺中浸渍1～10 min或超声清洗5～30 min	清除钢断口上的锈层
	⑤ 50%柠檬酸水溶液＋50%柠檬酸铵水溶液中浸渍或超声清洗	严重锈蚀的钢制零件断口
	⑥ 丙酮＋0.5%～1%盐酸中浸渍或超声清洗	严重锈蚀的钢制零件断口
	⑦ 在78 mL水＋16 g氢氧化钠＋6 g高锰酸钾中煮沸5～30 min后，取出放入60～70℃饱和草酸水溶液中清洗，再在丙酮中超声清洗	铁基、镍基、钴基高温合金的高温氧化断口
阴极电解法	① NaCl 500 g＋NaOH 500 g＋H_2O 5000 mL，电流4 A/cm²、电压15 V	钢制零件断口上的锈层

续表 5.2-1

类型	程序与内容	应用范围
阴极电解法	② $Na_2CO_3$30 g + $Na_2SiO_3$20 g + $Na_3PO_4$20 g + NaOH10 g，H_2O1000 mL，电流 2～5 A/cm^2	钢制零件断口上的锈层
	③ 1%H_2SO_4 + 2 g/L六亚甲基四胺 + NaOH500 g + H_2O500 mL，电压 15 V，电流 4 A/cm^2	钢制零件断口上的锈层
真空蒸发法	在真空中加热蒸发，除去断口表面层的低熔点金属覆盖物	液体金属致脆断口

1.1.2 断口的保存（保护）

断口在切割、运输、清理（清洗）等处理过程中，一定要注意如下几点。

1）要防止断口及其附近区域的显微组织因受热发生变化。

2）要防止断口的形貌特征受到机械的或化学的损伤。

3）在实施切割或打开裂纹的操作之前，要对失效件的外观特征进行仔细的观察与测量，并要将观察与测量结果用文字和照相的方式详细地记录下来。

4）在清洗之前，要对断口进行仔细的观察与检查，必要时对断口表面的附着物进行分析测定，这将有助于揭示断裂失效的原因（如腐蚀产物、低熔点金属等）。

5）在检查和清洗断口时，切忌用手去触摸断口表面，更不能将两个匹配断面对接碰撞，以免使断口表面产生人为的损伤。

6）为防止断口表面在运送与保存过程中遭受腐蚀与损伤，可在断口表面上涂抹一层保护材料，如醋酸纤维丙酮溶液、可剥涂料等；也可将断口直接浸泡在无水酒精溶液中。

7）断口在送交有关部门分析检验时，一定要保持断裂件的原貌。应加强断口保护，避免腐蚀、污染及损伤，并尽量提供详细的失效经过、失效件履历、现场检查情况等有关材料。

1.1.3 断口的选取（切割）

对大型设备、零件，失效后的残骸块大、质量大，难以运输和在实验室观察，需要选取典型断口进行分析。在选取需进一步分析的断口时，应先对断口进行宏观分析，确定首断件；然后进一步确定断裂的起始部位。在切割时，应先将需要分析的部位保护起来，然后进行切割；切割时，尽量使用锯、切等不会产生高温的机械方法。需要使用火焰切割、砂轮切割等会产生高温的切割方法时，切口位置应离开需分析部位一定的距离，同时对切割区域进行冷却，以确保需重点分析的部位不会因高温而产生氧化、组织变化、性能变化等二次损伤。

1.2 断口诊断仪器设备和技术

在现代分析研究中，各种仪器设备是必备的手段。只有熟练掌握、合理利用好各种相关仪器设备，才能得出科学、客观的分析结论。

1.2.1 断口宏观诊断仪器设备和技术

（1）宏观诊断仪器设备

宏观诊断是指在小于40倍的条件下对断口进行观察判断的技术方法。宏观诊断的主要手段是人的肉眼、普通放大镜和体视显微镜。

体视显微镜是失效诊断中常用的观察仪器，是获取有关宏观失效信息的重要手段。图5.2-1是一普通体视显微镜的外形图。

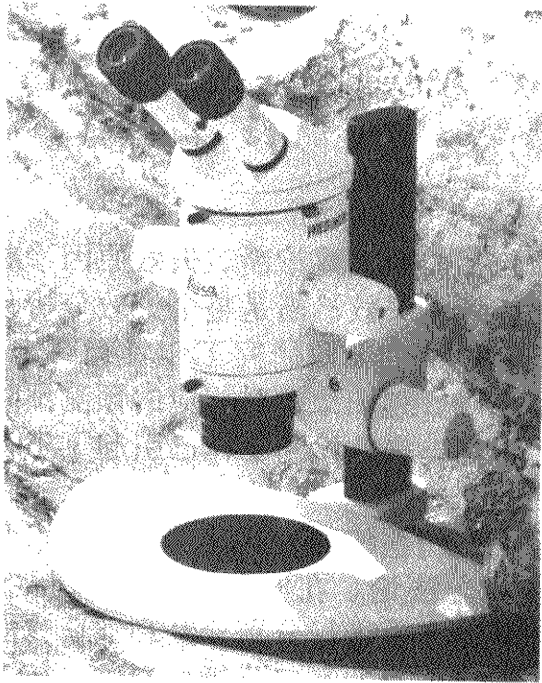


图 5.2-1 体视显微镜

体视显微镜的主要特点是景深大、观察的影像具有较强的立体感，能够观察较为粗糙的表面，给人以整体而真实的概貌。由于体视显微镜观察具有很强的真实感和立体感，观察者完全可以凭借直观感觉来理解在体视显微镜中的景象。

体视显微镜的放大倍数可以从几倍到100倍左右。

（2）断口宏观观察技术

对断口进行宏观观察时，应先用肉眼对断口进行观察，掌握断口的全貌和特征规律；然后借助放大镜或体视显微镜进行整体观察和重点观察。观察时应该尽量在自然光下进行；借助灯光时，应注意观察的角度，并旋转断口或灯光，从不同的角度来观察断口的形貌特征与规律。

1.2.2 断口微观诊断仪器设备和技术

（1）金相显微镜

金相显微镜是研究和检验金属材料的基本手段，用来对抛光表面的形貌（如金相组织、夹杂物分布、裂纹走向及形状等）进行显微放大观察。图5.2-2为一金相显微镜的外形。金相显微镜的放大倍数在几十倍至2 000倍之间。由于其景深较小，在较高倍数时不能观察粗糙的表面（如断口等）。

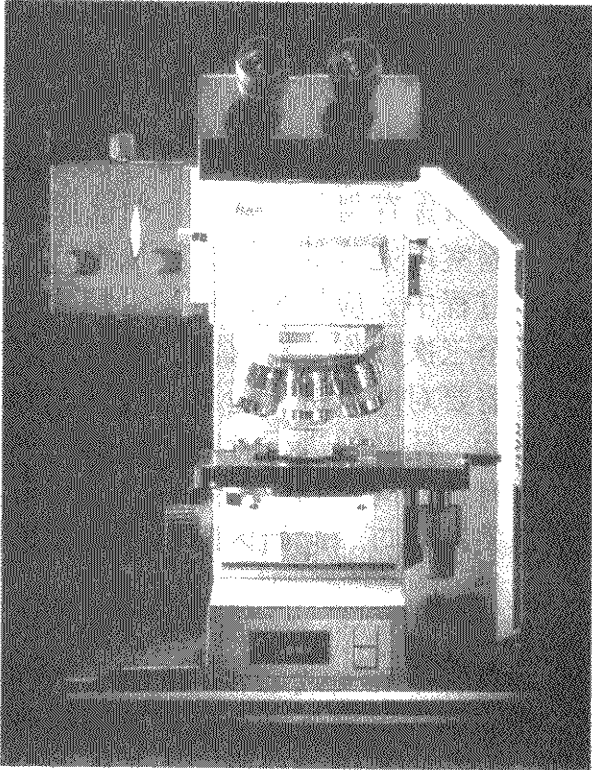


图 5.2-2 金相显微镜

在使用光学显微镜时，应先调整好视度和目距，以适应自己眼睛的视度（远、近视程度）和两眼之间的距离；同时调整好聚焦、照明等。

（2）扫描电子显微镜

1）工作原理 扫描电子显微镜（图5.2-3）的基本工作

原理是，由电子枪（灯丝）射出的电子束经过几百伏到几十千伏的加速电压加速、电子光学系统聚集后，射到样品表面，并在样品表面进行光栅扫描（与荧光屏上的扫描同步）。由于电子束很细，且具有很高的能量，电子束将射入表面一定的深度范围。高能的入射电子将与固体原子相互作用，产生一系列物理信息，如二次电子、背散射电子、透射电子、吸收电子、俄歇电子、荧光电子、特征 X 射线等。当这些特征信息被相应的接收器接收，经信号放大器处理后，送到阴极射线管调制其亮度，便可在荧光屏上显示出图像。扫描电子显微镜显示的是样品的二次电子像，由于二次电子的产额与样品的形貌有关，因此，可进行形貌观察。

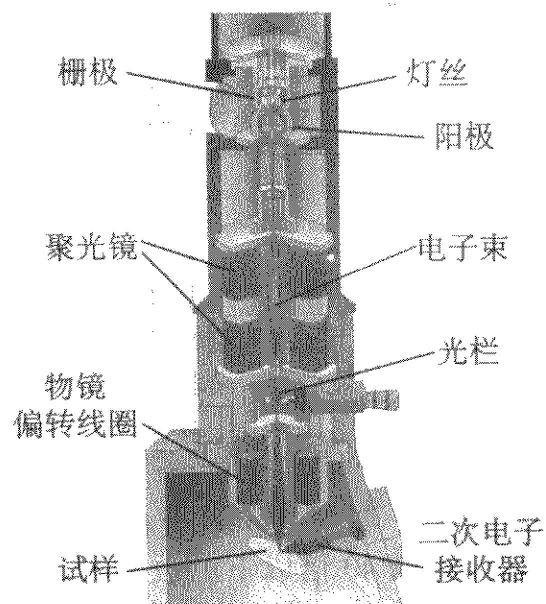


图 5.2-3　扫描电子显微镜工作原理图

2）特点　扫描电子显微镜是一种观察物体表面形貌的电子光学仪器，它主要有如下特点：

① 分辨率高，可优于 5 nm；

② 放大倍数范围广，从几倍到几十万倍；

③ 景深大，适于观察粗糙的表面（如断口分析）；

④ 可对样品直接观察而无需特殊制样；

⑤ 可以加配 X 射线能谱仪或波谱仪，将形貌观察和微区成分分析结合起来。

3）扫描电镜的基本观察技术

① 首先对断口用扫描电镜的较低放大倍数做初步的观察，以求对断口的整体形貌、断裂特征区有全局的了解与掌握，并确定重点观察部位。

② 在整体观察的基础上，找出断裂起始区，并对断裂源区进行重点深入的观察与分析，包括源区的位置、形貌、特征、微区成分、材质冶金缺陷、源区附近的加工刀痕及外物损伤痕迹等。

③ 对断裂过程不同阶段的形貌特征要逐一加以观察，找出各区断裂形貌的共性与个性。

④ 断裂特征的识别。在断口观察过程中，发现、识别和表征断裂形貌特征是断口分析的关键。在观察未知断口时，往往是和已知的断裂形貌加以比较来进行识别。各种材料在不同的外界条件下的断裂机制不同，留在断口上的形貌特征也不同。掌握这方面的知识与经验，是进行断口观察的前提与基础。

⑤ 扫描电镜断口照片的获得。一般地讲，一个断口的观察结果要用如下几部分的照片来表述：断口的全貌照片、断裂源区照片和扩展区、瞬断区的照片。

⑥ 断口的全貌照片可提供断裂形貌的整体观念。在全貌照片上应标明各特征区及微观形貌特征照片的位置。

⑦ 利用背散射电子像观察，可观察到第二相、夹杂物等对裂纹萌生与扩展的影响。

(3) 透射电子显微镜

透射电子显微镜是观察材料精细组织结构的电子光学仪器，具有很高的分辨率（可达 0.1 nm），放大倍数可从几百倍到几百万倍。图 5.2-4 是一透射电子显微镜外形。

由于透射电子显微镜是利用透射电子成像的，所以就要求样品的厚度足够薄，以使电子能够穿过。透射电子显微镜对样品的如此要求，使得块状样品无法直接观察，必须进行超薄切片或复型，做断口观察只能用复型的办法。但因其观察的是样品的复型，且制样技术较复杂，以及其他一些原因，所以只在某些特殊情况下才用来做断口分析。

透射电子显微镜和扫描电子显微镜都是失效分析中常用的设备，各有其优缺点。透射电镜具有分辨率高，放大倍数大的优点。但由于透射电镜观察的是断口的复型，而断面的复型工序较复杂，影响因素多，容易造成假象，同时，很难将所观察到的部位与实际断口上的位置一一对应起来。再者，透射电镜同时所观察到的复型面积很小，一般均在 3 mm^2 以下，不但不能对断口进行连续观察，更不能观察断口的全貌。因此，目前仅在某些特殊情况（例如观察断口的精细特征形貌、分辨较细的疲劳条带等）下，使用透射电镜来观察复型断口的特征。

扫描电镜制样简单，使用方便，可以直接观察尺寸较大的样品（如 100 mm × 100 mm），并具有聚焦景深大，分辨率较高，放大倍数范围宽，并可连续调节，可与其他设备配合，进行化学成分、晶体取向测定等一系列优点，从而在失效分析中得到了广泛的应用。

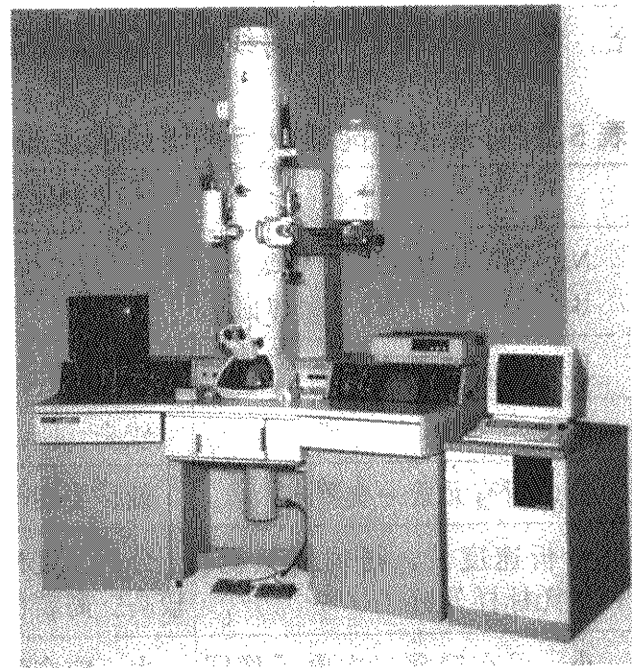


图 5.2-4　透射电子显微镜

(4) X 射线能谱仪

X 射线能谱仪的最大优点是不损伤被测件表面，可同时适用于光滑表面和粗糙断口表面的元素分析，是目前失效分析中应用最广泛的微区成分分析仪器。其分析原理见示意图 5.2-5。

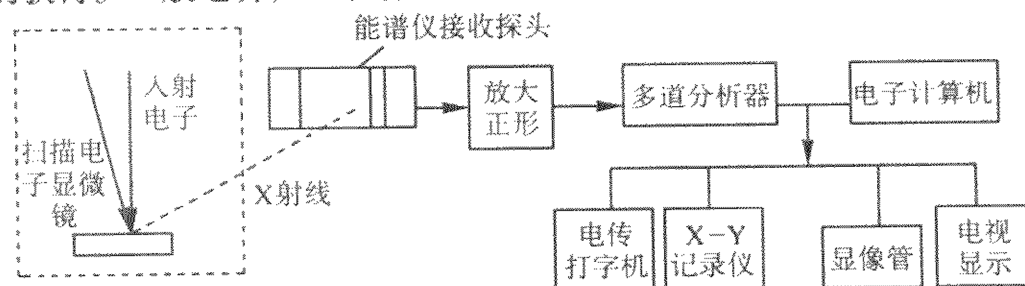


图 5.2-5　X 射线能谱仪原理示意图

不同的元素在高能电子的激发下会产生其特定波长（或能量）的X射线，即特征X射线，X射线微区成分分析就是利用电子束激发样品产生的特征X射线来进行的。

能谱仪是基于X射线的粒子性，将其按照光子能量进行色散的仪器。电子束激发样品产生的X射线被能谱仪探头接收后，经各种处理，按其能量和强度在荧光屏上显示出谱图。根据谱图上波峰所在的位置，可判断该样品中含有哪些元素（即定性分析）。根据波峰的面积，再经过复杂的数学处理，即可计算出各种元素的相对含量（即定量分析）。

X射线能谱仪不仅可以分析某一区域的元素平均成分，还可以分析样品表面某一区域某一元素的分布情况（面分布），也可对某几种元素进行沿指定线路的线分布分析。

由于能谱仪的探头需要在低温下工作和保存，所以其探头必须常年冷却（用液氮或其他方法冷却）。

X射线能谱仪定量分析属于半定量范畴，其测量出来的元素成分数值与背底的扣除技术密切相关，因此分析的误差相对较大。总的规律是重元素的分析误差小，轻元素的分析误差大，尤其是对N以前的轻元素的分析误差大，因此，不能用其代替常规的成分分析。但X射线能谱仪可进行微区的成分分析，对不同区域成分的对比分析可判断是否存在异常。

(5) X射线波谱仪

根据X射线的波动性，按其波长做出X射线强度的谱图，测出元素含量的仪器称为X射线波谱仪。电子探针的原理与波谱仪相同。

在进行微区成分分析时应注意：微区成分分析的结果只能代表分析部位的局部成分，而不能代表样品总体的成分；X射线微区成分定量分析的准确性和样品的制备有关。再者，由于微区成分分析的灵敏度和精确度的限制，其分析结果不能代替用化学分析方法所得的结果。

X射线波谱仪和能谱仪各有自己的优点和不足，它们可以互为补充。表5.2-2比较了X射线波谱仪和能谱仪的优点和不足。

表5.2-2 X射线能谱仪与波谱仪的比较

项目	X射线能谱仪	X射线波谱仪
分析元素范围	Na (11) ~U (92)（铍窗） Be (4) ~U (92)（超薄窗）	Be (4) ~U (92)
分辨率	低，重叠峰多	很高，很少有重叠峰
分析灵敏度	较低	较高
分析效率	高，所有元素一次测出	低，要逐个元素检测
定量分析	分析速度快，精确度较低。可做无标样分析	分析速度慢，精确度较高。需要标样
所需束流	小，可在观察二次电子像同时进行分析	大，应在背散射电子像下分析
其他	需液氮（或其他方法）冷却	—

(6) 铁谱仪和铁谱分析技术

铁谱技术是美国于1971年研制成功的一种对摩擦机械零件进行动态监测的试验检测技术，可用于设备维修、工况监测、摩擦学、故障诊断和失效分析等领域。

铁谱分析技术的基本原理是，不同材料的摩擦副在不同的摩擦条件和润滑条件下，会形成反映摩擦零件的摩擦、磨损和润滑状态的不同形状、大小、颜色和分布的磨损粒子。如对轴承零件，初期疲劳（微裂纹）时的磨损粒子为尺寸在5 μm以下的球形粒子；当零件表面已发生分层剥落或剥皮时，磨损粒子为有一定厚度的平板状粒子。因此，对磨损粒子的形状、大小、颜色等进行鉴别就能分析确定摩擦零件的磨损状态。

运用铁谱分析技术，对摩擦环境中含有磨损粒子的润滑油进行铁谱分析的仪器叫做铁谱仪。铁谱仪的工作原理见示意图5.2-6。铁谱仪工作时，润滑油被从系统中抽取，稀释后沿倾斜的玻璃载片从高处向低处缓慢流动，磁性和几何尺寸各异的磨损粒子被梯度磁场分离开，沉积在玻璃载片的不同部位，形成磨损粒子分布的“谱”。用固定液将该“谱”固定，即形成铁谱载片。在两色显微镜下和扫描电子显微镜下观察玻璃载片，即可得到磨损粒子的形状、大小、颜色和分布，从而诊断出摩擦副所处的磨损状态。

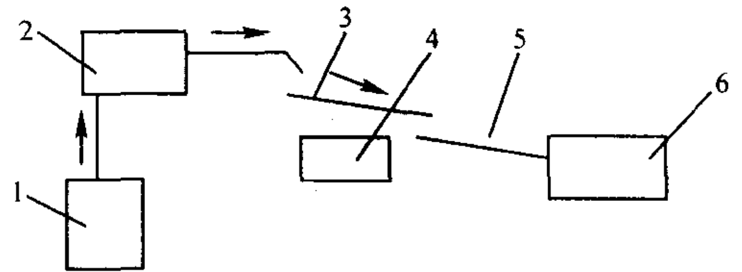


图5.2-6 铁谱仪工作原理示意图

1—润滑油系统；2—泵；3—玻璃载片；4—电磁铁；5—导流管；6—储油槽

1.3 断口形貌诊断技术和方法

断口和裂纹分析一般包括宏观分析和微观分析两个方面。前者系指用肉眼或40倍以下的放大镜、体视显微镜对断口或裂纹进行观察分析，可有效地确定断裂起源和扩展方向，初步判断断口的断裂性质；后者系指用光学显微镜、电子显微镜等对断口或裂纹进行观察、鉴别与分析，可有效地确定断裂类型及机理。宏观分析和微观分析是不可分割的整体，两者不能互相取代，只能互相补充、相互促进。过分依赖微观分析而不重视宏观分析，只根据几个微观视场的特征就作出判断，往往会导致全局性的判断错误；反之，忽视微观分析而只进行宏观分析，则可能得出浮浅的乃至错误的结论。

1.3.1 断口形貌的宏观诊断技术和方法

断口形貌的宏观分析，是指在各种不同照明条件下用肉眼、放大镜和体视显微镜等对断口进行直接观察与分析。断口宏观分析在掌握断口整体形貌特征和规律的基础上来进行进一步的工作，因此，断口的宏观分析一般在放大倍数40倍以下进行。

(1) 断口宏观分析的任务

断口宏观分析的主要任务是：判断断口的基本特征、变形情况和裂纹的宏观走向；确定断裂的类型和方式，为断裂失效模式诊断提供依据；寻找断裂起源区和断裂扩展方向；估算断裂失效应力集中的程度和名义应力的高低（疲劳断口）；观察断裂源区有无宏观缺陷等。

总之，断口的宏观分析可为断口的微观分析和其他分析工作指明方向，奠定基础，是断裂失效分析中的关键环节。

(2) 断口宏观分析的内容

断口宏观分析需要解决如下主要问题：

1) 断口的平直情况和断口的主要特征形貌；

2) 断口的颜色（氧化色、腐蚀产物颜色、夹杂物颜色、光亮情况等）；

3) 断口与主正应力（或主切应力）方向的关系；

4) 断口与成形方向（轧制方向、流线方向）的关系。

以上四方面是断口宏观诊断的主要依据，因此，断口宏观诊断应包括以下六方面的内容：

1) 断裂的位置及其结构特征、周围的工作环境；

2) 断裂位置及其附近的变形程度；

3) 断裂区域的痕迹特征；

4）断裂源的位置、特征及裂纹的走向；

5）断口的宏观形貌特征；

6）断口的颜色及附着物等。

(3) 断口的宏观形貌特征

断口是断裂过程信息的记载，因此，断口特征是构件所经受的应力、环境、材料三要素综合作用的结果。构件的材料不同，经受的环境条件不同，承受的载荷不同，则断口表现出的形貌特征也不相同；反之，断口的形貌特征不同，可以判断出构件所经受的应力、环境和材料三者综合作用的差异。

1）断口三要素　对韧性金属材料一次过载造成的延性断裂，宏观上的基本特征通常表现为三个特征区：即纤维区、放射区和剪切唇区。这三个特征区就是大家常说的断口三要素。图5.2-7是光滑拉伸断口的三要素示意图，图5.2-8是实际的拉伸断口形貌，可见明显的三个特征区。

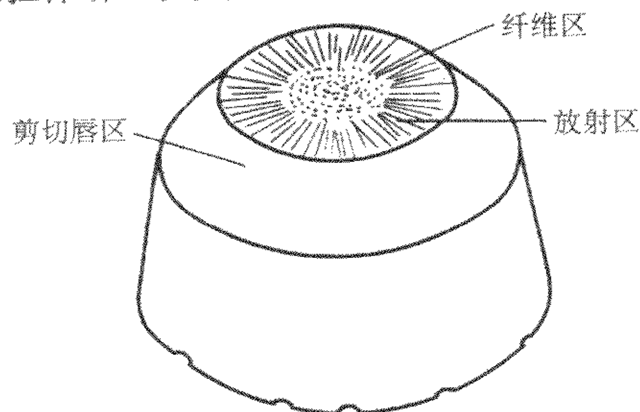


图5.2-7　光滑圆试样拉伸断口三要素示意图

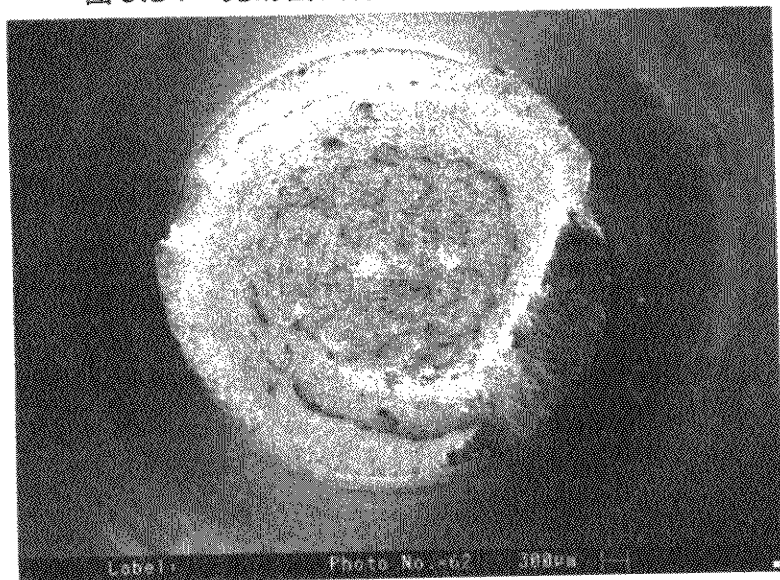


图5.2-8　光滑试样拉伸断口形貌

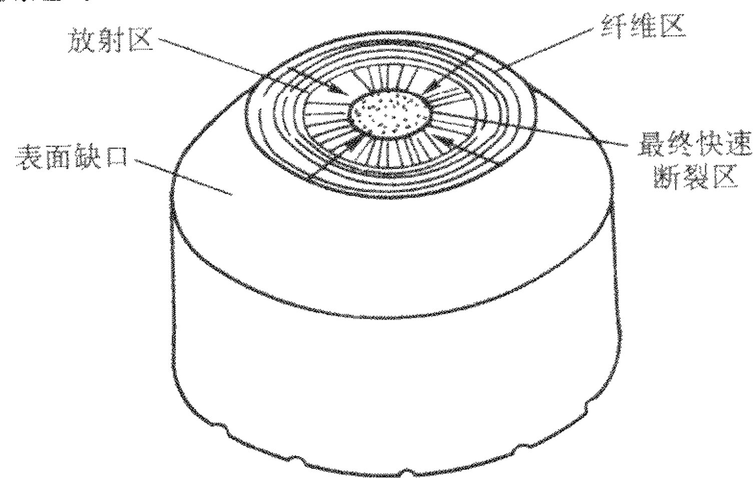
① 纤维区　一般位于断口的中央，是材料处于平面应变状态下发生的断裂，呈粗糙的纤维状，属正断型断裂。纤维区的宏观平面与拉伸应力相垂直，断裂在该区形核。纤维区中心或放射条纹收敛处为裂源位置。

② 放射区　该区紧接纤维区，是裂纹由缓慢扩展向快速扩展的不稳定扩展转化的标志，其特征是放射线花样，放射线发散的方向为裂纹扩展方向。裂纹扩展方向是由纤维区指向剪切唇方向。如是板材零件，放射区的宏观特征为人字条纹，其反方向为裂纹的扩展方向。

③ 剪切唇区　断裂过程的最后阶段出现剪切唇区。剪切唇区表面较光滑，与拉伸应力轴的交角约45°，属切断型断裂。它是在平面应力状态下发生的快速不稳定扩展。

需要注意的是，并不是所有的一次过载断口均具有断口三要素特征，有时只表现出两个要素特征，其中剪切唇区是最后断裂区。

零件形状对断口三要素有很大影响。对圆形光滑拉伸断口，纤维区一般位于断口的中央，放射区紧接纤维区，断裂过程的最后阶段出现剪切唇区。而对带缺口的圆形拉伸试样，断口三要素的分布与光滑圆形试样不同：试样中心部分，基本上是放射区；纤维状区在试样的周围形成环状；裂源在缺口底部萌生，裂纹扩展方向刚好与光滑试样相反，从周围开始向中心扩展；这类断口基本上无剪切唇。图5.2-9是其示意图。矩形板状拉伸试样断口的三要素特征见示意图5.2-10，纤维区呈现椭圆形且位于中心部位，放射区的形状往往为“人字纹”，剪切唇区为矩形框，与自由表面相接。冲击试样断口的三要素特征见示意图5.2-11，纤维区在缺口中部呈半圆形，放射区呈现半轮辐形，剪切唇区在非冲击面的三侧边缘。



黑箭头表示裂纹扩展方向

图5.2-9　缺口圆形拉伸试样断口示意图

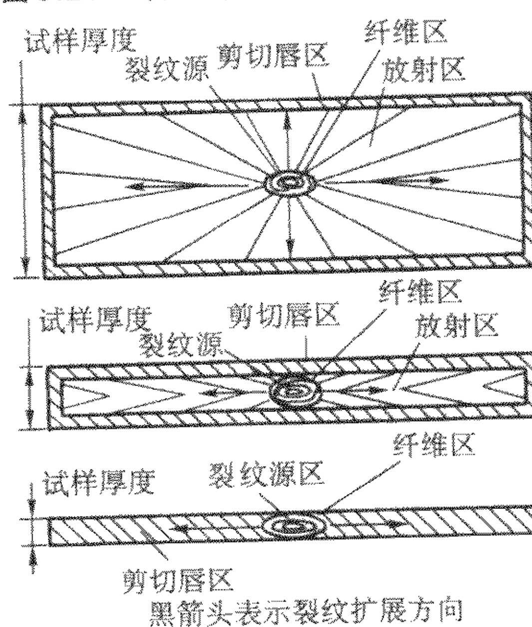
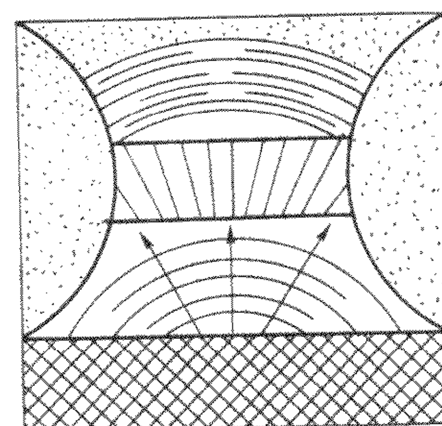


图5.2-10　矩形拉伸试样断口示意图



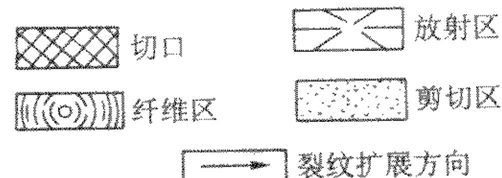


图5.2-11　冲击试验断口特征示意图

2）典型断口的宏观形貌特征

① 断口的宏观分类　断口的分类很多，按其宏观特征，一般可进行如下分类。

按断口上的宏观塑性变形可分为：

延性断裂断口：断口及其附近可见明显的塑性变形。

脆性断裂断口：断口及其附近无明显的塑性变形。

按断裂方式可分为：

正断断口：断口表面与最大正应力垂直。

切断断口：断面与最大正应力方向呈45°。

② 几种典型断口的宏观形貌特征　尽管断口的分类很多，在实际的宏观失效分析中，一般将断口主要分为延性断裂断口、脆性断裂断口和疲劳断裂断口（尽管疲劳断口也属于脆性断裂断口的一种，但由于其比较常见，因此，把其单独列为一种进行讨论）。表5.2-3列出了这三种典型断口的宏观形貌特征，根据这些特征，一般可诊断出断口的宏观类型。

表5.2-3　几种典型断口的宏观形貌特征

断口特征 \ 断口类型	延性断裂断口		脆性断裂断口		疲劳断裂断口	
	切断型	正断型	缺口脆性	低温脆性	低周疲劳	高周疲劳
色泽	较弱的金属光泽	灰色	白亮色，接近金属光泽	结晶状金属光泽	白亮色	灰黑色
断面粗糙度	较光滑	粗糙锯齿状	极粗糙	粗糙	较光滑	光滑
放射线	一般无，但高强钢中有时会出现	无	明显	不太明显	较不明显，板材有近似人字纹	明显，且细腻
弧形线	无	无	无	无	一般可见疲劳弧线，但在恒载时无	一般可见疲劳弧线，应力幅变化越大越明显
与主应力的交角	约45°	约90°	约90°	约90°	裂纹扩展速率小时接近90°（K_{I}），大时接近45°（K_{III}）	约90°

(4) 断口宏观诊断技术和方法

1) 断口观察技术　对断口进行宏观观察时，应注意观察的角度和照明的角度，自然光是最好的照明源，观察时应从不同的角度进行观察。宏观断口分析的第一步是用肉眼观察断口形貌特征及失效件的全貌，包括断口的颜色、变形程度，断口之外的损伤痕迹等，然后对主要特征区用放大镜、体视显微镜等进行进一步的观察，确定重点分析的部位。

2) 断口观察需要重点注意的特征　在断口形貌的宏观分析过程中，要特别注意以下几个特征。

① 断口上是否存在放射花样及人字纹花样。这种特征一方面表征裂纹在该区的扩展是不稳定的、快速的；另一方面，沿着放射方向的逆向或人字纹尖顶可追溯到裂纹源所在位置。

② 断口上是否存在弧形迹线。这种特征表明裂纹在扩展过程中，由于应力状态（包括应力大小的变化、应力持续时间）的交变，断裂方向的变化，环境介质的影响，以及裂纹扩展速率的明显变化等在断口上留下的此种弧形迹线，如疲劳断口上的疲劳弧形等。

③ 断口的粗糙度。一般说来，断口越粗糙，即表征断口特征的花样越粗大，则剪切断裂所占的比例越大；如果断口细平，多光泽，或者花样越细，则晶间断裂、解理断裂所起的作用越大。

④ 断面的光泽与色彩。如，准解理、解理断裂的金属断口在阳光下转动断面进行观察时常可看到闪闪发光的小刻面。如果断口做相对摩擦、氧化以及受到腐蚀时，金属断口的色泽将完全不同。

⑤ 断面与最大正应力的交角（倾斜角）。不同的应力状态、不同的材料及外界环境，断口与最大正应力的夹角不同。例如，在平面应变条件下断裂的断口，与最大正应力垂直；在平面应力条件下断裂的断口，与最大正应力呈45°交角。

⑥ 判定特征区的划分、分布、面积大小等。

⑦ 材料缺陷在断口上所呈现的特征。若材料内部存在缺陷，则缺陷附近存在应力集中，因而在断口上留下缺陷的痕迹。

3) 断口上断裂源位置的确定　虽然大家都知道，零件的表面、次表面、应力集中部位、缺陷处、损伤处等均是断裂最容易起始的部位，但由于受力状态、材料性质、失效模式、环境介质等因素的不同，致使断口表现出不同的形貌特征，从而衍生出不同的花样，给失效分析带来了难度。断口宏观分析的一个重要任务就是要确定断裂起源的位置。在一般情况下，从宏观特征来说，断裂的起始位置一般位于断口上的以下位置：

① 纤维区的中心；
② 放射花样的收敛处；
③ 人字纹的最尖顶处；
④ 断口的平坦区内；
⑤ 无明显塑性变形的区域，或无剪切唇形貌特征区；
⑥ 疲劳弧线的最小半径处；
⑦ 腐蚀氧化最严重的部位；
⑧ 台阶高差最大处。

4) 断裂扩展方向的分析判别　断裂失效分析中，当裂源的位置确定后，其裂纹扩展的宏观方向随之确定。一般裂纹的宏观扩展方向与断口的以下特征方向相同：

① 裂纹源区指向最后断裂区的方向；
② 放射线发散的方向；
③ 纤维区指向剪切唇区的方向；
④ 与疲劳弧线相垂直的放射状条纹分散方向；
⑤ 人字纹的人字张开的方向；
⑥ 断口的平坦区指向斜断口的方向；
⑦ 无塑性变形或塑性变形小的区域指向变形大的区域的方向；
⑧ 台阶高差减小的方向；
⑨ 氧化、腐蚀减轻的方向。

1.3.2　断口微观诊断技术和方法

(1) 断口微观诊断的任务和内容

微观分析诊断是整个断口诊断的重要组成部分，是宏观诊断的补充和完善。特别是在断裂的机理诊断中，微观诊断具有不可替代的作用。断口的微观诊断包括微观形貌诊断和微区成分、结构分析两方面的内容。

断口微观诊断主要包括以下具体内容：

1) 断口周围的塑性变形大小或有无；
2) 断口的边缘锐利情况；
3) 断口与零件形状或应力集中的情况；
4) 断口各特征形貌面积的比例；
5) 断口与晶面、晶向之间的关系；
6) 断口与晶界的关系；
7) 断口与显微组织的关系（是否沿相界面或弱相断裂）；

8）断裂源区的情况（是否有材质缺陷或几何缺陷等）；

9）断口的化学成分、或杂质元素的分布情况；

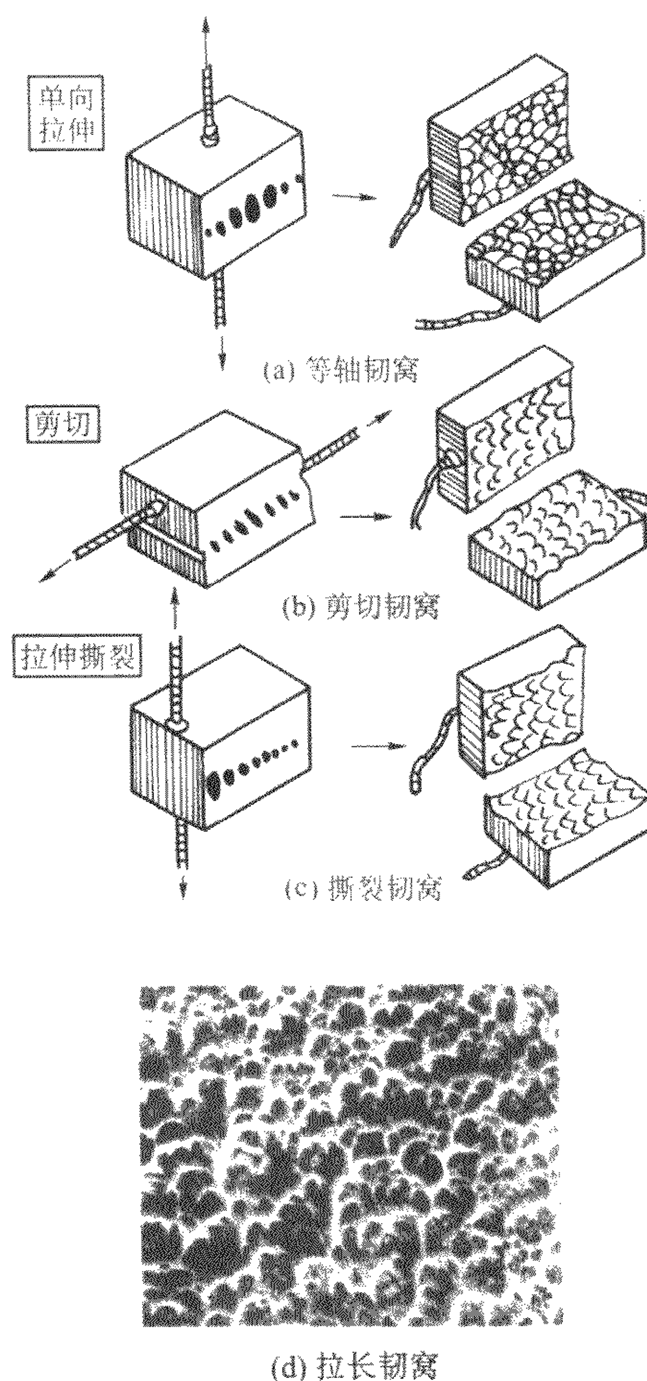
10）断口上二次裂纹的有无或多少、分布情况。

（2）断口的微观形貌特征

在断口上常见的微观特征不外乎韧窝，滑移特征（包括滑移线、滑移带、蛇形花样、涟波花样），解理特征（包括解理台阶、河流花样、舌状花样、扇形花样、鱼骨状花样、瓦纳线），准解理特征，沿晶断裂特征，疲劳断裂特征（疲劳条带、二次裂纹带、韧窝带）等断裂特征花样。下面分别对其进行分析与解释。

1）韧窝特征　韧窝特征是金属延性断裂的主要微观特征形貌，是材料在微区范围内塑性变形产生的显微孔洞经形核、长大、聚集，最后相互连接而导致断裂后在断口表面所留下的痕迹。虽然韧窝是延性断裂的主要微观特征，但其他断裂模式上也可观察到韧窝，因此，不能以韧窝特征作为延性断裂的充分判据，而只能作为必要判据来应用。

随着零件受力状态的不同，韧窝可能有不同的形状，即韧窝的形状可反映零件的受力状态。韧窝的最基本形态有等轴韧窝、剪切韧窝和撕裂韧窝三种，如图5.2-12所示。其中等轴韧窝近似圆形，是在正拉应力作用下形成的；剪切韧窝和撕裂韧窝都属于拉长韧窝，形状呈抛物线形，是在切应力作用下形成的，其中抛物线开口的方向指向切应力作用的方向。当零件受剪切应力作用时，形成的匹配断口上韧窝的拉长方向相反，即韧窝的抛物线开口方向相反；而当零件受撕裂应力作用时，匹配断口上的韧窝拉长方向相同，即韧窝的抛物线开口方向指向同一方向。对于靠第二相强化的材料，往往在韧窝的窝底可看到强化第二相小颗粒。图5.2-13是典型的韧窝特征。



(a) 等轴韧窝

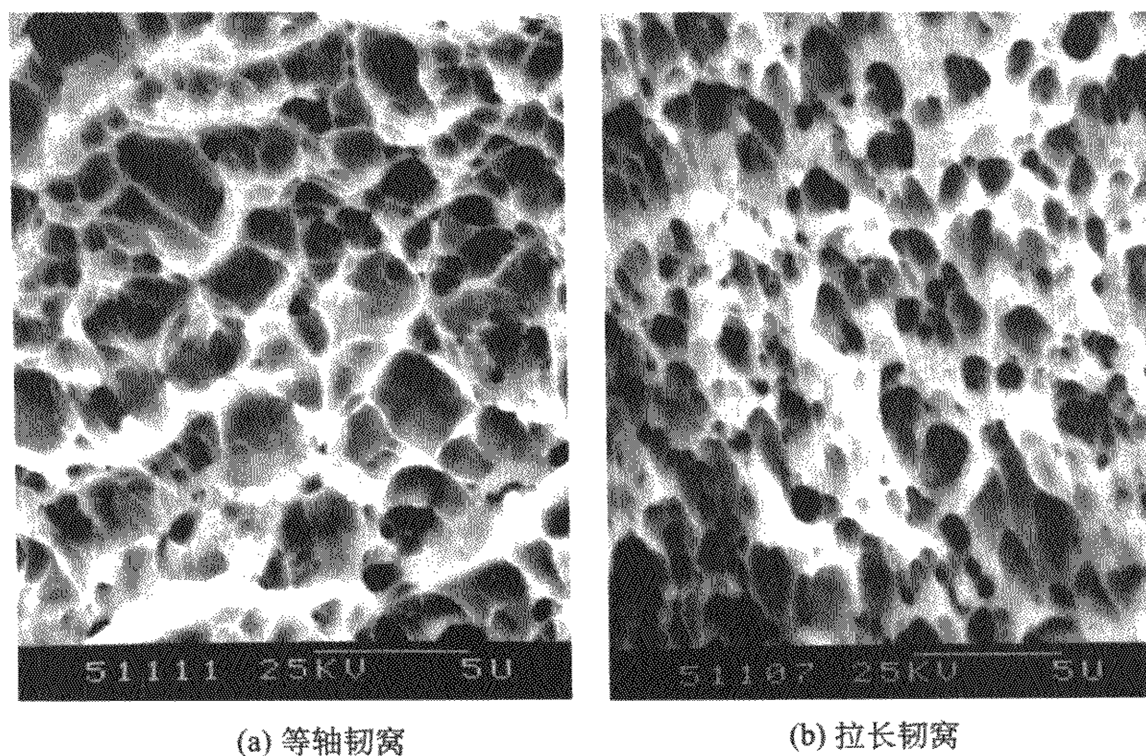
(b) 剪切韧窝

(c) 撕裂韧窝

(d) 拉长韧窝

图5.2-12　韧窝的三种基本形态示意图

虽然韧窝的大小（包括平均直径和深度）受第二相质点的大小、密度、基体的塑性变形能力、变形硬化指数、外加应力大小、状态及加载速率等因素的影响，但可反映材料韧性的好坏。对同一材料的不同形态，相同条件下，韧窝尺寸越大，则材料的塑性越好。



(a) 等轴韧窝　　(b) 拉长韧窝

图5.2-13　典型的等轴韧窝拉长韧窝

一些材料一般情况下的过载断裂和延迟断裂的瞬断区均为韧窝特征。

2）滑移特征　滑移特征包括滑移线、滑移带、蛇形花样和涟波花样，是在正应力作用下，金属沿滑移面滑移分离的主要微观特征，属于金属延性断裂的一种微观特征。而在宏观上，滑移分离的基本特征是断面倾斜，呈45°角；断口附近有明显的宏观塑性变形。

① 滑移线（或滑移带）　金属材料的塑性变形方式主要有滑移、孪生、晶界移动和扩散蠕变等四种，其中滑移变形是最重要的、最常见的变形方式。按照金属学理论，金属内部存在有大量的位错，在外力作用下，这些位错将沿滑移面移动（滑移），当位错滑移到晶体表面时，将在表面形成台阶或滑移线痕特征—滑移线；大量的滑移线或滑移台阶组成的条带就称为滑移带。图5.2-14是典型的金属表面滑移线（滑移带）特征。实际上滑移线（或滑移带）是高倍下的一些线痕特征，由于金属滑移总是在最密排的晶面上、沿最密排的方向滑移，因此在单晶材料上，滑移线只在一组平行晶面上的特定晶向上产生，故表现出平行、平直的特征。对

多晶材料，滑移总是在多个滑移系中同时进行，因此，有多个滑移面参加滑移，位错之间互相切割，致使滑移线不再是平直的直线，而是弯曲的折线或交叉线。

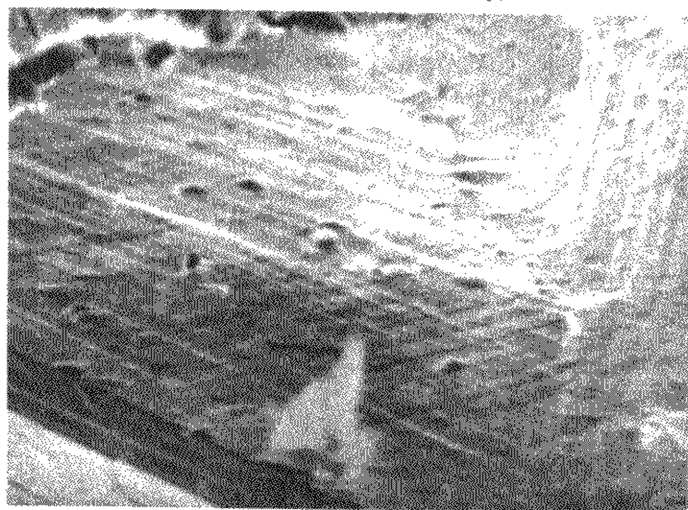


图 5.2-14　金属表面滑移线（滑移带）特征

② 蛇形花样　蛇形花样属于滑移的一种特征。当多晶材料塑性变形较大时，会产生交滑移，导致的滑移面分离断口上可观察到起伏弯曲的、类似蛇形的滑移线特征，称为蛇形滑移花样。

③ 涟波花样　当多晶材料的塑性变形程度进一步加大时，蛇形花样就会因变形而平滑化，形成较平的涟波花样。

实际材料中的缺陷、裂纹、孔洞、第二相粒子附近均经常可观察到这种滑移特征，有时在韧窝的侧表面上也可观察到这些滑移特征（见图 5.2-15）。



图 5.2-15　气孔边缘涟波花样

3）解理特征　金属在正应力作用下，由于原子结合键破坏而造成的沿一定晶体学平面（解理面）快速分开的过程称为解理断裂。解理断裂属于脆性断裂的一种，解理面通常是表面能量最小的晶面，不同的晶体结构具有不同的解理面，如体心立方结构的解理面为{100}，密排六方结构的解理面为{0001}，金刚石晶体的解理面为{111}；而面心立方晶系的金属一般不发生解理断裂。

解理断裂区宏观上没有明显的塑性变形，在太阳光下转动时可观察到反光的小刻面，属于脆性断裂。严格意义上说，解理断裂面上是没有任何解理特征花样的，但在实际材料中，由于各种因素的作用，解理面局部均会发生微观的塑性变形，从而形成解理台阶、河流花样、舌状花样、鱼骨状花样、扇形花样及瓦纳线等特征，它们是解理断口上常见的典型微观形貌。图 5.2-16 是典型的解理断裂断口特征形貌。

① 解理台阶　解理台阶是平行的、具有高度差的不同解理面断裂、汇合而形成的具有高度差、类似台阶的特征。如图 5.2-17 所示。解理台阶形成的机理主要有两种：一是解理裂纹与螺位错相交截而产生一高度为柏氏矢量的解理台阶；另一个机理是两个相互平行的、处于不同高度上的解理裂纹通过次生解理或撕裂方式互相连接而形成台阶。

② 河流花样　河流花样的形貌特征类似河流，由多个支流起源，汇合形成干流的特征。图 5.2-18 示出了河流花样的形成过程。河流花样也是解理断裂的一种重要的微观形貌特征。解理断裂在扩展过程中，为了减少所消耗的能量（阻力），处于不同平面的解理裂纹将逐步汇合，以减小撕裂面积，从而形成河流花样，即由较小台阶（支流）汇合到较大的台阶（较大的支流或干流）；河流的流向与裂纹扩展方向一致。因此，根据河流花样的流向，可判断裂纹微观区域上的扩展方向。

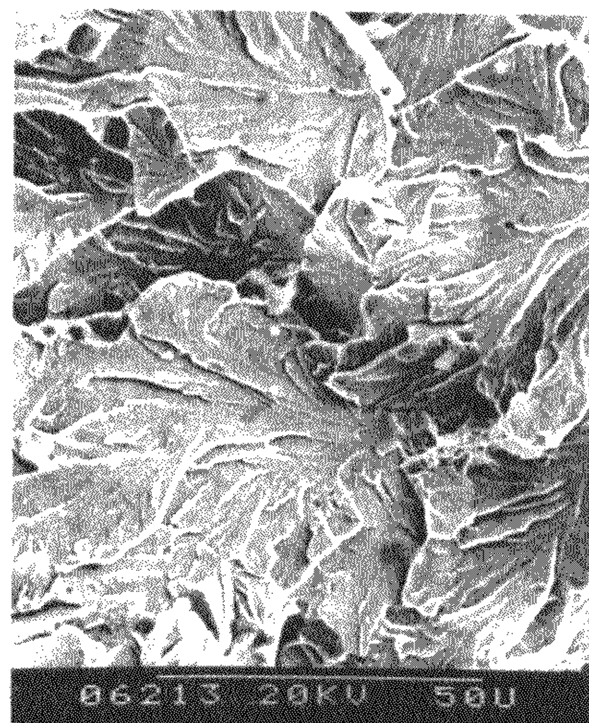


图 5.2-16　典型解理断裂断口特征形貌

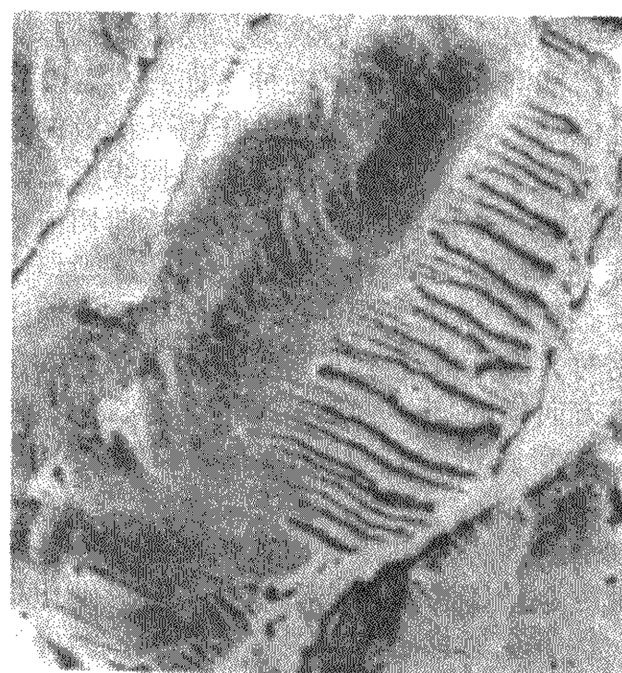


图 5.2-17　解理台阶



图 5.2-18　河流花样的形成示意图

在实际材料中，可观察到河流花样在裂纹扩展过程中存在有很多的变化。当解理裂纹穿过晶界或亚晶界时，河流花样将发生变化。穿过扭转晶界时，河流将在晶界重新形成，即重新出现大量的支流，从而呈现河流激增的现象；裂纹与小角度倾斜晶界相交时，河流会连续穿过晶界而不发生变化；当河流穿过普通大角度晶界时，因相邻晶粒的位相差大，河流也会重新形成，而出现大量的河流。

③ 舌状花样　舌状花样也是解理断裂的典型微观特征之一，其形貌类似"舌头"状，故得名"舌"状花样。图 5.2-19 是其典型形貌。舌头状花样是解理裂纹沿变形孪晶与基体间的界面扩展形成的、或是解理裂纹在扩展中局部发生

二次解理或滑移分离，或二次解理和滑移分离的混合。



图5.2-19 解理舌特征

④ 扇形花样 扇形花样也是解理断裂的典型微观特征之一，是裂纹从晶界或晶界附近起始、向外呈扇形扩展的一种形貌特征，其外观类似扇子形或羽毛形状。

⑤ 鱼骨状花样 鱼骨状花样是一种类似鱼的骨头形状的断裂形貌特征（见图5.2-20），也是解理断裂的典型微观形貌特征。

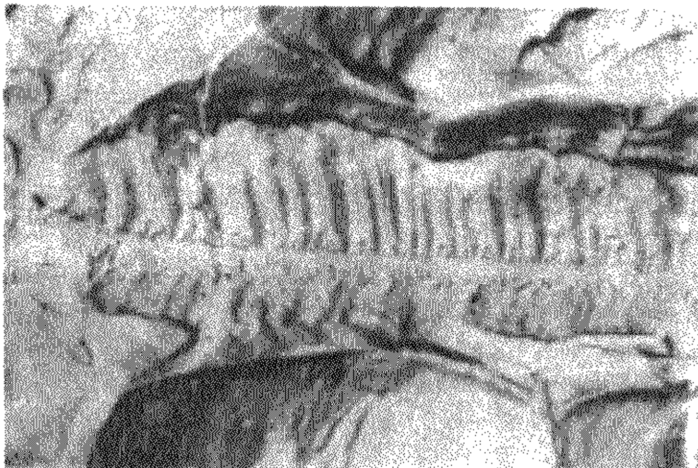


图5.2-20 鱼骨状花样

⑥ 瓦纳线特征 瓦纳线是在极脆的材料中，在解理断裂面上出现的类似宏观断口上看到的“人字纹”的线痕特征。它是裂纹扩展时，裂纹前沿与弹性应力波相互作用下出现的。目前只在粗晶粒的 Al－Ni－Co 永磁铁合金和 Fe－25Cr－5Al 中发现过。

4）准解理断裂特征 准解理断裂特征是介于解理断裂与延性断裂之间的一种过渡断裂形式，其宏观上基本无明显的塑性变形（或变形较小），断口平整，具有脆性断裂特征；其微观形貌有河流花样、舌状花样及韧窝与撕裂棱等。图5.2-21是典型的准解理形貌特征。

准解理断口与解理断口既有相似，又有不同，两者的对比情况见表5.2-4。

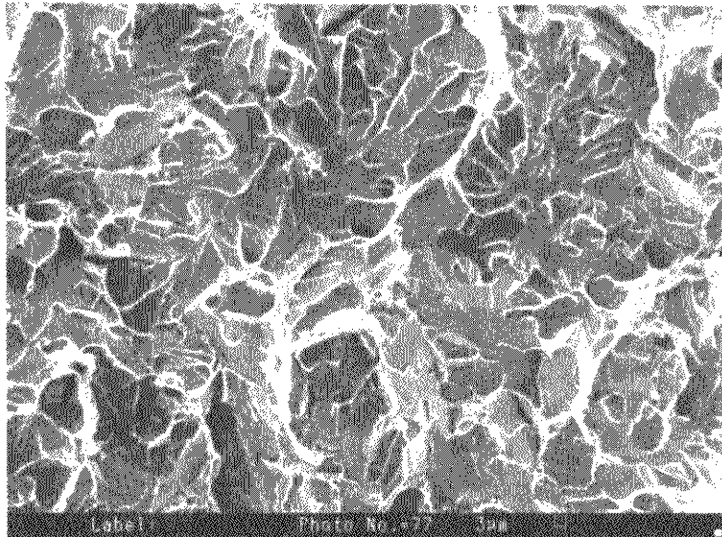


图5.2-21 准解理断裂微观形貌

表5.2-4 解理断口与准解理断口的形貌特征对比

	内容	解理断口	准解理断口
相同处	宏观形貌	无宏观塑性变形（或变形较小），断口平整，呈脆性断裂特征	
	微观形貌	河流花样、舌状花样	
不同处	源区位置	起源于晶粒边界或相界面上	起源于晶粒内部的孔洞、夹杂、第二相处
	裂纹扩展方向	由晶界向晶内扩展	由裂源向四周扩展，呈菊花状
	裂纹扩展平面	沿解理面扩展	不严格沿解理面扩展
	尺寸	在调质钢中，解理面相当于回火马氏体尺寸	在调质钢中，准解理面相当于原始奥氏体晶粒尺寸
	微观形貌	没有韧窝撕裂特征	局部有韧窝、撕裂特征

5）沿晶断裂特征 沿晶断裂特征属于脆性断裂的一种，又称为晶间断裂，是多晶材料沿晶粒界面分离，即沿晶界面发生断裂的现象。图5.2-22是沿晶断裂的典型形貌特征。

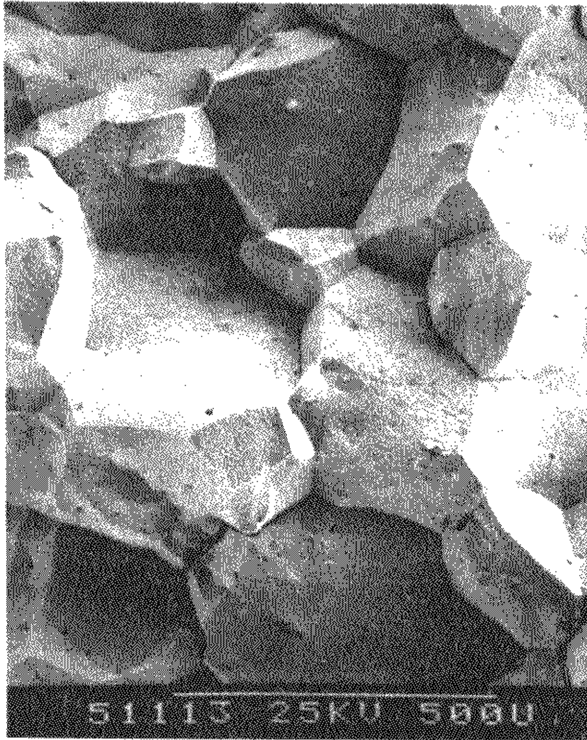


图5.2-22 沿晶断裂微观形貌

一般情况下，在等强温度以下，晶界的强度即键合力要高于晶内，因此，断裂扩展的路径应是穿晶的而不是沿晶，以使裂纹扩展的阻力最小。当杂质元素在晶界富集、沿晶界析出脆性第二相、环境介质沿晶界浸入金属基体、温度超过等强温度等情况时，则可使晶界弱化，晶界的键合力被削弱，从而使裂纹容易沿晶界扩展。扩展过程中，裂纹均沿与局部应力垂直的晶界扩展。

沿晶断裂又可分为沿晶韧窝断裂和沿晶脆性光面断裂。沿晶韧窝断裂是在沿晶面上可看到细小的浅韧窝，是由晶界沉淀的分散、细小颗粒作为韧窝核，然后以剪切方式形成孔洞，最后孔洞连接形成细小的裂纹而分离。沿晶脆性断裂沿晶面上光滑、干净，无微观塑性变形特征，呈冰糖块状形貌，如图5.2-22所示。回火脆、氢脆、应力腐蚀、液体金属致脆以及因过热、过烧引起的脆性断裂大都为沿晶光面断裂特征；而蠕变断裂等往往为沿晶韧窝特征。

沿晶面上具有线痕（鸡爪痕）特征的沿晶断裂形貌是氢脆断裂的典型形貌特征，见图5.2-23，其中沿晶面上的线痕（鸡爪痕）是一种撕裂痕迹特征。沿晶面上具有核桃纹特征

的沿晶断裂是应力腐蚀断裂的典型形貌特征，见图 5.2-24。液体金属致脆的沿晶面上一般可看到致脆的金属残留痕迹。

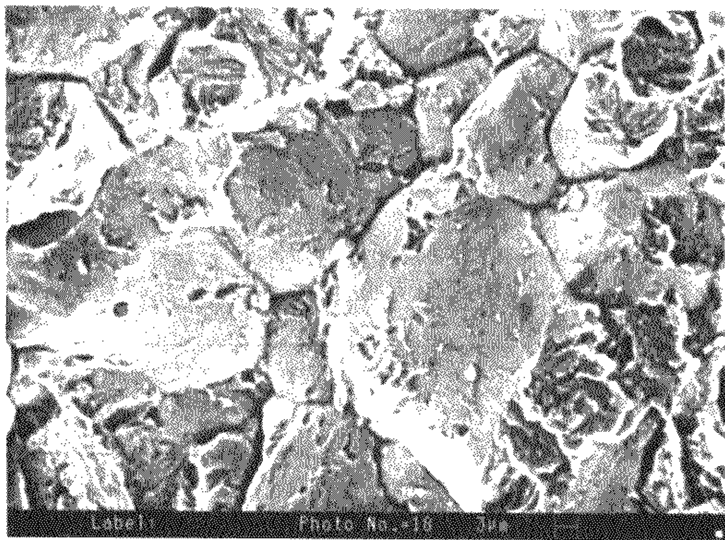


图 5.2-23 沿晶面上线痕特征

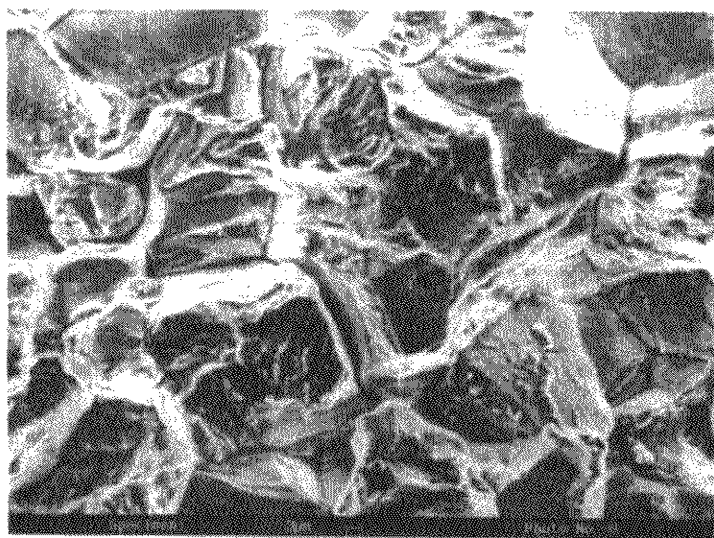


图 5.2-24 沿晶面上核桃纹特征

6）疲劳断裂特征 疲劳断裂过程可分为疲劳裂纹萌生、稳定扩展和失稳扩展三个阶段。疲劳条带是疲劳裂纹扩展第二阶段的最重要的显微特征，也是疲劳断裂断口的基本形貌特征，有人称之为疲劳辉纹。图 5.2-25 是典型的疲劳条带形貌特征。凡在断口上具有疲劳条带特征的，即可判断该断口为疲劳断裂；但断口上没有疲劳条带的，也不能排除其是疲劳断裂的可能，意即疲劳条带是诊断断口为疲劳断裂的充分判据，但不是必要判据。

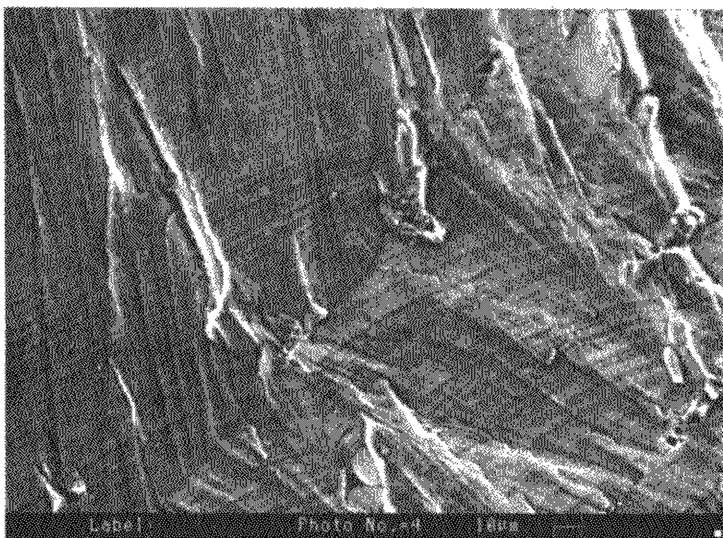


图 5.2-25 典型疲劳条带特征

疲劳条带在形貌上一般具有如下主要特征：

① 疲劳条带是一系列基本上相互平行的、略带弯曲的波浪形条纹，并与裂纹局部扩展的方向垂直；

② 疲劳条带间距随应力强度因子幅的变化而变化，一般离源区越远，条带的间距越大；

③ 一般源区附近条带特征不如扩展中后期的明显；

④ 匹配断口两侧的疲劳条带特征基本对应；

⑤ 局部区域的条带扩展方向与裂纹的宏观扩展方向可以相同，也可以不同，甚至可能出现相反的情况；

⑥ 疲劳断口通常由许多大小不等、高低不同的小断块所组成，各断块上的条带不连续，且不平行；

⑦ 每一条带代表一次应力循环；

⑧ 条带表示应力循环时裂纹尖端的位置。

一般韧性材料容易形成疲劳条带，而脆性材料则比较困难。

疲劳条带根据其形貌特征又可分为延性疲劳条带和脆性疲劳条带。脆性疲劳条带的断裂路径呈放射状扇形，条带被放射台阶分割称短而平坦的小段。材料本身较脆、腐蚀环境、应力交变频率较低及应力强度因子幅 ΔK 较低的情况下，容易形成脆性疲劳条带。延性疲劳条带长而起伏，不被放射台阶分割，连续性好。在高的应力强度因子幅 ΔK 下形成的条带大多属于延性疲劳条带，韧性较好的材料上也容易出现延性疲劳条带。

在疲劳断口上除了常见的疲劳条带之外，有时还可见到轮胎花样。轮胎花样貌似汽车轮胎在泥地上辗压后留下的花样而得名，是由于疲劳裂纹的两边断口在循环应力的作用下，作规则的反复接触（辗压）时，一侧的棱边或硬质点在相对一侧留下的痕迹。这种痕迹随着疲劳裂纹的扩展，两侧匹配面间的残留应力逐渐松弛，或为制约裂纹的同时扩展而发生应力的再分配，使得匹配断口发生逐渐的相对位移，于是痕迹就排列成行，形成轮胎痕迹。

对韧性材料，当疲劳应力强度因子幅 ΔK 很大时，疲劳断裂的循环周次可能很少，此时断口上的微观形貌可能为排列规则的平行韧窝带，如图 5.2-26 所示，其基本特征为韧窝，但这些韧窝呈带状、平行分布。这也是疲劳断裂的一种微观形貌特征，一般在韧性较好的大应力低周疲劳断裂断口上容易看到。

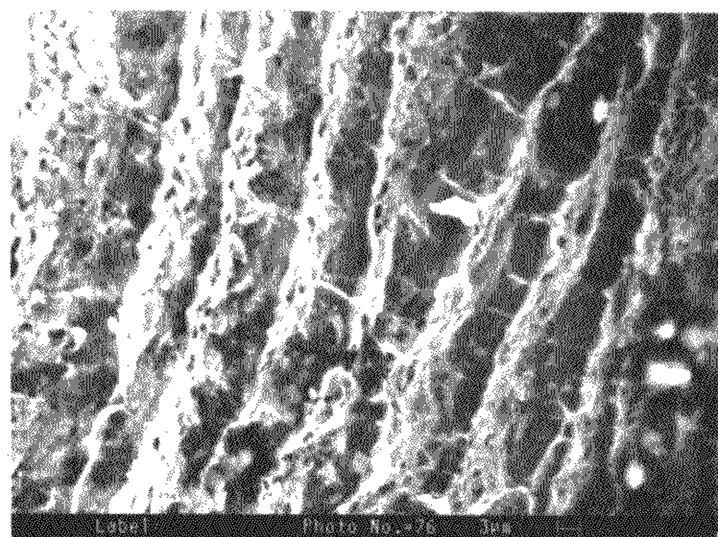


图 5.2-26 疲劳韧窝带特征

对高强韧性较差的材料，如 30CrMnSiNi2A，其疲劳断口上一般观察不到疲劳条带特征，而只能观察到平行的多条二次裂纹特征，如图 5.2-27 所示。这种平行的二次裂纹带也是疲劳裂纹扩展第二阶段常见的一种微观形貌特征。

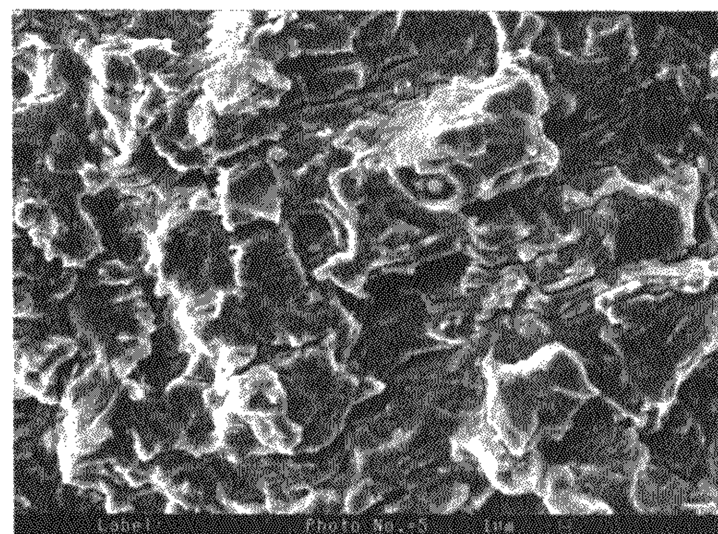


图 5.2-27 疲劳二次裂纹特征

（3）断口微观诊断技术

1）断口的微观形貌观察技术 断口的微观观察一般是

借助扫描电子显微镜、透射电子显微镜等工具来完成的，由于放大倍数较高，在电子显微镜下一般难以准确确定观察的位置，特别是用透射电子显微镜进行观察时，还需要制备复型样品，更加难以确定观察到的形貌特征的位置，因此，应在微观分析之前，熟练掌握断口的宏观形貌特征，并以此来指导微观观察。具体的电子显微镜观察技术已在断口诊断仪器设备中详细讨论过，在此不再重复。

2）断口的微区成分分析技术　断口的微区成分分析是断口微观诊断的重要组成部分，是微观形貌诊断的补充，对诊断失效原因和机理具有重要的指导作用。

断口的微区成分分析一般是根据需要来选择进行的，并不是断口上的所有区域均要进行微区成分分析。选择微区进行成分分析的原则是：

① 微观形貌异常区域；

② 裂纹起始区（源区）；

③ 断口上覆盖有外来物特征；

④ 断口上覆盖有腐蚀产物的区域；

⑤ 对基体成分进行定性判断时，应选择没有污染的瞬断区进行分析。

进行微区分析时，应根据特征区域的大小，合理地选择分析区域大小，如面区域或点区域。为了分析某一个或一些元素在某一区域的分布情况，可进行元素的面分布分析；为了确定某一元素沿某一特征线的含量变化情况，可进行元素的线分布分析。

3）断裂扩展的微观方向判别　解理与准解理断裂微观扩展方向的判别：

① 河流花样的合并方向就是解理裂纹扩展方向；河流花样的发散方向为准解理裂纹局部扩展方向；

② 在解理或准解理的显微断面上，扇形或羽毛状花样的发散方向为裂纹的局部扩展方向。

疲劳裂纹微观扩展方向的判别：

① 与疲劳条带相垂直的方向为裂纹局部扩展方向；

② 轮胎花样间距增大的方向为疲劳裂纹局部扩展方向。

1.4　断口定量诊断技术和方法

在失效分析中，有时不仅需要对断口进行定性诊断，而且需要对断口进行定量分析，以推断材料性能及导致断裂失效的一些基本参数。同时，断口的定量分析也可为确定零件的安全寿命与检修周期提供科学依据。断口定性诊断是定量分析的前提，而定量分析则是断口定性分析的深入和发展。

断口定量分析涉及的领域非常广，内容很丰富。金属断口的定量分析主要包括断口表面的成分、结构和形貌特征的定量参数的描述和表征等内容。

1.4.1　断口表面的成分分析

断口表面的成分定量分析是指对断口表面的平均化学成分、微区成分、元素的面分布及线分布、元素沿深度的变化、夹杂物及其他缺陷的化学元素比等参数进行分析和表征。

在断口分析中，断面或表面微区成分分析很重要，尤其是对于腐蚀、夹杂、成分偏析或外物损伤等因素造成的断裂。这类分析仪器主要有：俄歇电子谱仪、离子探针、电子探针、X 射线能谱仪、X 射线波谱仪等。由于俄歇电子谱仪、X 射线能谱仪、X 射线波谱仪可装配于扫描电子显微镜上，在断口微观形貌观察的同时进行微区成分分析，因而得到了广泛的应用，其中尤以 X 射线能谱仪的应用最为普及。

X 射线能谱仪一般与扫描电镜或透射电镜同时使用，收集电镜中高能电子轰击样品时产生的特征 X 射线，通过对其能量进行分析，便可知样品分析部位的成分。图 5.2-28 是能谱仪分析的典型谱线图。能谱仪可进行定性和定量分析，具有使用方便、速度较快、不损伤样品、对样品无特殊要求、可在观察微观形貌的同时进行成分分析等优点，是失效分析中最常用的微区成分分析仪器。其分析区域可从 1 微米到几毫米，分析元素范围为钠 Na（11）到铀 U（92）（有些能谱仪分析元素范围可为铍 Be（4）到铀 U（92））。

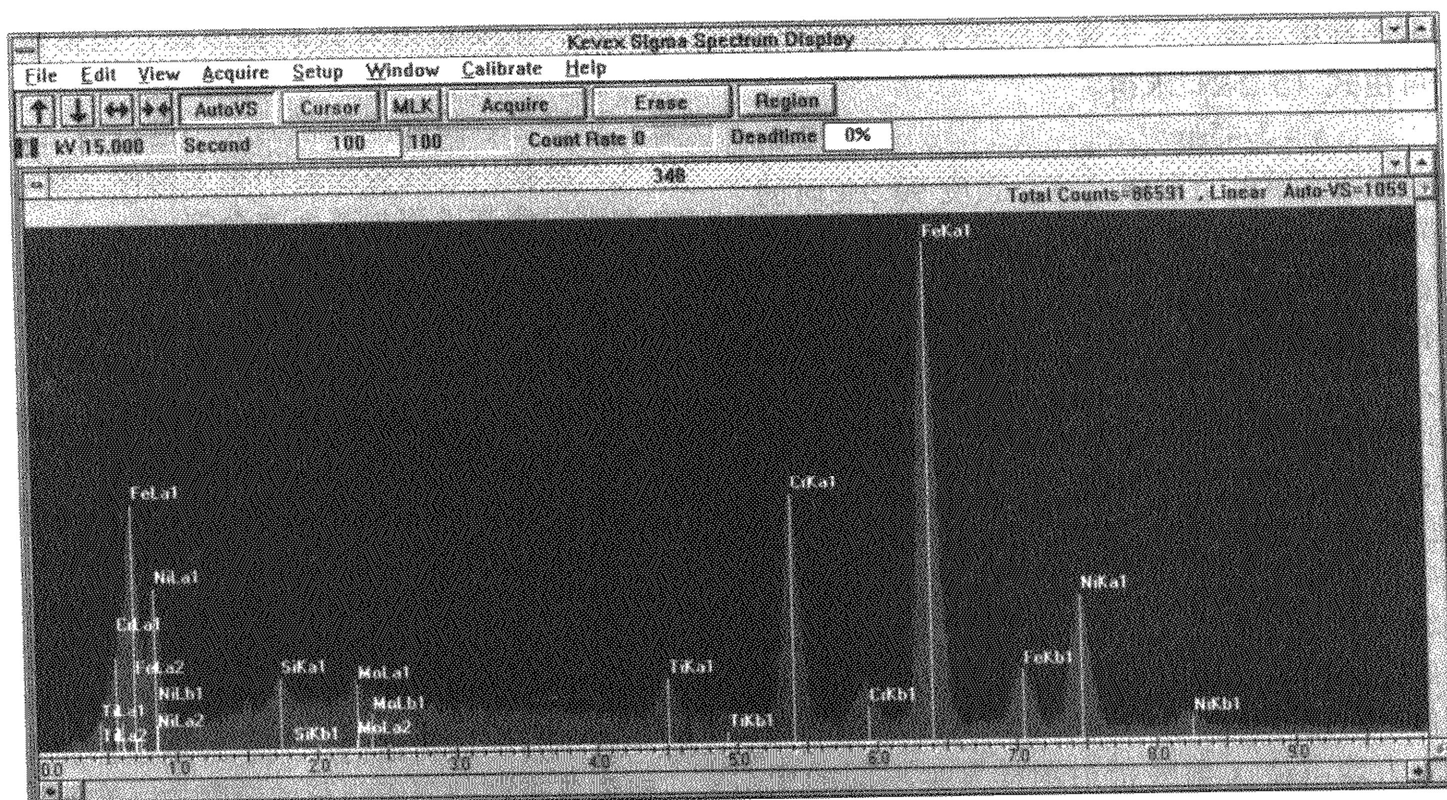


图 5.2-28　不锈钢的能谱仪谱图

与 X 射线波谱仪和电子探针相比，X 射线能谱仪操作方便、分析速度快，适于粗糙及平整表面，这是其优点；但缺点是分辨率和探测灵敏度及分析精确度较差，分析轻元素困难。

不同元素有不同的 X 射线特征波长，X 射线波谱仪就是通过探测 X 射线的特征波长来进行成分分析的。它可分析从铍 Be（4）到铀 U（92）的元素。X 射线波谱仪的突出优点是分辨率很高，但为了精确聚焦，需将样品上的分析点的几何位置严格处于聚焦圆上，限制了其使用；同时，在低速流和低激发强度的情况下还难以使用，这是影响其广泛应用的两个致命弱点。

俄歇电子谱仪可分析 5 个原子层以内厚度的极薄层上从锂 Li（3）到铀 U（92）的元素，弥补了其他分析仪器不能分析超轻元素的不足，但由于其采用的是剥层分析法，会破坏样品表面的结构状态，因而也限制了其使用范围。

1.4.2 断口表面结构分析

断口表面结构定量分析的对象是断口所在面的晶面指数、断口表面微区（夹杂、第二相等）的结构和残余应力。目前，主要的分析仪器是X射线衍射仪。

1）点阵参数测定　任何结晶物质在一定状态下均有一定的点阵参数，当外界条件（如温度、压力、其他外加因素）改变时，点阵参数也将发生相应的改变。通过测量点阵参数，可解决金属学及热处理中存在的很多问题，如固溶体的类型、宏观应力的测定、表面层的分析等。

点阵参数一般用X射线衍射法来测定。这是一种间接的测定方法，首先按衍射花样求出待测金属某一晶面（HKL）衍射线的布拉格角 θ，然后通过布拉格方程求出相应的晶面间距，再根据晶面间距求得点阵参数。对立方晶系，点阵参数 a 有如下计算公式：

$$a = \lambda(H^2 + K^2 + L^2)^{1/2}/(2\sin\theta) \qquad (5.2\text{-}1)$$

2）物相分析　每一种结晶物质，都有其特定的结构参数，包括点阵类型、晶胞大小、单胞中原子（离子或分子）的数目及其位置等，这些参数在X射线的衍射花样上均有所反映，而且衍射花样像人的指纹一样，没有两种物质的衍射花样是完全相同的，因此，可根据衍射花样来鉴别物相。几种物相混合后，其所得的衍射线条是各个物相衍射结果的简单叠加，据此，可以将混合物中的各个物相分别出来。同时，各相衍射线的强度随该相的含量增加而提高，因此，又可对物相进行定量分析。

需要注意的是，若多相混合物中某个物相的含量过少时，该物相有可能不能产生自己完整的衍射图样，甚至不出现衍射线条。一般来说，由重元素组成的物相较易出现衍射图样；物相结构简单，其衍射线条也容易出现；当物相的晶粒过细或存在显微应力时，衍射图样难出现。

由于X射线衍射分析的是表面一定深度内的金属，因此，被分析样品需要有一定的厚度。金属零件的化学热处理层、氧化层、电镀层等往往由于太薄而可能使其中某些相分的线条不能在衍射图样中出现。至于被分析层究竟需要多厚才能符合要求，可由式（5.2-2）求得。

$$x = 3.45\sin\theta/\mu_1 \qquad (5.2\text{-}2)$$

式中，x 为试样表层所贡献的衍射强度达总强度99.9%时的表层厚度；θ 为掠射角；μ_1 为材料的线吸收系数。

当某些相所占的体积分数很小，如钢中的碳化物、沉淀相、夹杂物等，将钢直接进行X射线衍射一般难以发现这些相分的衍射线条。此时，需事先将这些相电解分离下来，即选择恰当的溶液和电流强度，使基体相溶解，让欲分析的微量相沉淀下来。

3）应力测定　应力按其存在的范围大小可分为三类：宏观应力、微观应力和超微观应力。宏观应力是指存在于整个工件或较大区域（宏观区域）并取得平衡的应力；当应力只在晶粒范围内取得平衡时，该应力称为微观应力；存在于晶界、滑移面、位错附近等更微小区域中的应力称为超微观应力。由于各类应力均可使X射线衍射图样发生不同的变化：宏观应力使衍射线位移，微观应力使衍射线变宽，超微观应力使衍射线强度减弱。因此，通过测量这些变化，可区别出应力的类型和数值。目前已有专门的残余应力测定仪供选用。

用X射线法测定应力具有以下优点：

① 属于非破坏性的检验方法；

② 可测定表层（10～35 μm内）的应力，这对测定镀层、涂层、氧化层、表面热处理层的应力非常有效；

③ 可测局部小区域的应力，可测应力在表面的分布及应力沿层深的分布；

④ 可测宏观残余应力、微观应力和超微观应力；

⑤ 可测定复相中各个相的应力。

在测定时，需要注意试样状态和形状对测定结构的影响。由于X射线可贯穿的深度很小，衍射反映的只是试样表层的晶体结构，因此，若试样表面存在污垢、氧化皮、涂层、加工硬化层等，则不能得到真实的应力值。粗糙表面由于凸出部分有应力释放、而且这些部分对衍射的贡献较大，从而使得所测应力值偏低。晶粒过大，参与衍射的晶粒数目减少，使得衍射强度不稳定，衍射峰异常，峰值不能准确确定，所测应力值不可靠，重复性差，一般要求晶粒直径在30 μm以下；晶粒过小，使衍射线变宽，测量的精度下降。对有择优取向（织构）的材料，其衍射峰强度强烈依赖于试样表面法线与反射晶面法线的夹角，需合理选择此夹角为入射角，并将其固定，以获得最大衍射强度。对具有不同表面曲率半径的试样，应通过变化X射线束的宽度来测定。

1.4.3 断口形貌特征的定量分析

（1）断口形貌特征定量分析的内容

断口形貌特征的定量分析是对断口表面的各种花样，包括各种断口特征花样区域的相对大小及材料组织、结构、性能及导致发生断裂的力学条件、环境条件之间的相互关系进行分析。目前断口形貌定量研究进行得较好的是疲劳断口的定量分析，利用的主要参量是断口上的宏观疲劳弧线和微观疲劳条带，但前提是能够对疲劳弧线或疲劳条带的间距进行测量。

材料和结构的疲劳破坏是一种渐变的过程，当多次重复变化的载荷作用后，虽然应力值始终没有超过材料的强度极限，甚至比弹性极限还低的情况下，就可能发生破坏，这种在交变载荷作用下材料或结构的破坏就是疲劳破坏。由于疲劳破坏是一种渐变的过程，疲劳裂纹是逐步扩展的，因此，在疲劳断口上就会留下扩展过程的痕迹。通过对各种痕迹特征的分析，就有可能反推裂纹扩展的过程。所谓疲劳断口定量分析，主要是指对其断裂形态，包括疲劳源点、疲劳弧线、疲劳台阶、疲劳条带等的位置、数量、间距及疲劳源数量、扩展区、瞬断区面积的大小进行量化、测量与计算，并据此来反推引起疲劳断裂的相关构件及断裂过程的顺序、应力大小、疲劳扩展与萌生寿命等。

如果说在疲劳试验后，断口定量分析工作仅是作为试验的一种重要补充，那么在疲劳失效分析中它则具有特殊的意义。因为通过对疲劳断口的定性、定量分析，往往可以找到失效的原因。特别是根据裂纹扩展规律反推出的初始裂纹长度 a_0 更是损伤容限设计所关心的，以它计算出的剩余疲劳寿命，可用来确定构件使用中的安全检查周期。

（2）疲劳断口定量分析的理论依据

利用断口上疲劳弧线和疲劳条带进行定量分析的理论依据是：宏观上的每一条疲劳弧线相当于裂纹扩展过程中载荷或应变发生一次大的改变，如发动机的一次起动循环、飞机的一次飞行等载荷谱的加载、环境条件的改变；微观上的每一疲劳条带则相当于载荷或应变的一次循环，如发动机叶片振动一次等。由于疲劳断口在不同的放大倍数下可观察到不同形貌和数量的疲劳条带，因此准确地确定何种疲劳条带与何种载荷之间具有对应关系是断口定量分析的关键所在。

钟群鹏等人通过对疲劳断裂形貌特征和扩展区、瞬断区物理数学模型的研究，证明疲劳断口的形貌特征及扩展区、瞬断区的大小与疲劳应力幅之间具有定量关系。

（3）疲劳断口定量分析的现状

疲劳断口是疲劳破坏过程的见证，断口上的特征真实地记录了疲劳裂纹扩展的过程。虽然人们对断口的定量分析可

追溯到很早以前，现今的断口定量分析也已有了一定的发展，然而迄今为止的断裂失效分析还基本上是以定性为主，研究还主要集中在疲劳机理、规律、寿命预测等，对疲劳断口特征的定量研究还相对较少，对从断口研究疲劳的寿命更少。这一方面是由于断裂构件的复杂性，另一方面也受其他学科发展水平的制约。由于从断口反推寿命的可靠性大，人们已经认识到了它的重要性，开始了疲劳断口的定量反推计算研究。由于实际失效件载荷的不确定性和疲劳失效过程的复杂性，应力复杂，超载迟滞、低载加速效应难以识别，使用环境千差万别，时常有腐蚀介质的影响，使对实际失效件的断口定量反推还没有取得理想的突破，目前的研究主要集中于失效过程已知的试验件断口上。

近年来进行的疲劳断口寿命估算及反推结构原始疲劳质量，都是用定量分析方法确定宏、微观断口参数与宏观力学参数之间的关系。

断口定量分析中，目前最重要和最有价值的是利用疲劳断口参数进行寿命估算，利用疲劳断口弧线、条带间距反推疲劳寿命已成功地应用于以下几个方面：①低周疲劳寿命估算；②恒载与谱载下的疲劳寿命估算；③起源于先天性“裂纹”或缺陷的高周疲劳寿命估算。

在长期的实践中，人们通过对大量试验数据的总结分析，基于不同的疲劳裂纹扩展理论先后总结出了数百个疲劳定量公式和方法。需要注意的是，所有这些定量计算公式均是在试验条件下所得数据的一种经验总结。

1）早期的疲劳裂纹扩展理论　1963年，Paris等人应用应力强度因子概念，根据实验得出了著名的Paris公式。该公式因其形式简单，得到普遍应用，特别是在高循环低应力的高周疲劳情况下，即$10^{-6}\sim10^{-3}$ mm/循环下，非常适用，它也是目前高周疲劳断口寿命反推估算的基础。

严格地说，该公式只对脆性材料有效，原则上不能应用于韧性材料；但在实际上，只要塑性很小，K的计算是可以接受的，并为大多数实验结果所证实。

Paris公式的基本形式是：

$$da/dN = c\ (\Delta k)^m \qquad (5.2\text{-}3)$$

式中，da/dN为裂纹扩展速率；Δk为应力强度因子范围（$\Delta k=\Delta k_{max}-\Delta k_{min}$）；$c$、$m$为常数。

后人在自己的实践中，在他的基础上进行了发展，对此公式进行了修正，但增加了许多参数，使参数确定需要大量的试验。

目前，研究中用得最多的仍是Paris公式。

2）中期裂纹扩展理论　20世纪60年代以来，人们利用不同分析方法及物理模型得到了一些裂纹扩展的经验公式。

Bilby与Heald从位错连续分布理论推出了公式：

$$\frac{da}{dN}=\frac{5}{3G\beta\sigma_b^2}\left(\frac{\pi}{4}\right)^3K \qquad (5.2\text{-}4)$$

式中，σ_b为材料抗拉强度；G为切变模量；β为断裂塑性功；K为应力强度因子。

Forman、Elber等从连续弹塑性断裂力学推出了Forman公式：

$$da/dN = c\ (\Delta k)^m/\left[\ (1-R)\ Kc-\Delta K\right] \qquad (5.2\text{-}5)$$

式中，R为应力比；Kc为临界应力强度因子。

Elber公式：

$$da/dN = c\ (\Delta k_{eff})^m \qquad (5.2\text{-}6)$$

式中，Δk_{eff}为有效应力强度因子幅。

3）近期的裂纹扩展理论　近年来，又有许多半经验裂纹扩展方程和解析方程被推导出来，如Irving公式，Duggan方程。他们均有更加复杂的数学形式、涉及到更多的材料性能参数。即使在良好的实验条件下，确定这些参数也非易事，那么应用其对疲劳断口进行定量计算就更难了。

由于实际使用中，载荷几乎都是随机变幅的，大小载荷循环交替出现，载荷交互作用，使裂纹的扩展行为大大复杂化了。对于载荷的交互作用，其首要影响是“超载迟滞”，即在等幅载荷中间，加上一个“超载”，就使得裂纹扩展停滞一段时间，然后才恢复到正常等幅载荷下的扩展速率。其次是“反向超载加速”，即指在某个疲劳循环的谷值载荷比其后循环谷值载荷更低，当这种“反向超载”紧跟在“超载”之后，往往会大大减少、甚至完全抵消“超载迟滞”的影响。

为了很好地描述载荷间的这种相互作用，近年来研究人员提出了一些修正模型和公式，主要有Wheeler模型、Willenborg及Chang模型、何庆芝模型。虽然这些模型在描述载荷间的交互作用时有一定的改善，但都没有考虑“反向超载”的情况，且模型中的所谓材料参数均不是常数，随超载不同而变化，因此限制了其使用。陈志伟等为了能考虑“反向超载”影响，提出了Chen-Lee模型来修正裂纹尖端的有效应力强度因子，得到了较好的实验室结果。其修正系数公式为：

$$U=\ (K_{max}/K_{ap})^m \qquad (5.2\text{-}7)$$

式中，K_{max}为当前循环的峰点应力强度因子；K_{ap}为边界应力强度因子；m为可调整的材料常数。

有效应力强度因子为：$\Delta K_f=U\Delta K_{ap}$

综观以上考虑载荷交互作用的各种模型，所要确定的各种参数均较多，特别是对实际随机使用的零部件来说，超载的程度、频率等均是难以确定的参数，因此这些模型和公式用于研究实际使用的各种失效件仍有相当的难度。

由于疲劳断口定量反推计算所掌握的信息，特别是对使用中的失效件，反推计算时掌握的往往仅仅是失效件本身，要确定过多的参数显然是不可能的。因此，目前疲劳断口定量反推计算还是多采用早期的裂纹扩展公式，因为他们具有简单的形式，直观可测的参量，关键是对于不同的数据，如何选择较好的拟合方式，得到可接受的估算结果。

使用过程中得到的实际失效件断口是疲劳断口的特殊类别，首先，所能得到的实际失效件疲劳裂纹扩展过程参量很少，长期的失效分析实践表明，事故发生后，失效分析人员得到的往往仅仅是失效件断口本身；其次是由于使用过程有别于试验过程，因而它具有不同于一般疲劳试验断口的显著特点：作用于机件的应力复杂，超载迟滞、低载加速效应在断口上难以识别；使用环境千差万别，时常有腐蚀介质的影响等。因此探讨疲劳失效件断口定量分析模型有一定的难度，相应的定量分析方法也研究得很少，但从需求的角度来说，疲劳失效件断口定量分析研究是一富有活力的领域。

断口是构件在服役、使用、试验等过程的真实、详细记录。疲劳断口所隐含的定量信息最丰富、详细，真实地记录了疲劳失效的全过程，因此，寻找服役过程与断口信息之间的关系是断口定量分析的关键之一。只有确定了两者之间的对应关系，才能定量地分析出寿命来。因此，需要在对断口特征、构件的工作特点、受力特点分析的基础上，确定出工作过程与断口特征的对应关系。一方面，从断口上的特征点再到服役过程中找出对应的使用时间点；另一方面，对失效件的受力特征进行分析，结合断口上的点、线、面的特点，确定两者之间的对应关系。

确定裂纹扩展速率随裂纹长度的变化规律是对疲劳断口进行定量分析的另一关键之所在。在确定了服役过程与断口特征量的对应关系后，从断口上测量相关特征量数据，绘制相应的裂纹长度a与裂纹扩展速率da/dN图；观察两者之间的关系趋势，运用前人的研究成果，确定合适的分析模型和计算方法，对疲劳特征量进行计算；将计算结果与特征点数据对比分析，并对确定的对应关系进行修正，重新测量、

计算。最终给出疲劳裂纹的扩展速率、扩展寿命等。

(4) 疲劳断口特征与载荷历程对应关系的确定

确定断口特征与载荷历程的对应关系，即确定断口上什么特征值对应什么样的载荷、寿命值，这是疲劳断口定量反推计算的关键之一。

疲劳断裂是一个过程，构件在疲劳交变载荷作用下，将萌生疲劳裂纹，并扩展。一般交变载荷循环一次，裂纹将向前扩展相应的量，并停顿；如此往复，裂纹将反复扩展、停顿，在断口上留下相应的一系列痕迹。最常见的疲劳弧线、条带就是这种特征痕迹。它们的一些量的东西，如数量、间距、方向等就反应了应力水平、循环数量、寿命、裂纹的扩展速率等。

为了确定两者之间的关系，下面对各自的特点进行分析，以掌握它们各自的规律，从中确定可能的联系。

1) 疲劳载荷的基本概念　造成疲劳破坏的交变载荷，称为疲劳载荷。一般分为确定的和随机的两种。在确定的疲劳载荷中，幅值不变的叫等幅疲劳载荷；幅值按一定规律变化的叫变幅疲劳载荷；幅值随机任意变化的叫随机疲劳载荷。实际结构遇到的通常是随机载荷。

2) 疲劳断口的特征及规律　疲劳断口上一般可分为三个区：疲劳源区、疲劳扩展区和瞬时断裂区。由于疲劳裂纹扩展速率、扩展寿命主要与疲劳扩展区有关，下面重点描述疲劳扩展区的特征。

疲劳扩展区的最基本特征是疲劳弧线或疲劳条带。金属疲劳断口宏观形貌的基本特征是疲劳弧线，它是一种以疲劳源为中心，与疲劳裂纹扩展方向相垂直的呈半圆形或扇面形的弧形线。一般认为，疲劳弧线是由于外加应力大小、方向发生较大的改变，使得疲劳裂纹扩展的速度和方向发生变化，在破断面上留下的塑性变形的痕迹。显然，疲劳弧线的形状、间距等与应力有着必然的联系。疲劳弧线的形状和间距变化与受力状态、材质及环境介质等有关。

放射棱线指向裂纹扩展方向。

金属疲劳断口微观形貌的基本特征是疲劳条带。当疲劳实验是在均匀变化应力条件下进行时，疲劳条带的间距在局部范围内是相当均匀的。对于一定循环应力下的疲劳断口，每一条疲劳条带即是由一个应力循环周期产生的。但在实际中，疲劳条带的形成还受许多因素影响，依据它进行定量分析还是十分复杂的技术难题。

当疲劳裂纹扩展至临界尺寸后发生失稳快速破断，称为瞬时断裂，瞬时断裂的区域就称为瞬断区。瞬断区面积的大小取决于载荷的大小、材料的性质、环境介质等因素。通常情况下，瞬断区面积越大，则表示所受载荷越大或者材料越脆。

3) 工作载荷与断口特征的对应关系

① 颜色、成分与工作历程、时间的对应关系　裂纹形成以后，断口将暴露于空气中，必然要受到大气介质或工作环境介质如燃气、油液等的氧化、腐蚀，从而在断口颜色上表现出差异；断口暴露的时间越长，颜色上的差异越大。在发动机高温部件上如二级涡轮叶片断口上，常见的不同颜色的彩色弧形带（见图 5.2-29）就是在裂纹扩展过程中，裂纹不同长度位置在高温燃气中暴露的时间不同的一种反映。

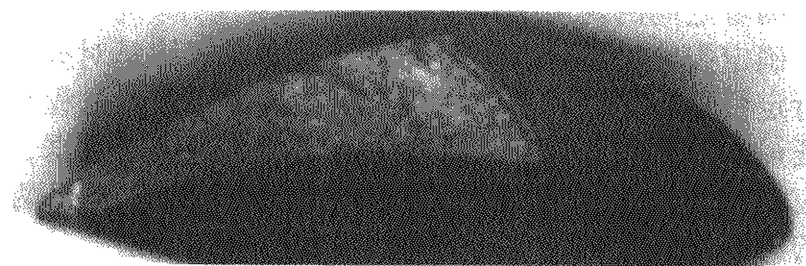


图 5.2-29　涡喷发动机涡轮叶片疲劳断口上的彩色弧线

如果在裂纹扩展过程中经历过化学处理过程，则这种差异表现得更加明显，在化学处理过程的前后，断口上颜色出现突变。如构件在出现裂纹后进厂进行大修，修理时进行了阳极化处理或电镀处理，陈旧性裂纹的部位，阳极化溶液或电镀液通过裂纹张口进入，腐蚀了断口，从而形成具有阳极化颜色的区域（见图 5.2-30）。



图 5.2-30　飞机起落架旋转臂横梁断口上陈旧性裂纹特征

同样，裂纹不同长度部位暴露的时间不同，其表面的化学成分也会表现出明显的差异。由表面成分差异的大小也可初步判断时间。

必须注意的是，从断口颜色、成分的差异来判断裂纹萌生的时间应该与使用历程、使用条件联系起来，只有在了解了使用历程、使用工况后才会取得有意义的结果。

② 宏观上的对应关系　宏观上，疲劳断口一般分为三个区：疲劳源区、扩展区和瞬断区；当疲劳裂纹扩展到临界尺寸时，裂纹快速扩展，构件发生瞬时断裂。因此，扩展区与瞬断区的交界线即与构件瞬时断裂时的寿命点相对应。

疲劳弧线一般是疲劳断口常见的断口特征，疲劳理论认为，疲劳弧线是裂纹扩展过程中，载荷或应变发生大的改变时，在断口上留下的塑性变形痕迹，因此每一条疲劳弧线（宏观）与裂纹扩展过程中载荷或应变发生的一次大的改变（如载荷谱的加载、环境条件等的改变）相对应。在飞机、发动机构件上，飞机每一次地－空－地循环、发动机每一次起动停车均要发生载荷的大的改变，因此，从飞机、发动机的受载特点及构件疲劳断口特点的对比分析可以看出，飞机、发动机构件疲劳断口上疲劳弧线是分别与飞机的地－空－地循环（起落循环）、发动机的起动循环相对应的。

对载荷谱试验的情况下，分析对比载荷谱特点及断口特征，也可确定相应的对应关系。

测量疲劳弧线的数量、间距就可以确定载荷的变化次数、大小。载荷的变化值越大，弧线间距越大。对载荷谱或使用历程中载荷进行分析，分别确定出不同载荷级别的时间、大小、变化幅度，与断口上疲劳弧线的粗细、间距的大小等进行分析对照，对两者之间进行宏观的一一对照，即可确定两者之间的对应关系。

③ 微观上的对应关系　疲劳断口的微观特征主要是疲劳条带，因此，从微观上反推疲劳裂纹扩展寿命可利用的主要数据就是疲劳条带间距。利用断口疲劳条带间距（微观）进行定量分析的理论依据是：每一疲劳条带相当于载荷或应变的一次循环。

(5) 疲劳弧线或疲劳条带间距的测定方法

利用疲劳条带间距进行断口定量反推的前提是能够对疲劳断口上的疲劳条带间距进行测定。测量疲劳弧线或疲劳条带的间距一般采用实体光学显微镜、扫描电子显微镜或透射电子显微镜，沿着主裂纹扩展的方向，在一定放大倍数下测量出距源区不同距离处的疲劳弧线或疲劳条带的间距。由于疲劳条带的扩展速率主要与局部微小区域的应力和材质相关，因此，疲劳条带的间距不可避免地存在一定的分散性，测量时应取一定范围内的平均值作为某一点的间距。在用扫描电子显微镜进行疲劳条带间距测量时，还应考虑疲劳条带法线与裂纹扩展方向不同所带来的位向误差和断口表面粗糙

带来的投影误差，需要分别进行相应的角度修正和投影修正，以使测得的数据为疲劳条带的真实数据。透射电子显微镜测出的条带间距一般与真实疲劳条带的平均间距较为接近，一般不需修正，但缺点是必须采用二次复型，不能直接观察测量，程序复杂，工作量大。

(6) 断口反推疲劳裂纹扩展速率的基本方法

若令每一载荷循环下的疲劳裂纹扩展量为 u，则

$$u = da/dN$$

或

$$dN = da/u \quad (5.2\text{-}8)$$

式中，a 为裂纹长度；N 为循环次数。

由上面的分析可知，疲劳弧线间距或疲劳条带间距与裂纹的扩展量是相当的。因此，只要在断口上若干离源区不同距离长度的裂纹 a_i 处测量疲劳弧线、条带间距，或沿一定长度的直线上测量疲劳条带数目，便可作出每一载荷循环下的裂纹扩展量（或单位裂纹长度上的疲劳条带数）与裂纹长度相互关系的试验曲线，进而可求得疲劳裂纹扩展寿命 Np：

$$Np = \int_{a_0}^{a_f} da/u \quad (5.2\text{-}9)$$

式中，a_0 为微裂纹开始扩展时的尺寸；a_f 为发生瞬断时的裂纹尺寸。

对于载荷谱加载，上式依然适用，只是此时 N 为载荷谱的数目，da 为疲劳弧线间距。

当构件存在裂纹，在疲劳载荷作用下，裂纹不发生扩展的条件是裂纹尖端的应力强度因子变程 ΔK 小于对应应力比下的裂纹扩展门槛值 ΔK_{th}，即：

$$\Delta K < \Delta K_{th}$$

式中 $\Delta K = \Delta K_{max} - \Delta K_{min}$。

当构件承受一稳定的交变载荷时，裂纹扩展速率可用下式表达

$$da/dN = c(\Delta K)^m$$

$$da/dN = 0(\Delta K < \Delta K_{th})$$

在一些特殊条件下，构件中虽然存在裂纹，但只有在某种特定条件下，才存在 $\Delta K > \Delta K_{th}$。由于在一个振动周期内疲劳裂纹扩展速率是变化的，因此不能用断口定量测定疲劳条带间距来求得疲劳裂纹扩展寿命。由于在过度转速下，发生瞬间振动导致裂纹扩展，每一次振动扩展均会在断口上留下疲劳弧线，即表明发生振动的次数。此时疲劳裂纹扩展寿命 Np，即裂纹从 a_0 扩展到 a_c 所经历的载荷发生大的变化的次数，可通过下式计算：

$$Np = \int_{a_0}^{a_c} \frac{dH}{da \times P\{\Delta K > \Delta K_{th}\}} da \quad (5.2\text{-}10)$$

式中，dH 为疲劳弧线间距；P 为发动机转动部件发生共振的概率。

利用断口反推计算疲劳裂纹扩展寿命的关键是得出疲劳裂纹扩展速率 da/dN 的数学表达式。有关描述疲劳裂纹扩展速率的数学表达式很多，但较为实用且应用最为广泛的仍是人们所熟知的 Paris 公式。另一个较为广泛应用的表达式是根据图 5.2-31 建立的经验公式：

$$1/u = A + Be^{-a/C} \quad (5.2\text{-}11)$$

式中，A、B、C 为常数，其值可从图 5.2-31 求得；$1/u$ 为单位长度的疲劳条带数目。

由 $dN = da/u$ 积分，得：

$$N = \int dN = \int da/u = C[Aa/C + B(1 - e^{-a/C})] \quad (5.2\text{-}12)$$

从 Paris 公式也可求出疲劳裂纹扩展速率与裂纹长度之间的关系

$$da/dN = c\ (\Delta K)^m \quad (5.2\text{-}13)$$

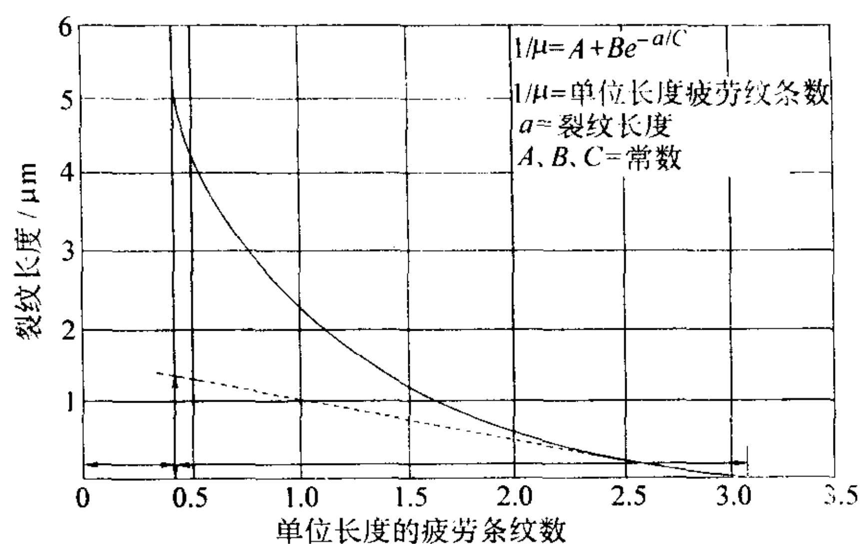


图 5.2-31 经验关系式

$$\Delta k = Y\Delta\sigma\ (\pi a)^{1/2} \quad (5.2\text{-}14)$$

式中，c、m 为材料常数；Y 为与裂纹有关的构件形状因子；$\Delta\sigma$ 为最大应力与最小应力之差；a 为裂纹长度。

对给定构件及恒定载荷，则有：

$$da/dN = c(Y\Delta\sigma\sqrt{\pi})^m \times a^{m/2} = c_0 a^{m/2} \quad (5.2\text{-}15)$$

式中，$c_0 = c\ (Y\Delta\sigma\sqrt{\pi})^m$ 为常数。将上式积分，可得：

$$Np = \int_{a_0}^{a_f} \frac{da}{c_0 a^{m/2}} = \frac{2}{(2-m)c_0} a^{1-m/2} \Big|_{a_0}^{a_f}$$

$$= \frac{2}{(2-m)c_0}[a_f^{1-m/2} - a_0^{1-m/2}] \quad (5.2\text{-}16)$$

常数 C_0 和 m 可由如下方法确定，即对式（5.2-15）取对数

$$\lg(da/dN) = \lg c_0 + (m/2)\lg a \quad (5.2\text{-}17)$$

则 lg（da/dN）与 $\lg a$ 为直线，截距为 $\lg c_0$，斜率为 $m/2$。

因此，对于不同裂纹长度 a_i 所对应的（da/dN）$_i$，则可按式（5.2-17）进行拟合或分段拟合，求出 c_0 及 $m/2$，按式（5.2-16）求得 Np。

对于随机载荷目前尚未有描述裂纹扩展速率的数学表达式，应该根据实际情况分析随机载荷的一般规律，进而建立相应的裂纹扩展速率的数学模型。

在具体的断口反推疲劳寿命中，应按如下步骤进行。

1）断口分析　断口分析是疲劳断口定量反推的基础，其基本方法与常规断口相同，主要是要判断断口的性质、疲劳源点的位置、疲劳裂纹的走向、疲劳扩展区与瞬断区的交界位置、有无陈旧性裂纹等，同时还要判断疲劳特征、有无明显的过载迟滞特征。

2）载荷、历程分析　载荷、历程分析的目的是确定载荷的范围、变化规律及特征点，为寻找断口特征与载荷、寿命之间的关系打下基础。一般需要注意大载荷出现的时间、幅值；载荷的交变频率，如低周还是高周、高低周复合；失效件使用过程中的特殊工况条件、时间，大修等处理记录；较全面地掌握失效件的履历情况。

3）确定载荷与断口特征的关系　断口上的特征是构件经历的使用条件和载荷变化过程的真实、全面、准确的反映，因此，断口特征与载荷、使用历程之间应有一定的对应关系。

在确定过程中，应特别注意观察断口颜色，颜色的明显变化点，一般与大修或使用中的特殊处理时间点相对应；大的弧线或条带间距一般与大的过载之间有对应关系；裂纹走向的突然改变一般与载荷方向的改变有关。在确定了断口上的特殊寿命点后，就可以此来验证和修正反推估算的结果，提高准确性，减少计算的误差。

4）数据获取　在确定了断口特征与使用载荷、历程的对应关系后，利用实体光学显微镜、扫描电子显微镜或透射电子显微镜等仪器，观察、测量断口上的相关数据，如疲劳

区面积、瞬断区面积，沿着裂纹扩展方向的疲劳弧线、条带的数据等。

5）对数据的拟合、估算　将从断口上获得的数据在坐标纸上作图，观察其趋势，并采用合适的模型进行拟合处理，确定裂纹扩展速率随裂纹长度的变化规律。

目前许多研究者已从试验中总结出了一些经验公式，但能够合理地描述使用中随机载荷下的裂纹扩展速率的数学表达式还没有，因此，对失效件的裂纹扩展速率的描述还应根据实际情况分析随机载荷的一般规律，进而建立相应的数学模型。

综上所述，利用疲劳断口形貌反推失效件裂纹扩展寿命，首先应通过对失效件的断口分析和履历、载荷分析，找出断口形貌与载荷谱之间的对应关系，并从断口上实测出对应于每次载荷循环数的裂纹扩展长度 da/dN；然后根据测得的一组数据绘制裂纹长度与谱循环数（或与循环次数）的关系曲线，即疲劳裂纹扩展曲线；最后，对实测曲线拟合，将其反推至整个裂纹扩展区，并以此求得整个断口的裂纹扩展速率、扩展寿命。

在反推过程中应以分析确定的特征寿命点来验证计算结果，以提高反推的准确性。

(7) 断口反推疲劳原始质量

按照损伤容限理论，结构件在服役前均带有初始裂纹或缺陷，由该裂纹或缺陷扩展达到临界裂纹的寿命即为结构的总寿命。损伤容限是指结构材料在正常工作载荷下，仍然保证工作安全可靠所容许的最大损伤程度或大小。在实际零件中，必然存在有某些缺陷，如夹杂、气孔、位错等材质缺陷，加工刀痕、划伤、毛刺等表面加工缺陷，以及残余应力等表面完整性因素。按照损伤容限设计思想，这些存在于零件中的初始缺陷对疲劳寿命的影响可等效为一当量裂纹长度，称为当量初始裂纹 a_0，可作为表征零件质量的参量（当量初始质量），对疲劳断裂失效，则称为原始疲劳质量 a_{0i}。

根据上述观点，零件在开始就存在当量初始裂纹 a_{0i}，经过 N 次循环后，裂纹扩展至 a_N，因此有：

$$a_N = a_{0i} + \sum \Delta a_i$$

则

$$a_{0i} = a_N - \sum \Delta a_i \tag{5.2-18}$$

式中，a_{0i}为当量初始裂纹尺寸；Δa_i 为每次载荷循环的裂纹扩展量；a_N 为载荷 N 次循环后裂纹逐次扩展的总量。

显然，只要求出裂纹的总扩展量即可确定当量初始裂纹长度。

利用断口形貌反推零件当量初始质量的基本原理、方法和步骤如下。

1）通过对模拟件实验断口或失效件的断口分析，找出断口形貌与载荷谱之间的对应关系，并从断口上实测出对应于每次载荷循环数的裂纹长度。

2）根据测得的数据绘制裂纹长度与谱循环数（或循环次数）的关系曲线—疲劳裂纹扩展曲线。

3）拟合实测曲线，并将与其相吻合的裂纹扩展规律反推至时间为零，即 $N=0$，此时对应的裂纹长度即为该零件的 a_{0i}。

4）对同类零件的 a_{0i}进行统计分析，根据 a_{0i}的分布密度确定规定概率下同类零件的 a_{0i}。

需要说明的是，a_{0i}仅是损伤容限设计中使用的一个参考值，是零件在服役前的内部各种缺陷、表面加工缺陷及表面非完整性因素等对疲劳寿命影响的一个综合反映参考量，虽然在理论上，该值仅是材料、加工工艺和装配质量的函数，与载荷因素无关，但在实际中，许多因素对其均有较大的影响，如载荷谱顺序、载荷大小、断口金相尺寸范围、反推使用的裂纹扩展曲线形式、拟合实测曲线的精度等。

(8) 疲劳断口反推失效构件的应力

1）利用疲劳裂纹长度及瞬断区大小反推构件承受的疲劳应力水平　由于构件疲劳裂纹临界长度 a_c 与最大交变应力 σ_{max}存在如下关系：

$$K_{\mathrm{I}} = \sigma_{max} Y \sqrt{\pi a} \tag{5.2-19}$$

式中，K_{I} 为裂纹尖端的应力强度因子；a 为裂纹长度；Y 为与裂纹有关的形状因子；σ 为构件承受的应力。

当发生快速断裂时，$K_{\mathrm{I}} = K_{\mathrm{Ic}}$，$a = a_c$，$\sigma = \sigma_c$，则有

$$K_{\mathrm{Ic}} = \sigma_c Y \sqrt{\pi a} \tag{5.2-20}$$

式中，K_{Ic}为材料抵抗裂纹失稳扩展的能力，是材料常数。

因此，在材料一定的情况下，就有

$$\lg a_c \propto -\lg a_c \tag{5.2-21}$$

通过试验测得一系列的 $a_c-\sigma_c$ 数据，拟合求得关系式，将失效件的值代入即可估算出失效件所承受的疲劳应力水平。钟群鹏等人对 30CrMnSiA 旋转疲劳试样不同应力集中系数 K_t、不同应力水平下的 207 个试样断口的特征参数进行了测量与统计分析，得出了 $a_c-\sigma_c$ 的关系式，对失效件估算的应力水平，其相对误差仅为 4.6%。说明利用 a_c 和 σ_c 的关系来确定 σ_{max}是可行的。

疲劳瞬断区是疲劳裂纹达到临界尺寸后发生的快速破断，其面积大小一般认为受材料的断裂韧性 K_{Ic}控制。因为 a_c 与瞬断区面积是相关的，因此通过试验测得一系列的 $A-\sigma_c$数据，拟合求得关系式，将失效件的值代入也可估算出失效件所承受的疲劳应力水平。

必须注意的是，以上方法只能求得疲劳载荷的最大应力水平，但不能推算疲劳载荷的范围。

2）利用疲劳条带间距反推失效件的疲劳应力　疲劳裂纹扩展的第二阶段的速率可用 Paris 公式来表达：

$$da/dN = c(Y\Delta\sigma\sqrt{\pi})^m a^{m/2}$$

断口反推法可求得失效件疲劳裂纹扩展速率 $(da/dN)_{sx} \sim a$，若用试样模拟同种材料的疲劳裂纹扩展速率 $(da/dN)_{sy} \sim \Delta K$，即可求得该材料在 Paris 公式中的材料常数 c、m，在裂纹稳定扩展阶段，则有：

$$\frac{(da/dN)_{sx}}{(da/dN)_{sy}} = (\Delta\sigma_{sx}/\Delta\sigma_{sy})^m [Y_{sx}/Y_{sy}] \tag{5.2-22}$$

若已知 Y_{sx}的表达式或数值解，则可求得 $\Delta\sigma_{sx}$。

对一些形状及受力均复杂的大型构件，Y_{sx}难以用表达式来表达，则可用与失效件等同的模拟试验件进行疲劳试验，求得 $(da/dN)_{sy} \sim a$，则有：

$$\frac{(da/dN)_{sx}}{(da/dN)_{sy}} = \left(\frac{\Delta\sigma_{sx}}{\Delta\sigma_{sy}}\right)^m$$

$$\Delta\sigma_{sx} = \frac{(da/dN)_{sx}^{1/m}}{(da/dN)_{sy}^{1/m}} \Delta\sigma_{sy} \tag{5.2-23}$$

若知道实际构件所承受的应力比 $R=\sigma_{min}/\sigma_{max}$，则可求得

$$\sigma_{max} = \Delta\sigma_{max}/(1-R) \tag{5.2-24}$$

需要注意的是，就目前断口反推计算常用的检测仪器而言，在速率小于 10^{-6} mm/循环内分辨疲劳条带间距是异常困难的，因而工程上利用断口反推疲劳裂纹扩展寿命通常是指疲劳裂纹扩展速率在 10^{-6} mm/循环 $< da/dN < 10^{-4}$ mm/循环的有限范围内。

近期才发展并受到极大关注的扫描隧道显微镜分析技术为疲劳裂纹萌生及初期扩展寿命的定量计算提供了可能。

2　裂纹诊断技术和方法

裂纹是材料表面或内部的完整性或连续性被破坏的一种

现象，是断裂的前期；断裂则是裂纹发展的结果。裂纹诊断包括裂纹的无损检测、表面分析、光学金相分析及裂纹打开后的断口分析等内容。

2.1 裂纹的无损检测

裂纹的无损检测是生产和使用过程中质量检验、安全控制的一项重要内容，除肉眼的外观检查外，通常采用无损检测方法进行检查。常用的无损检测方法有：X 射线、磁力、超声波、荧光、着色、声发射、敲击测音法、工业 CT 等物理检测方法，其中磁力、荧光、着色等方法主要用来检查表面裂纹，而 X 射线、超声波、声发射、工业 CT 等可检测表面和内部裂纹。声发射靠捕捉裂纹扩展中发射的声信号来检测裂纹，因此只能用来检测正在扩展中的裂纹，对静态下的裂纹无能为力。在裂纹检测中，需要注意检测的方向，超声波、磁力必须垂直裂纹平面检测，X 射线检测须平行于裂纹所在的平面才能检测出来。有关裂纹检测的具体方法可参阅有关的专业书籍，在此不展开论述。

2.2 裂纹产生先后顺序诊断技术和方法

在断裂失效分析中，往往存在有多条裂纹，如果对所有裂纹均进行分析，往往是不现实的，也是不必要的。一般来说，在众多的裂纹中，它们的产生在时间上是有先后顺序的，而最先产生的裂纹往往是导致其他裂纹产生和整个事故（故障）发生的根本原因。因此，在分析中，首要任务就是从这些裂纹中确定首先产生的裂纹，即主裂纹，然后对主裂纹进行分析。

(1) 确定主裂纹（首断件）的原则

从众多的断裂件中确定主裂纹、找出首先断裂失效件所要遵循的原则是：根据各零件的功能特征，各相关零件的损伤痕迹，各零件的断裂形貌特征等加以综合分析判断。

1) 当各断裂件中，既有延性断裂，又有脆性断裂时，一般脆性断裂件发生在前，延性断裂件发生在后；

2) 当各断裂件中，既存在脆性断裂件，又存在疲劳断裂件时，则疲劳断裂件应为首断件；

3) 当存在两个或两个以上的疲劳断裂件时，低应力疲劳断裂件出现在前，而大应力疲劳断裂件出现在后；

4) 当各断裂件均为延性断裂时，则应根据各零件的受力状态、结构特性、断裂的走向、材质与性能等进行综合分析与评定，才能找出首先断裂失效件。

(2) 确定主断口或主裂纹的方法

有时会碰到在同一断裂失效件上出现多个断口或多条裂纹，这就要求从中准确地找出首先开裂的部位，即主断口或主裂纹。

常用的裂纹先后顺序判断方法主要有以下五种。

1) 塑性变形量大小确定法　当零件在断裂过程中断裂成多块，有的部位没有明显的塑性变形，而有的部位塑性变形明显，则无塑性变形的区域为首先断裂区域；当所有断裂部位均为延性断裂时，可根据各断裂部位的变形量大小来确定主裂纹，其中变形量大的部位为主裂纹，其他部位为二次或三次裂纹。

2) T 形法　当一个零件上同时出现两条或多条裂纹，这些裂纹构成“T”型关系时，可根据裂纹的相对位置关系来确定主裂纹。如图 5.2-32 中所示，横贯裂纹 *A* 为主裂纹，因为 *A* 裂纹阻止了 *B* 裂纹穿过 *A* 裂纹向前扩展。因此，*A* 裂纹形成在前，而 *B* 裂纹形成在后，为次裂纹。

3) 裂纹分叉法　机械零件在断裂的过程中，在出现一条裂纹后，往往会引生出多条分支裂纹或分叉裂纹，如图 5.2-33 所示。由于分支裂纹或分叉裂纹是从主裂纹引生的，因此，裂纹的扩展方向为从主裂纹向分叉或分支裂纹方向，分叉或分支裂纹汇集的裂纹为主裂纹。

4) 断面氧化颜色法　金属零件暴露在环境介质、高温下时将会被腐蚀和氧化，而且腐蚀和氧化的程度会随时间的增加而加重。由于主裂纹较次裂纹形成的时间早，因此主断面较次断面暴露在环境介质和温度下的时间就长，腐蚀与氧化的程度就要严重，腐蚀产物多、氧化颜色深，据此可以判断主裂纹的部位。

5) 疲劳裂纹长度法　当同一零件上出现多条疲劳裂纹时，一般可根据疲劳扩展区的长度、疲劳弧线和疲劳条带间距的大小来判断主裂纹。疲劳裂纹长、疲劳弧线或疲劳条带间距小的为主裂纹。

需要特别说明的是，以上的五种方法均是针对一般情况而言的，而实际失效事件中情况复杂多变，因此，在实际失效分析中，应该根据具体情况和相关条件，结合结构特点、受力状态、工作原理、裂纹扩展规律、表面痕迹特点、零件材料与性能等综合分析，才能给出准确的判断。

一般地讲，脆性断裂可用 T 形法或分叉法来判别主裂纹与二次裂纹；延性断裂则可用变形法来判别主次裂纹；环境断裂可根据断面氧化与腐蚀程度及颜色深浅来区分主次断裂；而疲劳断裂常常利用断口的宏观与微观特征形貌加以区分。

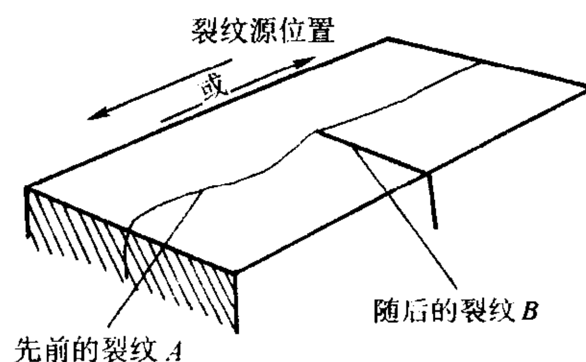


图 5.2-32　判别主裂纹的 T 形法示意图
A—主裂纹；*B*—二次裂纹

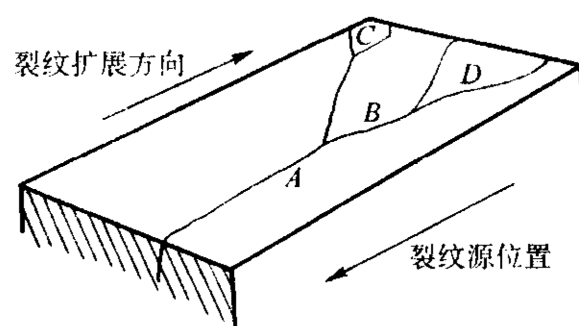


图 5.2-33　判别主裂纹的分叉法示意图
A—主裂纹；*B*、*C*、*D*—二次裂纹

2.3 裂纹的形貌诊断

2.3.1 裂纹的宏观诊断

(1) 裂纹宏观诊断的目的及内容

裂纹宏观诊断的目的是确定裂纹的位置、类型、外观形貌及张开情况等。裂纹宏观诊断的主要内容有：

① 裂纹产生的部位；

② 裂纹的平直情况、或分叉情况、或宏观走向的变化等；

③ 裂纹与主应力方向（或切应力方向）之间的关系；

④ 裂纹与材料成形方向（轧制方向、流线方向）的关系；

⑤ 裂纹的啮合情况（紧密配合，还是分离）；

⑥ 裂纹尖端的情况（尖锐或圆钝）；

⑦ 裂纹起始位置与零件形状的关系（是否有应力集中）。

确定裂纹产生的部位是裂纹宏观诊断首先必须解决的问题。裂纹一般容易产生于尖角、转折、几何尺寸突然变化处

等应力集中部位，受力最大部位，焊缝熔合区等组织薄弱部位，材料缺陷处。其次是确定裂纹的外观形貌、张开情况和匹配情况，并根据这些情况诊断出裂纹的类型、起始源区。

确定裂纹产生的部位，以及是否是龟裂纹、线裂纹、多条裂纹、起始源区、匹配情况，是否是引起应力集中的部位，再结合应力状态，初步判断裂纹产生的条件，如在应力集中部位，则可能与使用载荷有关；不在应力集中部位，则可能与材料性能、成分、缺陷和内应力等有关。

裂纹的起源位置和扩展途径决定了裂纹的宏观形貌，是构件局部受力状态、大小（外力）与材料强度（抗力）综合效应的结果，即应力－强度干涉作用的结果。它们往往是应力较大（动力大）、强度值较低（阻力小）的路径，如应力集中处、材料局部缺陷处等。

(2) 裂纹的宏观形貌

裂纹的宏观形貌很多，常见的有龟裂纹、线裂纹、环形裂纹、周向裂纹、辐射状裂纹、弧形裂纹等，下面分别对其进行讨论。

1) 龟裂纹　外观形貌类似于龟壳网络状分布的一类裂纹称为龟裂纹。一般情况下，龟裂纹是一种表面沿晶裂纹，其深度一般都不大。龟裂纹产生的原因是由于零件表面（或晶界）的成分、组织、性能及应力状态与中心（或晶内）不一致，在制造工艺过程或随后的使用过程中使晶界成为薄弱环节；同时，在制造工艺过程或随后的使用过程中还会产生很大的组织应力、热应力等内应力，在这些内应力的作用下，薄弱的晶界发生开裂，从而形成龟裂纹。根据龟裂纹的形成条件，可将其分为如下几类。

① 铸造表面龟裂纹　这是一种不常见的龟裂纹，是浇铸时融熔金属液与模壳材料反应，生成的硅酸盐夹杂物在铸件表面的初始奥氏体上析出的结果。也可能是在1 250～1 450℃在铸件表面形成的热裂纹。

② 锻造表面龟裂纹　在锻件加热过程中，由于加热温度过高或停留时间过长，使锻件晶粒严重粗化，脆性增加，严重时出现晶界氧化，以至在后续的锻造加工时沿晶界出现表面龟纹状开裂的现象。这种裂纹的表面一般有氧化色，无金属光泽；断口粗糙，颜色灰暗。

当钢中含硫量较高时，在晶界会形成低熔点的硫化铁和铁的共晶，高温锻造时，这些共晶物处于熔融状态，无法随其他金属一起塑性变形，从而引起锻造开裂。这种开裂的表面也呈龟裂状。

当钢中含铜量过高时（质量分数大于0.2%），在锻造加热过程中，表层的金属会发生选择性氧化，即铁首先氧化，从而使得铜在晶界聚集，沿晶界形成熔点低于基体的富铜网络，锻造加工时形成铜脆龟裂纹。

③ 热处理表面龟裂纹　表面容易脱碳的高碳钢零件，当淬火温度过高、时间过长时，表面会产生脱碳层，使表面淬火强度大大降低；由于含碳量不同的奥氏体具有不同的马氏体转变温度和体积变形，以及淬火时表面先冷却，而内部后冷却，使得组织的转变不同时，从而使淬火组织应力显著增加，在表面形成很大的多向拉应力而产生龟裂状裂纹。

这种淬火裂纹在大型零件和重复淬火的高碳钢零件上经常可以看到。

④ 焊接龟裂纹　焊接龟裂纹属于焊接热裂纹，是由于电弧焊时，起弧电流过大，引起局部热量过高，在热应力的作用下而产生的。焊接龟裂纹一般产生于1 100～1 300℃的温度范围内、焊缝刚刚凝固、晶界强度较小的情况下。焊接龟裂纹常常产生在焊缝区或由焊缝区开始向基体金属引伸，最后成为一种沿晶粒边界分布的网状裂纹。

⑤ 磨削龟裂纹　为了满足零件的表面要求，常常需要对零件表面进行磨削加工。由于对金属表面进行磨削加工时，不可避免地会产生大量的磨削热，这些磨削热可使被磨削金属表面迅速升温至很高温度（可达840℃，6 000℃/s），产生很大的热应力；当冷却不充分时，可使表层金属重新奥氏体化，随后又再次淬火成马氏体，从而在表层产生很大的组织应力。这些表面热应力和组织应力叠加、则可能导致磨削裂纹的产生。对有的零件，在淬火回火后，其组织中还残留有一定数量的奥氏体，在磨削热影响区中，这些残留的奥氏体可能转变为马氏体，引起局部体积膨胀，产生组织应力，当这种应力大于材料的拉伸强度极限时，则可导致磨削裂纹。

从以上磨削裂纹形成的两种机理分析中可以看出，不同的原始组织对磨削裂纹的敏感性是不同的。在相同的磨削制度下，具有带状或条状分布的碳化物的钢，由于容易引起磨削应力集中，因此容易产生磨削裂纹；碳化物呈网状分布于晶界的，因促进磨削应力集中，也易形成磨削裂纹；碳化物均匀分布的钢，因不引起应力集中，故对磨削裂纹不敏感。

常见的磨削裂纹除龟裂状以外，还有与磨削方向基本垂直的、规则排列的条状形态。产生何种形态的磨削裂纹，与磨削条件、零件形状、材料质量及零件的工艺履历有关。一般因材质原因产生的磨削裂纹的形态为龟裂状，因磨削条件产生的磨削裂纹的形态为与磨削方向基本垂直的、规则排列的条状。

⑥ 使用龟裂纹　在使用过程中产生的龟裂纹主要有热疲劳（如热压或压铸模具）龟裂纹、应力腐蚀龟裂纹和蠕变裂纹。

应力腐蚀龟裂纹是零件在使用过程中，由于拉应力和腐蚀介质的共同作用而产生的一种裂纹。其微观走向取决于金属和腐蚀介质，可能是沿晶的，也可能是穿晶的。

在高温下工作的部件可能产生蠕变裂纹。蠕变裂纹是金属在“等强温度”（晶界强度和晶内强度相等的温度）以上使用时，在低应力条件下，产生的沿晶界扩展的一种裂纹。研究表明，蠕变机理是先在晶界形成孔洞、细沟或微裂纹，通过扩散使原来溶解在金属中的氧或大气中的氧进入孔洞、细沟或微裂纹，从而使晶界逐渐氧化，降低裂纹发展所需的能量，使裂纹沿晶界扩展、断裂。因此，蠕变裂纹一般从金属表面开始，起始区具有沿晶界排列的孔洞特征。

2) 线裂纹　线裂纹是指近似直线状的裂纹。最典型的线裂纹是由于发纹或其他非金属夹杂在后续工序中扩展而形成的裂纹。它们一般沿材料的纵向发展，较长，在裂纹的两侧和金属基体上一般有氧化物夹杂或其他非金属夹杂物。

发纹是钢材内部存在的非金属夹杂物沿热加工方向延伸而形成的一种纵向线性缺陷。它在塔形试样和金属制品的表面上具有近似直线状的外形。

在生产中，虽然原材料的氧化物、硫化物及发纹等均符合要求，但在淬火中仍然可能产生纵向直线状裂纹。这种裂纹多产生于一些表面冷却比较均匀一致、芯部淬透的细长工件，产生的原因是由于芯部淬透的细长工件的组织应力和热应力等淬火应力共同作用的结果。淬火时，由于内外温度不均匀，表面首先开始冷却和收缩，但受到心部的限制，所以产生热应力，表面产生拉应力，芯部产生压应力。当表面温度降到马氏体开始转变温度以下时，首先开始马氏体转变，并带来体积膨胀。同样，表面的马氏体转变也受到心部的阻碍，使表面产生压应力，心部受到拉应力，这是组织应力作用的结果。当心部受到的拉应力大到超过该温度下的屈服强度时，金属塑性变形、应力松弛。继续冷却，芯部发生马氏体转变，引起芯部体积膨胀，因受到已完成了马氏体转变的表层的阻碍，结果使表层受拉应力，芯部受压应力。当冷却到某一温度范围内，由于热应力引起的拉应力和由于组织应力引起的拉应力相叠加，并超过材料的强度极限时，淬火裂

纹就会产生。由于芯部淬透的细长工件的表层切向应力总是大于轴向应力、故淬火裂纹是纵向的直线状裂纹。

对冷拔、热拔、深冲、挤压的制品，在表面还可能产生拉痕。这种拉痕是金属在拔制或挤压变形过程中，表面金属的流动受模具内壁的机械阻碍而产生的。拉痕均沿变形方向纵向线性分布，具有一定的深度和宽度，尾端具有一定的圆角，两侧较为平整，整个宽度基本一致，且一般与表面垂直，拉痕附近的组织与基体组织无区别。

当磨削工艺不合理时，也可能产生纵向的直线状裂纹。

3) 其他形状裂纹　常见的其他形状的裂纹还有环形裂纹、周向裂纹、辐射状裂纹、弧形裂纹等。

对化学热处理的零件，由于渗层的组织与成分突然过渡，在渗层内或渗层与中心组织的过渡层内引起热应力和组织应力，从而在中心部位出现圆周裂纹。大型复杂零件，淬火时，由于一些部位冷却速度较慢而未能淬透，在淬硬层与未淬硬层产生组织应力，可能在淬硬过渡层内或附近引起弧形裂纹。

2.3.2 裂纹的微观诊断

(1) 裂纹微观诊断的内容

当宏观分析方法判断裂纹的性质困难时，或为进一步确定裂纹的性质和产生的原因，需对裂纹进行微观诊断分析，即光学金相分析和电子金相分析，通过裂纹表面形态观察和金相磨片观察来进行。

裂纹微观分析的主要内容可概括如下：

① 裂纹的微观形态特征，其扩展路径是穿晶还是沿晶；主裂纹附近有无微裂纹；

② 裂纹处及其附近的晶粒度有无显著粗化、细化或大小极不均匀的现象；晶粒是否变形；裂纹与晶粒变形的方向是否一致；

③ 裂纹两侧是否存在氧化和脱碳现象，有无氧化和脱碳组织出现；

④ 裂纹附近是否存在碳化物或非金属夹杂物，其形态、大小、数量及分布情况如何；裂纹源是否产生于碳化物或非金属夹杂物周围，其扩展方向如何；

⑤ 裂纹处是否存在粗大过热组织、魏氏组织、带状组织，以及其他的反常组织；

⑥ 裂纹源区是否存在加工缺陷、材质缺陷、腐蚀损伤等；

⑦ 产生裂纹的表面是否存在白色加工硬化层或回火层。

(2) 裂纹微观诊断技术

通过对裂纹区及其附近的金相组织、晶粒度检查，可判断出裂纹起始的部位，定性判断出裂纹部位的受力大小、加工质量等。

在微观上，裂纹源区一般均是材料的薄弱环节，如零件的表面或次表面、应力集中处和材料缺陷处（有时可见到明显的缺陷）。对于一条主裂纹，由粗到细的形态就是裂纹的扩展过程。当存在放射状微裂纹时，其“收敛”点位置即为裂纹源。

裂纹的扩展途径是复杂的，既可能是沿晶的，也可能是穿晶的、还可能是沿晶与穿晶混合的。一般情况下，制造过程中产生的铸造热裂纹、过烧引起的锻造裂纹、回火脆性裂纹、磨削裂纹、焊接裂纹，使用中出现的冷热疲劳裂纹、蠕变裂纹、热脆裂纹，环境因素引起的应力腐蚀裂纹、氢脆裂纹等均是沿晶界扩展的；而疲劳裂纹、解理裂纹、延性断裂裂纹等使用中形成的裂纹和因冷却速度过大、零件几何尺寸突变等引起的淬火裂纹、焊接裂纹等制造裂纹都是穿晶裂纹。

根据裂纹及其周围的形状和颜色，可以判断裂纹经历的温度范围和零件的工艺历史，从而找到裂纹产生的具体工序。如裂纹两侧具有明显的氧化和脱碳现象，裂纹的形成肯定与制造热工艺过程有关。淬火工件的裂纹断口颜色发黑，氧化物层厚，说明淬火加热前即已存在裂纹。淬火前就已存在的裂纹，裂纹两侧常有脱碳现象。

在裂纹的外观分析中，还应该注意观察裂纹两侧的耦合情况，一般认为两侧的耦合性很好，但发裂、拉痕、磨削裂纹、折叠裂纹及经过变形后的裂纹，两侧的耦合性均较差。

一般情况下，疲劳裂纹的末端是尖锐的；拉痕、发裂纹的末端圆秃；折叠裂纹的末端粗钝。在金相磨片下观察，淬火裂纹细直、线状、棱角较多、末端尖细；两侧金相组织与其他部分无异常，无氧化、脱碳现象。铸造热裂纹呈龟裂纹状，沿原始晶界延伸，裂纹内侧一般有氧化和脱碳，末端圆秃。磨削裂纹一般细、浅，呈龟裂状或规则直线排列。

由于过热、过烧引起的锻造或热处理裂纹，往往晶粒粗大，并常在晶界处伴有析出物。局部应力超过材料的强度极限所引起的裂纹，裂纹处往往具有明显的塑性变形痕迹。裂纹表面的附着物对裂纹的分析也有一定的参考价值。如水淬时产生的裂纹，会出现红锈。

必要时，可对裂纹进行解剖分析。裂纹解剖分析的目的是分析研究裂纹的起始和走向、经过的路径、裂纹中有无其他物质、裂纹两侧附近区域有无材料变化，从而确定裂纹形成过程与材料显微组织之间的对应关系、断裂过程、断裂机理、变形程度、表面状态及其损伤情况等，以揭示零件在制造、加工等过程中产生的缺陷、使用状况和环境条件等对断裂失效的影响。如对夹杂物、脱碳、增碳、偏析、硬化深度、镀层厚度、晶粒大小、组织结构及热影响区等检查与分析。裂纹的剖面分析具有其他许多分析过程难以确定的优势，一般情况下应该进行此工作。分析过程中，一般需要沿与裂纹垂直的方向切割剖面、磨制金相试样来进行宏观、微观分析，必要时也可利用工业 CT 等无损检测手段在解剖前进行切片分析。为了尽量减少材料损失及热损伤造成的影响，应该尽量采用线切割方法来切割金相剖面，镶制金相试样来进行。

一般张开较大区域为裂纹的起始区，裂纹中夹有氧化物、腐蚀产物等，这说明裂纹形成后（或形成过程中）经历过复杂的环境过程，如高温、腐蚀环境等。裂纹的走向及两侧的材料变化情况往往可以对确定裂纹的性质有重要的帮助作用。如碳钢裂纹两侧脱碳，则说明裂纹为热裂纹或裂纹后经过了热过程；裂纹扩展过程中有无分叉现象，对区分氢脆与应力腐蚀有重要帮助，氢脆裂纹扩展过程中一般无分叉现象，而应力腐蚀裂纹往往有分叉现象。裂纹的扩展是沿晶还是穿晶等也是应该特别注意的问题。一般沿晶扩展的裂纹均与腐蚀介质的作用有关。

2.3.3 裂纹断口分析

(1) 裂纹的打开与断口切取技术

在对裂纹断口进行分析之前。必须人为地将裂纹打开，以获得需要的裂纹断口。有时为了实验室观察的需要，还需要对断口进行选取，需要切取断口。

裂纹和断口是断裂失效过程不同阶段的术语，零件在破断之前称为裂纹，而最终破断的断裂面称为断口。由于断口上较全面地记录了断裂过程中每一阶段的内部的、外部的，力学的、化学的、物理的诸多因素综合作用的结果（痕迹、形貌特征等），断口分析具有比裂纹分析更加全面、准确、直观的优点，因此，往往需要对裂纹人为地打开来进行裂纹的断口分析。

在打开裂纹前，应做好相关的记录、测量和照相，特别是裂纹与相关结构的相对位置、表面的痕迹特征等，以保证裂纹打开后，仍能准确确定裂纹的位置、结构特点、受力状态等，为进一步的分析提供便利。

打开裂纹的方法很多，但不管采用何种方法，都必须满足以下要求：

① 断面要保持原始的形貌特征，不能受到机械的、化学的损伤；

② 断口及其附近区域的材料显微组织不能因为受热发生变化。

打开裂纹的基本原则是，根据裂纹的位置及扩展方向来选择人为加力点，使零件沿裂纹扩展方向受力，使裂纹张开形成断口，而不会在打开过程中损伤断面。常用的裂纹打开方法有三点弯曲法、冲击法、压力法、拉伸法等。当零件厚度大、裂纹小，裂纹难以打开时，可正对裂纹，在裂纹的背面预先加工出一定深度的V形豁口，从而可容易地将裂纹打开。但需要注意的是，加工时不能触及裂纹尖端；当裂纹尖端难以确定时，可适当多留出一些余量，以保证裂纹尖端不会被破坏。打开裂纹时，最好采用一次性快速打开方法，而不用重复的、交变的、分阶段处理的方法，如振动疲劳、反复弯曲等。以免打开时在断面上形成的特征与原始断裂特征相混淆，造成不必要的混乱。

对大型结构件，如锅炉、管道、飞机机翼等，为了便于运输和深入的观察分析，需要将大型零件切割成小试样。常用的切割方法有：砂轮切割、火焰切割、线切割、锯切等，对会产生高温的切割，切割位置应与裂纹保持一定的距离，并用适当的方法进行冷却，以免裂纹附近的材料组织、性能因受热发生变化，断面特征产生化学损伤。

(2) 裂纹的断口分析

裂纹的断口分析与断裂面的断口分析的技术和方法均相同，适用于断裂面断口分析的方法和手段在裂纹断口分析中均可应用；两者的形貌特征、规律也相同。因此，裂纹的断口分析技术和方法可参照前面的断裂面断口分析部分，本节在此不再赘述。

但需要注意的是，一般裂纹是断裂的前一过程，裂纹断口的扩展区一般不是太充分，当裂纹很小时，有时甚至没有扩展区，因此，裂纹断口的特征有时不是太典型，判断起来会有一定的难度，需要慎重。其次，要区分开裂纹断口上的原始断裂特征区和人为打开区，人为打开区的特征应结合打开的方法来判断。一般的裂纹原始断裂区与后续的人为打开区在断口特征上是有明显的区分界限的，宏观、微观特征均有明显的差异，可以很容易地区分开。

2.3.4 裂纹定量诊断技术和方法

裂纹的定量诊断技术和方法与断口的定量诊断技术和方法基本相同，可参阅本章中“断口定量诊断技术和方法部分”，在此不再重复。

2.4 裂纹综合诊断

通过对裂纹的宏观、微观分析，可确定裂纹的部位、形态、裂纹源的位置，初步判断裂纹的形成时期和扩展途径，结合应力分析、制造工艺和使用条件、材料性能综合分析，可初步诊断出裂纹的性质及产生的原因。

(1) 裂纹的起始位置

裂纹的产生是应力作用的结果，其起始的位置取决于应力集中和材料强度两方面综合作用的结果。工件结构形状上易引起应力集中的部位，如工件截面尺寸突变、厚薄不均、孔槽边缘和尖锐棱角处等，往往是裂纹出现的部位。根据裂纹存在的部位和受力状态，可以初步判断裂纹产生的条件。

若裂纹不是起始于应力集中部位，则裂纹存在的部位必然与材料的缺陷和内应力的作用有关，需在应力分析的基础上，结合加工工艺和使用条件，从断口特征、裂纹周围的显微组织缺陷等方面进行综合分析，确定削弱材料强度和引起内应力以至产生裂纹的原因。

1) 材质原因引起的裂纹　金属的表面缺陷，如夹砂、斑疤、划痕、折迭、氧化、脱碳、粗晶环等，以及金属的内部缺陷，如缩孔、气泡、疏松、偏析、夹杂物、白点、过热、过烧、发纹等，不仅本身直接破坏金属的连续性，降低材料的强度和韧性，而且往往在这些缺陷周围造成很大的应力集中，使得材料在很低的平均应力下产生裂纹。

由金属的表面缺陷和内部缺陷为源的裂纹源处，一般可以找到作为裂纹源的缺陷特征。如由砂眼引起的疲劳裂纹，在零件表面或在裂纹断口上可以找到砂眼；由于切削刀痕引起的疲劳裂纹，裂纹源是沿着刀痕分布的；由于残余缩孔引起的锻造裂纹，裂纹从缩孔开始向外扩展，并沿纵向开裂；由于白点引起的淬火裂纹，不仅可看到由白点发展成的裂纹，而且还能看到许多呈短线状的白点。

金属材料光滑表面上的疲劳裂纹多起源于驻留滑移带，而当材料表面或内部存在缺陷时，缺陷本身就成为裂纹源，所以具有缺陷（表面或内部）的金属材料往往具有低的疲劳强度（寿命）。

虽然材料的偏析不破坏材料的连续性，但却使材料的力学性能变得不均匀，故偏析也能成为裂纹源。

2) 零件形状因素原因引起的裂纹　当零件由于某种原因，或者设计上的考虑不周，其几何形状上存在内外圆角、凸边或缺口时，在零件的制造和使用过程中，这些圆角、凸边或缺口的部位会产生应力集中，从而容易产生裂纹。

对需进行淬火处理的零件，不但圆角、凸边或缺口的部位容易产生淬火裂纹，而且凡是截面尺寸相差大的部位，都可因为冷却速度的差异而产生高的组织应力，加之应力集中作用而形成淬火裂纹。

在焊接件的应力集中处，也可能产生焊接裂纹。在深拉或冷冲压时，由于总的变形程度过大、或零件的圆角过小、或材料的晶粒度不均匀，往往在圆角根部（变形程度最大）产生裂纹或开裂。

3) 受力状况不同引起的裂纹　除了金属材料的质量和零件的几何形状影响裂纹的起始位置外，零件的受力状况也对裂纹的起源位置有重要影响。如零件的形状设计合理、材料质量合格的情况下，裂纹将在应力最大处起始。如单向弯曲疲劳情况下，裂纹一般起源于受拉一侧的最大应力处；双向弯曲疲劳情况下，裂纹一般起源于受力两边的最大应力处。

(2) 裂纹的走向

从宏观上看，裂纹的走向是由应力原则和强度原则决定的。

1) 应力原则　按照应力原则，裂纹应该沿着最大应力方向扩展。如金属脆性断裂、疲劳断裂、应力腐蚀断裂，裂纹的扩展方向一般都垂直于拉伸应力的方向。当韧性金属承受扭转载荷或金属在平面应力的情况下，裂纹的扩展方向一般平行于切应力的方向。如塔形轴疲劳情况下，在凹角处起源的疲劳裂纹，在与主应力垂直的方向上扩展，而并不与轴线相垂直，最后形成所谓碟形断口，裂纹的实际扩展方向与主应力的垂线基本垂直，即沿最大应力方向走向。但在局部区域可能有不重合的情况，那是由于材料缺陷引起的走向偏离。

2) 强度原则　当裂纹按应力原则在某一方向的扩展不利时，就会按材料的强度原则来扩展。所谓强度原则，就是指裂纹沿着材料最小阻力路线，即材料的薄弱环节扩展的原则。材料内部的薄弱环节可使按应力原则扩展的裂纹途中突然转折。

在一般情况下，当材质比较均匀时，应力原则起主导作用，裂纹按应力原则扩展；而当材质存在明显的不均匀性时，裂纹按强度原则扩展，强度原则起主导作用。当应力原

则和强度原则一致时，无疑裂纹将沿着一致的方向扩展。这就是存在于最大应力部位的缺陷对裂纹产生的影响很大的原因。

(3) 裂纹的形成原因和形貌特征

各种裂纹的形成原因很多，其形貌也各异。为了对比分析，将各种典型裂纹的形成原因及形貌特征列于表5.2-5。

表5.2-5 金属常见裂纹的名称、形成原因和形态特征

裂纹名称		裂纹形成原因	裂纹形态特征						备注
类别	名称		宏观形貌	源区位置	扩展途径	周围情况	末端	其他特征	
铸造裂纹	热裂纹	形成于约1 250～1 450℃的高温下，形成原因：①在浇铸后的冷却过程中，金属在该温度区间内冷却收缩应力过大；②铸件在铸型中收缩受阻；③冷却严重不均匀；④铸件设计不合理，几何尺寸突变；⑤有害杂质多，并在晶界富集降低金属的强度和塑性；⑥铸件表面和涂料相互作用	龟裂状（网状）	铸件最后凝固区或应力集中区	沿晶界扩展	有严重的氧化脱碳，有时还有明显的偏析、疏松、杂质和孔洞等缺陷	圆秃		
	冷裂纹	形成于较低温度，主要是由于热应力和组织应力造成	应力集中区	穿晶扩展	基本无氧化脱碳，两侧组织和基体相差不大	尖锐			
锻造裂纹	过烧裂纹	轧、锻前的加热温度过高	龟裂状	表面或形状突变处	沿晶扩展	有内氧化和脱碳	严重时呈豆渣状		基本组织也有过热过烧特征
	冷裂纹	终锻温度过低，材料塑性下降，或锻造温度在Ar_3～Ar_1两相区间，铁素体沿晶析出，进一步锻造时沿铁素体开裂	呈对角线或扇形	应力集中处或在晶界铁素体处	穿晶扩展		无明显组织变化		
	热脆裂纹	钢内含硫量过高，锻造加热时在晶界处的FeS－Fe共晶熔化，锻造时开裂	龟裂状	表面或应力集中处	沿晶扩展	有硫化物夹杂		晶界有硫化物夹杂	钢的硫化物级别高
	铜脆裂纹	钢内含铜量较高，或锻造加热时，毛坯表面渗入金属铜	龟裂状	表面或应力集中处	沿晶扩展	有铜夹杂或氧化铜夹杂		晶界有铜	
	折叠	表面突起部位被折叠	由表面开始倾斜	表面层		有氧化皮及脱碳层			
	加热不足	轧、锻前加热保温时间不够，芯部未热透；高合金钢中心碳化物偏析严重	放射状	锻件芯部	穿晶扩展	稍有氧化脱碳现象或碳化物偏析			有的碳化物偏析严重
	皮下气泡锻裂	皮下气泡未清除干净	与表面垂直	次表面皮下气泡处	穿晶扩展	有时有氧化现象			一般较浅
	锻坯缩孔未清除	钢锭切头不足	缩孔顺变形方向拉长	中心部位	沿晶	表面有氧化物			
	锻比大，锻速快	方坯对角线部位由中心起开裂，由变形热升温引起	交叉裂纹	锻件芯部开始	穿晶	有氧化层	尖锐		
焊接裂纹	冷裂纹	在温度100～300℃之间，因热应力和组织应力的共同作用下产生，特别是由于100～300℃温度范围内氢气析出及聚集作用的结果		应力集中处或组织过渡区（在热影响区内）	一般穿晶扩展	很少氧化脱碳			
	热裂纹	钢在1 100～1 300℃之间，因热应力作用产生。形成热裂纹的可能性与基体金属、焊条金属的成分有很大关系，一般合金钢或含碳量高、强度大的钢，含氧量大的铜合金及使用低熔点焊条的铝合金产生裂纹的可能性较大	有时呈蟹脚状、网络状或曲线状	一般起源于焊缝区内	沿晶界扩展	有氧化脱碳，有时还有焊料			
	熔合线裂纹	热应力过大或表面有残留氧化物等		一般在熔合线处	一般穿晶扩展				

续表 5.2-5

裂纹名称		裂纹形成原因	裂纹形态特征						备注
类别	名称		宏观形貌	源区位置	扩展途径	周围情况	末端	其他特征	
热处理裂纹	淬火龟裂纹	表面脱碳的高碳钢零件，在淬火时，因表面层金属的比容比中心小，在拉应力作用下产生龟裂纹	龟裂纹	脱碳表面层	沿晶扩展	很少氧化		一般限于脱碳层内	一般较浅
	淬火直裂纹	细长零件，在芯部完全淬透的情况下，由于热应力和组织应力共同作用而产生的纵向直线淬火裂纹	纵向直线裂纹	应力集中或夹杂处	穿晶扩展	很少氧化	尖细		
	过热裂纹	淬火加热温度过高，产生了过热或过烧，削弱了晶界的强度，致使在组织和热应力作用下沿晶开裂	网状或弧形	应力集中处	沿晶扩展	很少氧化	尖细		
	其他淬火裂纹	凹槽、缺口处因冷却深度较小，产生局部未淬透或软点附近的组织过渡或偏析区在拉应力作用下开裂	一般呈弧形裂纹	应力集中处或组织过渡区	穿晶扩展	很少氧化	尖细		
	回火裂纹	具有回火脆性的钢在回火脆性范围内回火，冷却速度又小或零件厚度太大，引起回火脆性，在随后的校直或使用中开裂		一般在应力集中处	主要沿晶扩展				
加工裂纹	磨削裂纹	由于磨削热引起组织转变（如残余奥氏体转变为马氏体）和应力再分配等原因引起	龟裂纹状或呈辐射状，或规则排列	在被磨削的金属表面层内	主要沿晶扩展	稍有氧化	呈喇叭口状		
	皱裂纹	由于表面张力而引起	龟裂纹	沿纤维方向	穿晶		尖细		
使用裂纹	应力腐蚀裂纹	在腐蚀介质和拉应力的共同作用下产生	有时呈网络状	与腐蚀介质接触并受拉应力的表面					
	蠕变裂纹	金属在高温下工作时		应力集中处	沿晶扩展	严重氧化			
	疲劳裂纹	在交变应力作用下产生		多数在表面应力集中处	主要呈穿晶扩展	有时可观察到有金属磨屑			
	延性撕裂	所受载荷超过金属的强度极限而开裂	断裂与切应力方向平行	一般在应力集中处	穿晶扩展				

3　痕迹诊断技术和方法

在失效分析中，痕迹分析往往是一个有效的分析手段和技术。通过痕迹分析，不仅可对事故和失效的发生、发展过程做出判断，而且可为事故和失效分析结论提供可靠的佐证和依据。因此，在分析中应充分重视痕迹分析。

3.1　痕迹及痕迹分析概述

广义地说，痕迹是指环境作用于系统，在系统表面留下的标记。在机械事故调查和失效分析中，系统指的是失效的构件或零件，而环境则是指外来的力学、化学、热学、电学等因素。因此，对于机械失效时所留下的痕迹，可以定义为：力学、化学、热学、电学等环境因素单独地或协同地作用于机械，并在它的表面或表面层留下的损伤性标记，称为机械表面痕迹，简称为痕迹。

对完整表面，痕迹的含义包括表面形貌（特征）的变化、成分的变化（或材料的迁移）、颜色的变化、表层的变化、表层组织、性能的变化、残余应力的变化，以及表面污染状态的变化等。痕迹分析即是对上述变化特征进行诊断鉴别，并找出其变化的原因，为事故调查和失效分析提供线索和依据的技术活动。

由于痕迹暴露于表面，容易受外来因素的干扰和破坏，因此，痕迹往往缺乏连续性，可以重叠，可以反复产生和涂抹，有时仅记录了最后一幕，因此，痕迹分析需要采用综合分析的手段。

在失效分析范畴内，痕迹的具体含义可以归纳为：

① 痕迹的形貌（或称花样），特别是塑性变形、反应产物、变色区、分离物和污染物的具体性质、尺寸、数量分布；

② 痕迹区以及污染物、反应产物的化学成分；

③ 痕迹颜色的种类、色度和分布、反光等；

④ 痕迹区材料的组织和结构；

⑤ 痕迹区的表面性能（耐磨性、耐蚀性、显微硬度、表面电阻、涂镀层的结合力等等）；

⑥ 痕迹区的残余应力分布；

⑦ 从痕迹区散发出来的各种气味；

⑧ 痕迹区的电荷分布和磁性等。

在一般情况下，痕迹分析应按如下程序进行。

1）寻找、发现和显现痕迹　这是痕迹分析的前提和基础，寻找痕迹时应以现场为起点，全面收集证据。一般应首先收集能反应整体破坏顺序的痕迹，其次收集具体零部件外部的痕迹，再收集零部件之间的痕迹，最后搜集污染物和分离物。搜集时应注意痕迹所在的部位、相对关系等。

2）痕迹的提取、固定、显现、清洗、记录和保存　痕迹的提取和固定方法有复印、制模法、静电法、AC 纸黏附法等。痕迹的记录可以用文字、示意图和照相法。其中照相法是最重要的方法，照相时应通过整体、局部、重点痕迹、痕迹特征点等一系列的照片来反应痕迹所处的部位、特征、相互关系。痕迹照相时，应注意角度和光线，以保证痕迹特征的清晰，同时没有假象显现。

3）鉴定痕迹　对具体的痕迹特征进行针对性检验，从而确定痕迹的性质、产生的时间和条件等。鉴定时，应遵循由表及里、由简而繁、先宏观后微观、先定性后定量，按形貌－组织结构－性能的顺序进行的原则。鉴定时，应掌握零件的工作原理、过程，有无异常发生等履历情况。

4）痕迹的模拟再现试验　模拟再现的难度很大，需要在上述工作充分完成以后才能进行。简单的模拟试验可以在模塑制品（如塑料、蜡、特制胶泥）上进行，必要时可在产品上进行，拆检同型号的、已使用过的产品的相应痕迹加以对比也是一种更真实的“试验”。

5）综合分析　由于失效往往是多种因素共同作用的结果，痕迹处于表面，更是经历了复杂的过程，因此痕迹分析需要采取综合分析的方法，要考虑到痕迹的形成过程、形成条件、影响因素，痕迹的可变性，痕迹与零件工作的关系、失效的关系等。

3.2　痕迹的发现和显现技术和方法

痕迹的发现和显现就是将隐藏的、不明显的痕迹特征揭示出来，以确定痕迹的分布及其规律，为进一步的痕迹性质鉴定提供目标。

为了发现痕迹，首先应弄清失效件本身的表面状况，包括表面颜色、原始形貌特征、粗糙度、材料成分与组织等原始基本情况，当对失效件不熟悉时，这些基本情况更需要仔细了解。由于痕迹是一种表面特征，因此，寻找痕迹应在零部件的表面进行，寻找的重点也应根据地点、机件大小等情况采取不同的方法。对大的机件，如飞机、汽车、大的压力容器等的现场调查中，应着重从零部件的表面颜色变化、表面结构轮廓变化、形貌变化等肉眼易见的特征上来发现痕迹；当零部件较小、可以在实验室借助一定的仪器设备来进行时，在以上检查的基础上，应重点从粗糙度变化、细小附着物、擦痕、划痕、材料成分、组织等方面来发现痕迹特征。

由于失效时的条件一般均很复杂，因此，零部件表面一般容易覆盖有很多的附着物，为了将痕迹暴露出来，需要将这些后来的附着物（痕迹）区分、去除。一般先用软毛刷将表面的浮尘、泥土等扫干净（尽量不要用水和其他溶剂），以显现出痕迹的颜色特征。为了准确显现机械痕迹，应将表面的附着物清除干净，可以采用清洗、粘揭的办法，使划痕、擦痕等特征暴露无遗。材料的成分、组织痕迹一般需要切取与表面垂直的剖面才能显现出来。

3.3　痕迹的提取、固定、清洗、记录和保存技术和方法

痕迹的正确解释，常取决于对外来物（污染）影响的排除程度，因为外来物的影响使痕迹特征模糊；可导致做出错误的结论。

（1）基本原则

为了得到痕迹的准确信息，对痕迹进行处理时，应遵循以下原则。

① 尽量避免机械损伤　在搜集、运输带痕迹的残骸时，应对残骸进行适当的包扎，以避免碰撞；也不要用任何东西去擦拭痕迹或用手去触摸痕迹。

② 防止化学损伤（腐蚀）　可把带痕迹的残骸放于干燥器中，或浸入无水酒精的密封容器中。不要在痕迹区涂防腐剂，以免干扰对痕迹区的鉴定。

③ 防止痕迹区的松散附着物的剥落　一般这些痕迹是说明痕迹成因的重要线索。不得已时，可将这些剥落的附着物收集以供分析。

④ 避免环境中的粉尘、纤维、水气等附着在痕迹上，以免给分析造成假象。

⑤ 涂层　当不必进一步分析痕迹上的外来物时，可用涂层的办法来保护痕迹。

（2）痕迹的清理技术

清理痕迹的目的是去除保护涂层、腐蚀产物、灰尘之类的松散沉积物。常用的清理技术有以下几种。

① 机械刷洗法　吹干燥空气，或用软毛刷清理。

② 有机溶剂清洗　主要用于去除痕迹表面的油污、有机污染物等。常用有机溶剂有：汽油、丙酮、三氯甲烷、甲苯、乙醚、石油醚等。有时可辅以超声清洗。

③ 弱酸或碱性溶液处理法　用来去除高温氧化产物。需要注意的是，无论采用何种溶液，都应只对表面沉积物起作用，而不能浸蚀基体材料。

④ 超声清洗法。

3.4　痕迹的鉴定

3.4.1　痕迹的形貌诊断

痕迹的形貌诊断是痕迹分析的重要内容。在痕迹形貌诊断中，应首先观察痕迹的整体分布特征与规律，并画出示意图，从而进一步明确其分布规律。此阶段的分析以肉眼观察为主，观察的重点是在确定了痕迹的整体分布规律以后，重点应放在痕迹性质的确定上，以有代表性的局部痕迹鉴定为主。

（1）形成痕迹的要素

一般来说，形成痕迹包含有三个基本要素。

1）造痕物　它是痕迹的制造者，是直接接触并作用于机械表面的物体或介质（也称痕迹形成物），能把自身的某些特征标记遗留在机械表面上。

2）留痕物　它是造痕物作用的对象（也称痕迹接收物），是痕迹的接收载体，在机械失效分析中，一般就是指机械表面。

造痕物与留痕物是相对的，有时在两个匹配的接触面都会留下痕迹。

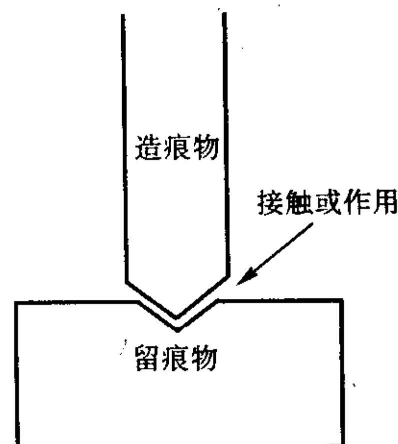


图 5.2-34　痕迹的形成

3）发生相互接触或非正常作用　只有相互接触或作用，才能留下痕迹。图 5.2-34 示出了痕迹的形成原理。当造痕

物与留痕物发生相互接触或作用，痕迹就会形成。对机械痕迹，造痕物与留痕物之间的这种作用就是力的作用，而且一般为压应力。

(2) 机械的表面特征

材料的表面状态在很大程度上决定了材料的表面性能。机械上的各种痕迹均发生并存在于表面，研究痕迹、识别和诊断痕迹就必须从材料的表面形貌和状态着手。

1) 表面加工纹理　任何机械产品的表面，都是加工的产物，不同的加工方法和工艺参数，会留下相应的表面形态特征，亦称作表面纹理，其中加工方法等因素形成的表面微观结构的主要方向，即表面结构纹理，俗称加工痕迹，如图5.2-35所示的表面切削加工痕迹。

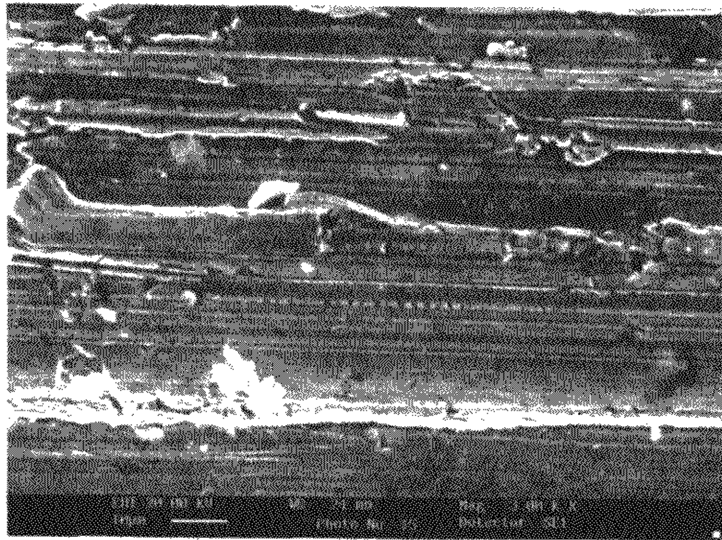


图5.2-35　表面切削加工痕迹

由于表面纹理对生产工序中的变化非常敏感，因此，可以作为一种重要的质量检验方法来应用。通过检查加工零件表面留下的纹理，常可发现工具或车床在使用和操作中的缺点。机械痕迹分析时，首先要熟识分析对象的表面纹理，也就是原始的表面痕迹，从而才有可能判断哪些痕迹是使用中形成的痕迹。

2) 表面不平度　机械表面的不平度包括宏观形状误差、粗糙度和波纹度。

3) 表面的结构　机械表面的痕迹与机械表面层结构有关。金属的加工表面不仅表现出一定的几何形状，而且表面的特殊结构和所处的环境还使其在物理、机械、电学以及化学性能等方面也区别于内部基体组织。一般来说，金属的最表层由厚度约3 nm的污染层、厚度为0.3 nm的气体吸附层和10 nm左右的金属氧化层组成。对单一晶体，表面会出现结构缺陷（如空位、台阶），还会发生晶粒长大和溶解、吸附或化学反应。

此外，金属表层中还存在化学缺陷，如金属夹杂物、成分偏析、外来间歇原子等，以及应力缺陷，如加工后的有害残余应力等。金属表面在加工过程中还要形成硬化层。

当表面有污垢或氧化皮时，不但影响次表层的强度和吸附情况，而且还影响表面的各种物理性能。

(3) 机械失效中的痕迹分类

根据痕迹形成的机理和条件的不同，可将痕迹分为以下七类：

① 机械接触痕迹；
② 腐蚀痕迹；
③ 电侵蚀痕迹；
④ 污染痕迹；
⑤ 分离物痕迹；
⑥ 热损伤痕迹；
⑦ 加工痕迹。

加工痕迹是在已知生产条件下的产物，因此规律性较强。

值得特别注意的是那些可能导致机械失效的非正常加工痕迹，即留在表面的各种加工缺陷，如挤压裂口、啃刀痕迹、磨削烧伤痕迹等。

(4) 机械表面接触时引起的各种效应

当机械表面在力、热、电等的作用下接触时，接触表面将发生一系列的材料、物理、机械和化学性能的改变，具体表现在如下几个方面：

① 塑性变形；
② 温度升高；
③ 相和组织的转变；
④ 化学成分的变化和表面膜的形成；
⑤ 材料的转移；
⑥ 材料的磨损；
⑦ 表面应力状态的转变；
⑧ 表层材料的性能改变。

从材料的角度分析，接触面的材料可能发生：

① 扩散和黏着转移；
② 脱离成为分离物离开表面；
③ 化学反应生成腐蚀产物；
④ 电侵蚀时的飞溅、烧蚀；
⑤ 附着物（如污染物、吸附物）的转移等。

(5) 机械接触痕迹

由于机械力的作用而在接触部位留下的痕迹称为机械接触痕迹，简称机械痕迹。机械痕迹的特点是，存在塑性变形或材料转移、断裂等现象，且这些现象均集中发生在接触部位；材料的塑性变形极不均匀。

依据接触方式（接触面的形状和尺寸，接触角和方位等）和相对机械运动方式的不同，机械接触痕迹又可分成以下五种：

① 压入性机械痕迹；
② 撞击性机械痕迹；
③ 滑动性机械痕迹；
④ 滚压性机械痕迹；
⑤ 微动性机械痕迹。

1) 压入性机械痕迹　造痕物压入留痕物时，法向载荷的作用缓慢而持续，变形速度一般很小，保持较长时间的接触状态或接触面不再分离，这时留下的痕迹称为压入性机械痕迹，简称压痕。压入性机械痕迹的典型形貌是容积性的压坑、压伤、压陷和压痕。

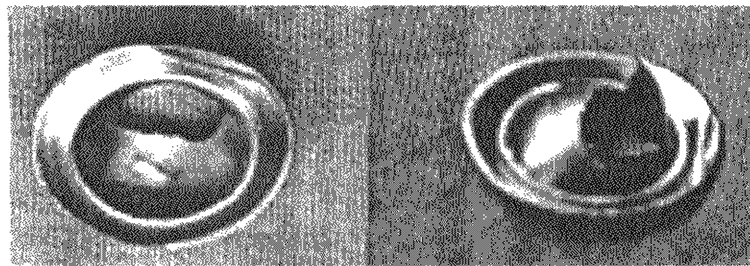
常见的最典型的压入性机械痕迹有压入法测量金属材料硬度时在金属表面留下的各种规则的印痕，在零件表面敲上的钢印编号、标记等。

在机械加工过程中留下的有害压痕，是机械失效的重要原因之一。例如，某型飞机座舱抛盖活门薄膜表面在制造时存在压痕，导致薄膜在使用中从压痕底部开始破裂（见图5.2-36），险些酿成事故。某型飞机用GC－11钢制造20框时经过冷敲，结果在镀锌后沿着冷敲点（压痕）与基体的交界处产生氢脆裂纹，然后沿敲击痕周边成放射状扩展，是因为交界处正是局部应变不连续处，表面存在有害的拉应力（压痕坑底表面为压应力）。某型飞机机翼大梁在加工过程中留下的冲眼（也是一种典型的压痕），曾使该梁早期疲劳破坏。

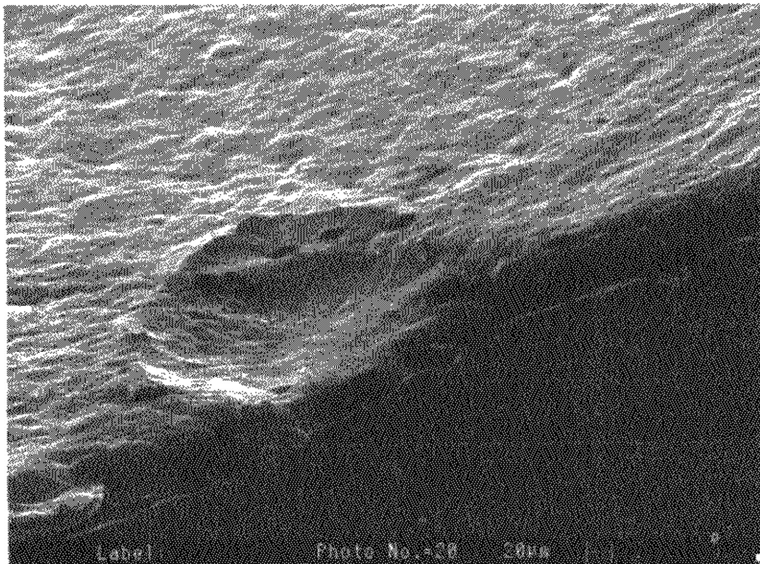
在机械事故检查和失效分析时，最常用的压入性机械痕迹分析有：

① 确定发生事故（或故障）时机件之间相对工作位置的卡压痕迹；
② 确定仪表指示位置的卡压痕迹；
③ 外来物的卡压痕迹；
④ 反映解体顺序的机件印痕。

在失效件上，压入性机械痕迹往往与其他痕迹同时存在，它们之间一般具有如下的关系。



(a) 破裂的薄膜宏观形貌



(b) 薄膜上的压痕放大

图 5.2-36　飞机座舱抛盖活门薄膜表面压痕导致薄膜破裂

① 钢珠在镀铬零件上压入，会造成圆形的压坑，并且在压坑的底部和边缘铬层呈现脆性的网状龟裂纹。

② 在既有凹陷变形又有贯穿凹陷变形的连续划痕时，则划痕产生于凹陷变形之前。

③ 重叠压痕先后顺序的判断，一般是后面的压痕覆盖前面的压痕；最后形成的压痕外形最完整（形貌完整、边界清晰、特征变异小）；最早形成的压痕外形最不完整；小坑可建立在大坑上，大坑可以覆盖小坑。

图 5.2-37 是重叠压痕形成先后顺序的判断示意图，图中 3 号压痕的外形最完整，2 号压痕次之，1 号压痕的外形最不完整。因此，可以判断，最先形成的是 1 号压痕，其次形成的是 2 号压痕，3 号压痕最后形成。

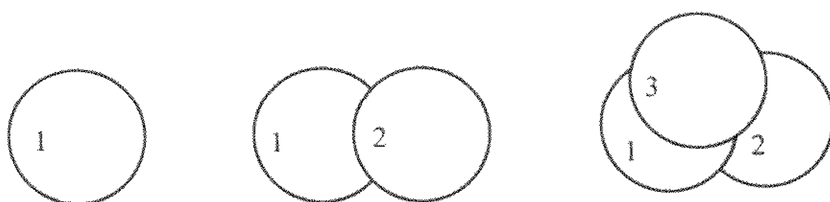


图 5.2-37　重叠压印痕先后顺序的判断

根据痕迹的以下特征可判断其是否属于压入性机械痕迹以及形成压入性机械痕迹的造痕物的外形特征。

① 压印痕一般形貌比较规则，边界比较清晰，与造痕物的接触部位的形状比较吻合，能较好地反映造痕物的几何特征。

② 压入性机械痕迹在垂直表面的方向上的变形最大，往往形成容积性的压印痕。

③ 压印痕的面积大小和深度与作用在压印头上的法向载荷成正比、与压印头的接触面积成反比；在高温状态下产生的压印痕要比室温下的大，且随载荷保持时间延长，压印痕将逐渐增大。

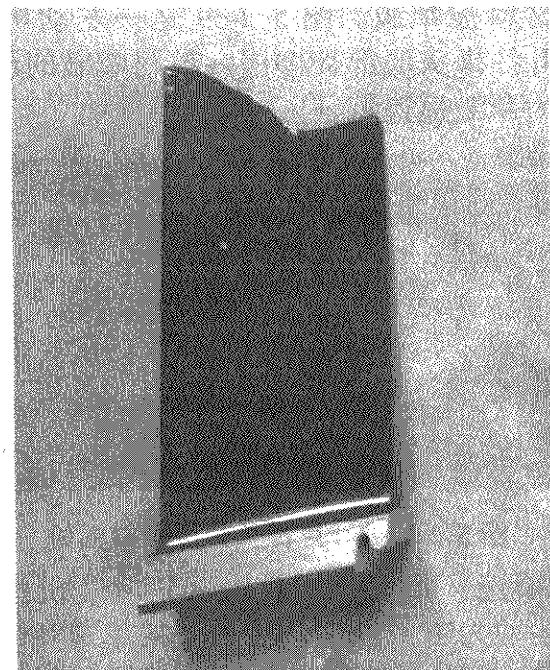
④ 纯粹的压印痕不会产生磨粒，在压印痕表面很少发现有材料粘着和转移，因此压印痕是最干净的痕迹。

⑤ 不同性质的材料，其压印痕不相同。

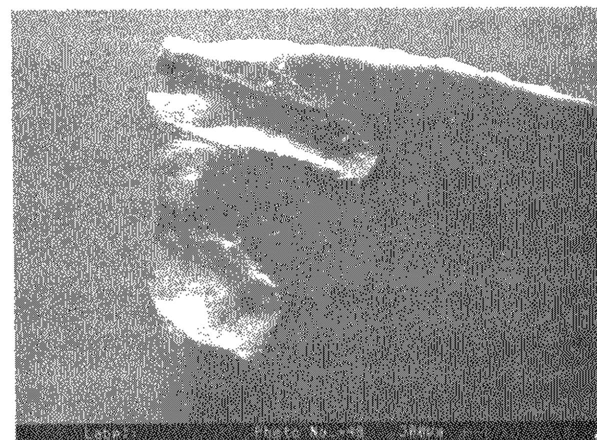
根据压印痕的以上特点，在众多紊乱的压痕中，要选择形状完整、特征清晰的一次形成（或最后形成）的痕迹部位作深入分析。

2）撞击性机械痕迹　造痕物原来不与留痕物接触，只在撞击时才接触；机械力作用的时间很短，变形速度较大，但在接触面之间以垂直接触面方向的相对运动为主。例如飞机坠毁时的接地痕迹，鸟撞击机翼前缘的痕迹，跑道上石子打伤发动机叶片的痕迹等。在撞击载荷下，不仅会产生比静载荷要严重得多的应力状态，而且材料的性能也有重大变化。

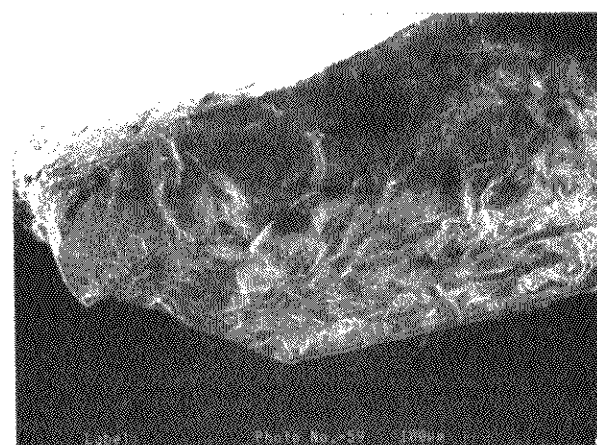
图 5.2-38 是一典型的发动机压气机叶片被打伤，导致疲劳裂纹从打伤痕迹边缘起始，叶片断裂的例子。对高速运动的机件，打伤往往可能导致严重的后果，因此，应采取有效的预防措施，以避免打伤的发生。



(a) 被打伤的叶片全貌



(b) a 中打伤痕迹的放大



(c) 从打伤痕迹处起源的疲劳断口

图 5.2-38　发动机压气机叶片断裂源处的打伤痕迹

撞击性机械痕迹又可分为：

① 外物打伤痕迹；

② 多次撞击表面疲劳痕迹；

③ 冲蚀损伤痕迹。

3）滑动性机械痕迹　造痕物和留痕物的接触面在痕迹形成过程中不断相对移动、分离，移动速度可快可慢，机械作用力和变形方向大体上平行于接触面。例如，机械表面的各种划痕、磨痕、刮痕，飞机轮胎在跑道上留下的拖胎痕迹，以及摩擦面上留下的摩擦痕迹和磨损痕迹。

应该说，滑动性的机械痕迹都是在摩擦过程中形成的；没有摩擦，就不会产生滑动性的机械痕迹。因此，滑动性的机械痕迹又可成为摩擦痕迹，它是机械表面最常见的机械痕迹。

滑动性机械痕迹按其形成机理和特征，可分为以下四类。

① 犁痕（即划痕）　犁痕是滑动性机械痕迹中最常见的一种，它还可分为以下三种。

a）犁沟痕迹　造痕物为钝犁头滑过软的机械表面，以塑性变形为主要机制（犁皱）而留下的犁沟，称为犁沟痕迹。其特点是，痕迹区只产生塑性变形，但不产生切屑。整条犁沟表面比较圆滑，沟边和前缘（犁沟结束部位）有材料堆积，沟底材料被犁皱。图5.2-39是一发动机叶片安装时强行装配，叶片榫头底部被榫槽内材料犁皱的形貌特征，箭头所示部位为犁皱痕迹特征。

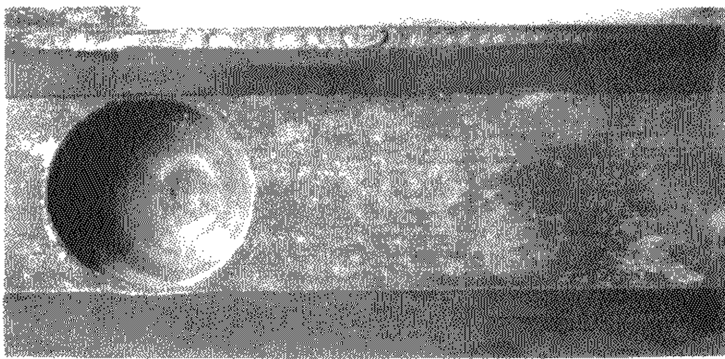


图5.2-39　叶片榫头边缘犁皱痕迹

b）犁削痕迹　犁刀头（或锐器）滑过机械表面，以微切削或切削运动为主要机制，而留下的犁槽和犁屑称为犁削痕迹。特点是，产生微切屑或切削痕迹区的塑性变形局限在更小的范围，变形更不均匀，槽边常有毛刺，槽底常有更细微的划痕，槽侧有剪切皱折。

c）犁碎痕迹　常发生在脆性材料中，包括裂纹和破碎。

犁痕也是整个机械痕迹中比较常见的一种痕迹，准确判断其形成的条件、方向和机理对诊断整个事故或失效的形成过程和原因具有重要的意义，其中尤以犁痕方向的判断更具实际意义，也是痕迹诊断中最基本的内容。

犁痕的判断，总的来说，需要从犁痕的起始、末端、沟边和沟底四个部位的宏观、微观特征去鉴别，其中特别要注意细微形貌和材料转移特征。犁痕的方向，一般可根据以下特点来鉴别。

a）直接犁入的犁痕，起点处不但没有材料堆积，而且往往出现凹陷，如图5.2-40左边所示。

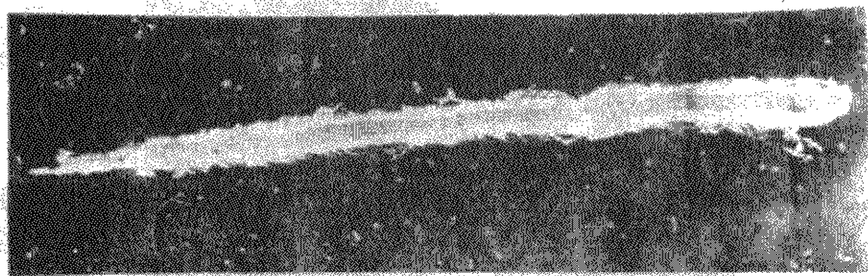


图5.2-40　犁痕方向的判断

b）如果先形成压印痕，再发展成划痕，则起点处会留下压印痕的特征

c）一次性的划痕，结尾往往带有突然性，在犁痕的末端有比较明显的材料堆积，如图5.2-41所示。

d）当划痕过程中途经表面凹凸处时，其形成方向可以借助该凹凸处材料的变形或堆积的位置的形状，以及划痕的中断特征来加以判断。图5.2-42示意地表示出了划痕经过凸起时，材料碎渣在凸起前缘的堆积情况。

e）一般金属材料向犁沟外侧的两边翻起，翻起的金属毛刺的倾斜方向为表面犁沟的形成方向。从图5.2-40可以看出，犁痕的方向为从左向右，犁痕两侧的金属毛刺呈约30°的方向向右倾斜，指示出了犁痕的形成方向。

f）撞击型的犁痕，由于造痕物对留痕物的作用力逐渐减小，因此，划痕宽度会由粗变细，深度会由深变浅，材料转移会由多变少；犁痕的宏观形状呈收敛形，收敛的方向就是犁痕的形成方向。

对橡胶类高弹性体，由于其弹性变形大，一般不易留下连续的犁痕，而在犁痕的轴线方向上，橡胶碎渣积聚成弓形排列，弓形突出的方向即为犁痕的形成方向。

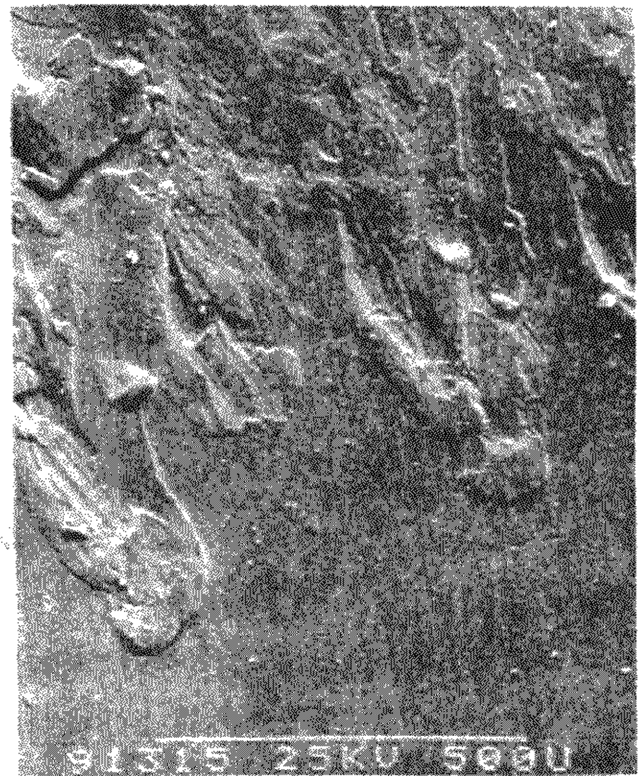


图5.2-41　犁痕末端的材料堆积

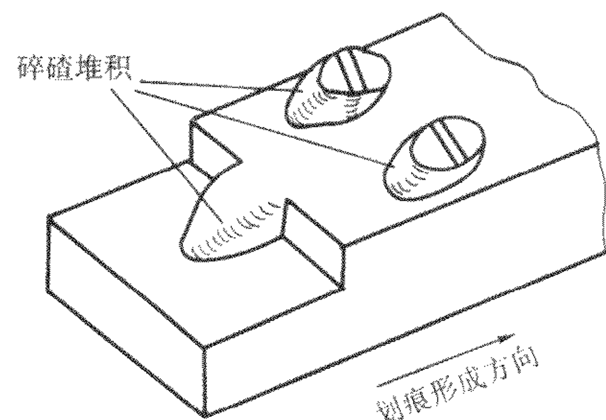


图5.2-42　碎渣在表面凸起处堆积

根据犁痕的其他一些特征，还可以对犁痕的形成过程和条件、造痕物的特征等进行判断。如犁痕平稳，沟宽基本不变，沟底为平行的细微划痕，沟边缘成脊状，则说明犁痕的形成过程中法向载荷基本稳定，没有明显的变化。如果犁痕的宽度逐渐变窄、深度逐渐变浅，末端出现隆起，则说明形成犁痕的法向载荷在逐渐减小；反之，则说明载荷在逐渐加大。

在实际失效件中，经常还会遇到多条划痕相交的现象，这就需要区别这些划痕形成的先后顺序。一般，相交划痕的先后顺序可以根据以下规律来进行判别。

a）如果第二条划痕的细划痕覆盖了第一条划痕的细划痕，这时第一条划痕沟底的细划痕在相交处突然中断（或发生转折、变形）说明第一条划痕先出现。在划痕沟底较好地保持有连续而平行于沟边方向的细微划痕一般是较晚出现的划痕。

图5.2-43是三条相交划痕的形成顺序示意图。3号划痕连续、完整，是最后出现的；而1号划痕在与2号、3号划痕的相交处均有中断，是最先出现的。

b）浅划痕遇到深划痕时，则浅划痕在深划痕的沟边出现不连续现象，浅划痕呈断续状。但有时可使深划痕沟底的细划道顺划痕方向凸起。在迎划痕的深划痕的沟边有可能产生涂抹材料的堆积。

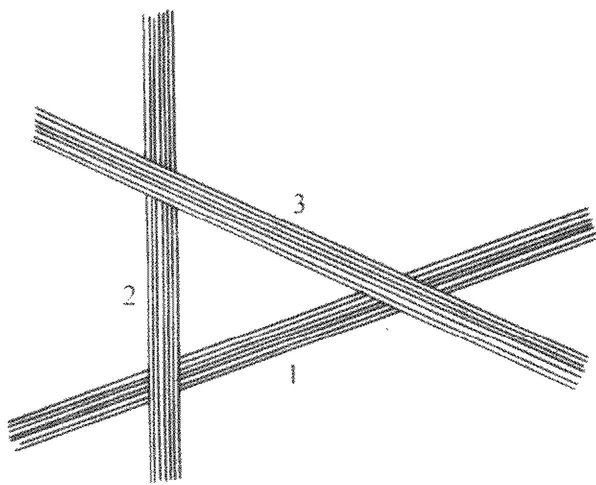


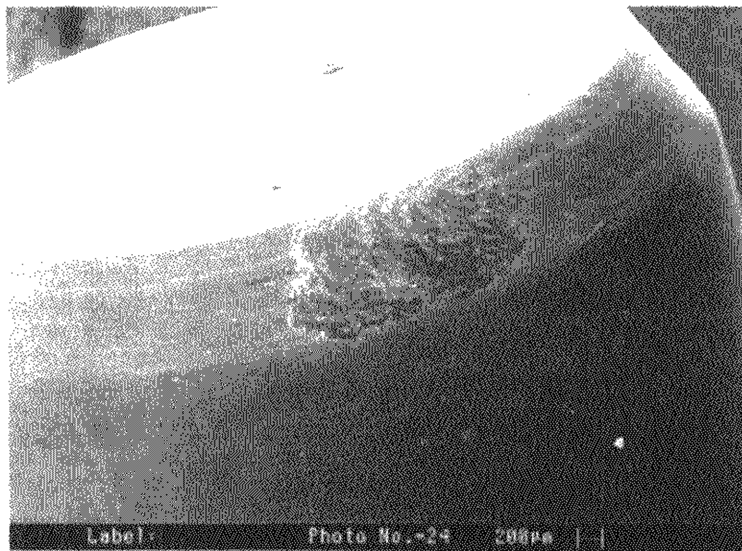
图 5.2-43　相交的划痕（顺序：1、2、3）

c）深划痕划过浅划痕时，将迫使浅划痕中断（犁断），交叉相遇处浅划痕的沟边顺着擦划的方向变形，在交叉处出现“收口”，由此也可反推深划痕的划痕形成方向。

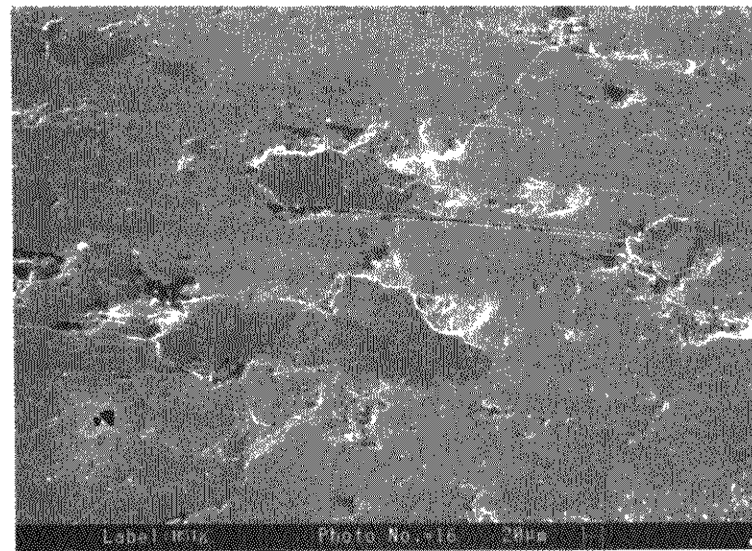
d）涂抹型划痕遇到原有划痕时，常使原有划痕覆盖而中断。

② 黏着痕迹　滑块在切向应力作用下滑过机械表面，接触面之间由于黏着接点发生局部剪切断裂，而留下的切断形貌和材料转移称为黏着痕迹。

黏着痕迹的主要特点是出现接触面的材料转移，凸起和转移处有相对应的撕脱（具有延性破坏特征凹坑），接触面变得比较粗糙。图 5.2-44 是一插销上黏着有其他金属的痕迹形貌。



(a) 黏着痕迹宏观形貌



(b) 金属黏着的微观形貌

图 5.2-44　黏着痕迹

在发生黏着接触时，通常是内聚力较弱的金属黏着和移附到内聚力较强的金属上。金属表面之间黏着接点断开后的韧窝结构，是典型的延性断口。

带漆层、镀层和涂层的造痕物，往往会在痕迹上留下相应的表面层，据此可找到造痕物。1908 年，世界上第一次正式的飞行事故调查就是根据垂尾铝合金张线上抹有螺旋桨桨叶上的漆，从而判断出桨叶先折断，然后打断了垂尾的张线，导致飞机失去操纵而坠毁的结论的。

实际上任何摩擦副都可能发生黏附，尤其在高速滑动时。

③ 摩擦疲劳痕迹　接触斑点在实际的随机交变接触应力作用下，由于微区疲劳断裂而留下的表面裂纹和表层材料的分离脱落（或磨屑）称为摩擦疲劳。其主要特点是接触表面出现各种封闭式微裂纹并产生磨屑。

摩擦疲劳痕迹实质上是一种复合痕迹，它不是一次或若干次微峰接触运动所能形成的，而是反复犁划、多次冲击或反复黏着等运动导致的裂纹和微断裂现象。

摩擦疲劳属于表面接触疲劳。

当摩擦材料表面的微体积受到一定的接触循环交变应力的作用时，在表层内将萌生裂纹，由于裂纹逐渐扩展到表面，导致表面产生片状或颗粒状磨屑。在摩擦时，由于表面温度升高，还可能产生热疲劳。磨屑脱落后表面形成豆状的坑，这种磨损形态即摩擦疲劳痕迹。

④ 摩擦腐蚀痕迹　在腐蚀环境（包括腐蚀性气体、液体介质）下，由于摩擦面材料起化学反应或电化学反应（生成腐蚀产物）而引起腐蚀和磨损称为摩擦腐蚀。图 5.2-45 是金属表面生产氧化膜，氧化膜不断形成和脱离引起磨损的示意图。

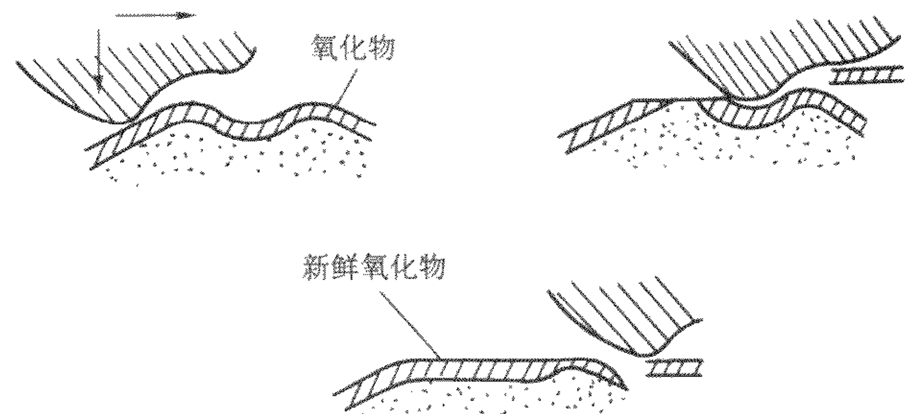


图 5.2-45　氧化膜不断形成和脱离引起的磨损

摩擦腐蚀是在腐蚀作用和机械作用共同参与下完成的。

4）滚动性机械痕迹　单纯的滚动不会在机械表面留下痕迹。工程上所说的滚动性机械痕迹，实际上都是滚压或滚滑性机械痕迹，简称为滚动痕迹。

工程中最常见的一次性滚动痕迹要算各种轮胎滚压痕迹和履带滚压痕迹了。

工程中最常见的多次性（或反复）滚压痕迹是滚动接触疲劳痕迹，其典型特征是麻点、点蚀、痘斑和片状剥落。

滚动接触疲劳是以点、线接触的摩擦副为对象，如滚动轴承和齿轮传动。因反复进行的滚动接触而引起的表面疲劳称为滚动接触疲劳。

当两接触表面是纯滑动时，裂纹源在表面；如为纯滚动，则裂纹源在表层下。

5）微动性机械痕迹　微动就是名义上相对静止的两个固体，其相互接触的表面在法向压力作用下互相挤压并产生往复的相对滑动。相对滑动幅度在 5～400 μm。

微动性损伤有：微动疲劳、微动磨损、微动腐蚀等。图 5.2-46 是一零件表面产生的微动疲劳形貌特征。

（6）腐蚀痕迹

金属腐蚀是指金属与周围介质发生化学或电化学作用，变成化合态物质，从而导致金属损伤的各种转变过程。这里出现腐蚀产物是一个必要条件。

当腐蚀作用使设备不能完成它设计的功能时，就是腐蚀失效。工程中常常遇到腐蚀与其他失效形式协同作用，产生更为严重的组合失效形式，如腐蚀磨损、腐蚀疲劳、热腐

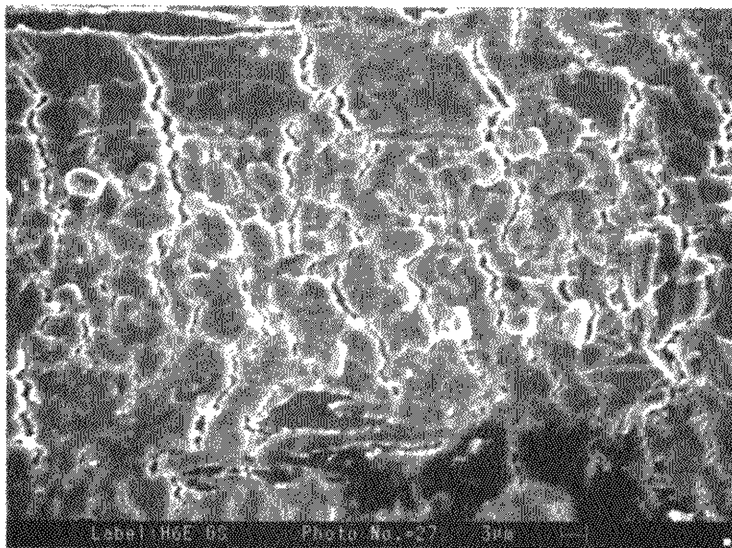


图 5.2-46　微动疲劳形貌

蚀、应力腐蚀等等。

金属的腐蚀不仅使材料的刚度、强度下降，而且有可能成为新的疲劳源（见图 5.2-47），使疲劳寿命也大大下降。就目前情况看，大中型飞机的主要承力件产生裂纹的故障不多，但结构腐蚀却相当严重。



图 5.2-47　裂纹产生于腐蚀坑

1）腐蚀的分类

① 按腐蚀的原理不同分类

a）化学腐蚀　金属表面与环境介质发生化学作用而引起的腐蚀。金属在干燥的气体介质中和在不导电的液体介质中发生的腐蚀，都属于化学腐蚀。

b）电化学腐蚀　金属在导电的液态电介质中由于电化学作用而导致的腐蚀。

② 按腐蚀破坏形式分类

a）全面腐蚀　腐蚀作用全面地分布在整个金属表面上，腐蚀速率大体相同。

b）局部腐蚀　腐蚀作用集中在某一定的区域内，而金属的其余部分却没有发生腐蚀。局部腐蚀一般为电化学腐蚀。

③ 按具体的腐蚀环境分类

a）大气腐蚀；

b）在电解质溶液中的腐蚀；

c）海水腐蚀；

d）土壤腐蚀；

e）熔盐腐蚀；

f）有机气氛腐蚀；

g）生物腐蚀；

h）特殊环境的腐蚀。

2）表面腐蚀痕迹的鉴别　零件是否发生了腐蚀，可以从表面颜色是否发生了变化、形貌是否有改变、表面层的化学成分是否有改变、物资结构是否有变化和导电、导热、表面电阻等表面性能的变化等方面来鉴别。

故障发生之后，如若怀疑是腐蚀与防护方面的问题所造成，往往要收集、提取有关腐蚀的各种信息和证据，其中腐蚀痕迹及其特征是最关键、最重要的。

各种金属表面的腐蚀特征见表 5.2-6。

(7) 污染痕迹

各种污染物附着在表面而留下的痕迹成为污染痕迹。由于污染物虽与表面接触，但并不发生相互作用，因此，污染痕迹实际就是污染物的自我影像。

表 5.2-6　各种金属表面的腐蚀痕迹特征

金属材料名称	腐蚀产物特征	腐蚀产物的颜色
钢及铸铁	开始是金属表面颜色发暗，腐蚀轻时呈暗灰色，进一步发展会变为褐色或棕黄色，严重腐蚀呈棕色或褐色疤痕甚至锈坑。去除腐蚀产物后底部呈暗灰色，边缘不规则	Fe（OH）$_3$ 黄色，Fe_3O_4 黑色 FeO（OH）棕色，Fe_2O_3 红色 FeS 黑色，$FeCl_3$ 暗褐色 $FeCl_2$ 暗绿色
发蓝（氧化）和磷化的钢件	通常呈黄褐色的锈层，也有的呈点状、斑状	同上。由于发蓝、磷化后钢铁表面呈黑色或灰色，锈蚀产物的颜色有加深
铜合金	铜锈呈绿色，也有呈橘红色或黑色的薄层；铝青铜的锈蚀可呈白色、暗绿色及黑色薄层，严重时呈斑点状或层状突起，除去绿色锈蚀产物后，底部可见麻坑；铅青铜的锈蚀有时呈白色，黄铜有脱锌腐蚀、锈蚀性破裂（季裂）等	CuO 黑色 Cu_2O 橘红色 CuS 黑色 $CuCl_2$ 绿色 Cu（OH）$_2$·$CuCO_3$ 绿色
铝合金	初期呈灰白色斑点，发展后出现灰白色锈蚀产物，刮去锈蚀产物后，底部出现麻坑；硬铝会出现局部腐蚀、剥蚀、晶间腐蚀	Al_2O_3 白色 Al（OH）$_3$ 白色 $AlCl_3$ 白色
锌、镉、锡及其镀层	初期呈灰白色斑点，发展后生成黑色、灰白色点蚀，并有灰白色锈蚀产物，除去锈蚀产物后有坑；锌、镉在有机气氛下，腐蚀产物如白霜，俗称“长白毛”；锌、镉、锡在应力及湿气作用下会产生晶须	ZnO、$ZnCO_3$、Zn（OH）$_2$、ZnS、Cd（OH）$_2$、$CdCO_3$、Sn（OH）$_3$、Sn（OH）$_2$ 均为白色 CdO、CdS 为黄色 SnS 灰褐色
镁合金	初期呈灰白色斑点，发展后在锈蚀处出现灰白色粉末，除去锈蚀产物后底部有黑坑。锈蚀一直沿阳极区伸入，呈深孔交错状	MgO、Mg（OH）$_2$、$MgCO_3$ 均为白色

续表 5.2-6

金属材料名称	腐蚀产物特征	腐蚀产物的颜色
铅及其镀层	一般呈白色或黑色薄层，也有呈红褐色或棕色。镀层腐蚀严重时，会露出基体金属	$Pb(OH)_2$、$PbCO_3$、$PbCl_2$ 均为白色，PbS 黑色，PbO 黄色，PbO_2 棕、褐色，Pb_3O_4 鲜红色
银及其镀层	在空气中易氧化变暗。常见的锈蚀呈暗灰、黑色，也有呈黄色或棕褐色	Ag_2O 褐色，Ag_2S 灰黑色，AgCl 白色
镍及其镀层	初期呈暗灰斑点，发展后锈蚀产物为绿色粉末状疏松物	$Ni(OH)_2$、$NiCl_2$、Ni_2CO_3 均为浅绿色
钴钨合金	锈蚀产物一般呈橘红色	$CoCl_2 \cdot 6H_2O$ 玫瑰红色
铜、银或合金	在 H_2S 气体中	黑色锈色
钢铁、铜表面镀镉	在高温、多雨的霉雨季节	在气孔处出现黑色斑点

污染痕迹分析的内容很多、范围很广，除常规的各种理化检验方法外，主要还有气氛分析、油液分析、腐蚀产物分析和磨粒分析。

另外，机械表面存在的各种异常油迹、烟迹、灰迹等，往往是机械失效的重要证据，在失效分析时要充分加以应用。

污染痕迹分析的主要内容有成分分析、油液中磨损产物的含量与增长速度分析以及磨损产物的粒度及形状分析等。

(8) 分离物痕迹

分离物主要是指接触面在物理、化学作用下从接触面上脱落下来的颗粒。它可以是机械表面的分离物，也可以是反映产物的脱落物。在机械系统中，最常见的分离物是磨屑。此外还有腐蚀产物、毛刺、剥落的涂层、镀层、烧熔溅痕等。

分离物痕迹分析主要是指分离物本身的形貌、成分、结构、颜色和磁性等方面的分析。目前颗粒鉴定已发展成为一门专项技术，其中铁谱分析技术及其应用已比较成熟。

(9) 热损伤痕迹

在热能的作用下，接触部位发生局部不均匀的温度变化而在表层留下的痕迹称为热损伤痕迹。金属表面层局部过热、过烧、熔化、烧穿，漆层及非金属表面烧焦都会留下热损伤痕迹。

热损伤痕迹一般可从以下几方面进行分析。

① 颜色的变化，一般金属在不同的温度区间，具有不同的氧化膜颜色。材料不同，温度区间不同，氧化膜的颜色也不同。

表 5.2-7 列出了不锈钢、钛合金、金属表面硝基清漆层不同温度区间的颜色变化情况。

表 5.2-7　金属材料在不同温度区间的颜色变化

材料		温度区间与颜色						
不锈钢	温度/℃	430～480 开始变色，随温度上升						
	颜色	从黄褐色变为淡蓝色、蓝色、黑色						
钛合金	温度/℃	200	300	400	500	600	700～800	900
	颜色	银白色	淡黄色	金黄色	蓝色	紫色	红灰色	灰色
金属表面硝基清漆层	温度/℃	<200	250～300	350～425	>600			
	颜色	无明显变化	黄色	黑色	起皱			

② 表面层成分、结构的变化，包括氧化膜的形成、合金元素的扩散、富集和贫化。

③ 金相组织的变化，如表面脱碳、晶界贫化、晶界熔化、再结晶、γ′相长大、共晶相熔化等。

④ 表面性能的改变，如显微硬度、耐磨性和耐腐蚀性等。

⑤ 形貌特征，如变形、龟裂、碳化等。

图 5.2-48 是涡轮喷气发动机涡轮叶片由于热损伤导致变形、裂纹的形貌特征。

热损伤一般可分为热冲击、热磨损和低熔点金属热污染三种类型。

1) 热冲击　物体在急剧地加热或冷却过程（温变速度大）中温度急剧地变化（温差大），由此产生冲击热应力，造成破坏的现象叫做热冲击。

热冲击损伤具有以下一些痕迹特征：

① 表面可能烧熔，出现铸态熔坑、几何花样、交叉滑移等；

② 表面烧蚀变色，失去金属光泽；

③ 表面龟裂，萌生热疲劳裂纹，并且出现多条热疲劳裂纹。

表面观察热疲劳裂纹多呈分叉的龟裂状，裂纹内充满氧化物。宏观上断口呈深灰色，并为氧化物覆盖，裂纹源系多源，从表面向内部发展。

磨片观察，裂缝内充满氧化物，其边沿则因高温氧化使基体元素贫化、硬度降低。裂纹多为沿晶型或沿晶 + 穿晶型。

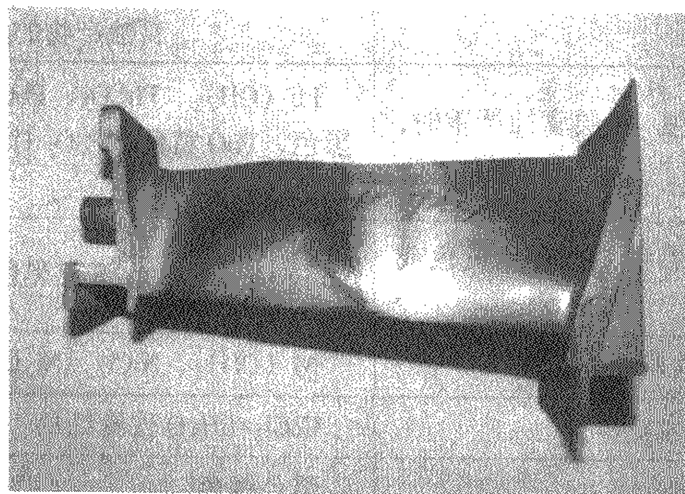
2) 热磨损　当固态物体相互滑过时，甚至在中等的载荷和速度条件下都可能产生非常高的表面温度。这些高温局限在离摩擦界面很近的物体的极薄的表层内，而基体的温升是比较小的。

当滑动摩擦的摩擦表面相对移动速度较高（$v = 3$ m/s），且单位压力也较大（$p \approx 2.5$ MPa）时，使金属摩擦表面层的温度急剧增高，引起热磨损。

当产生热磨损时，互相作用的表面，在表面接触区域之间产生金属的亲合力。发生热磨损时，摩擦表面由裂纹、金属的黏着粒子和涂抹粒子所覆盖。

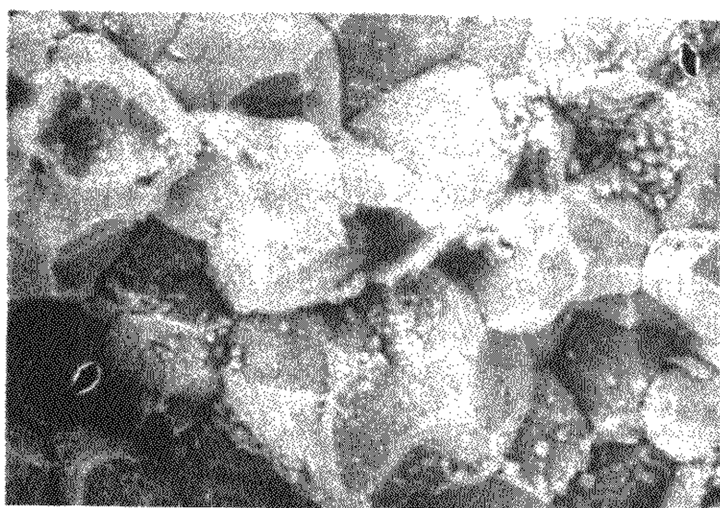
3) 低熔点金属的热污染　低熔点金属受热液化时若与固体金属偶表面直接接触，常使该固体金属浸湿而脆化，在拉伸应力的作用下，从表面起裂，而裂纹尖端吸附低熔点液态金属原子，进一步降低固体金属的晶体结合键强度，导致裂纹脆性扩展。

低熔点金属的热污染只要环境温度接近低熔点金属的熔化温度就可发生，常导致接触金属的脆性断裂，一般称作液体金属脆。这种脆化裂纹一般是分支裂纹或与主裂纹相连接的网状裂纹，裂纹源区为低熔点金属所覆盖，带有不同的色彩，常可检测出低熔点金属元素。高强钢的镉脆断口上就具有这样的特点：沿晶，源区覆盖有镉，并可检测到镉元素的存在，见图 5.2-49。


(a) 涡轮导向叶片热损伤变形、裂纹痕迹


(b) 涡轮转子叶片热损伤变形痕迹

图 5.2-48　叶片热损伤痕迹


图 5.2-49　镉脆断口微观形貌

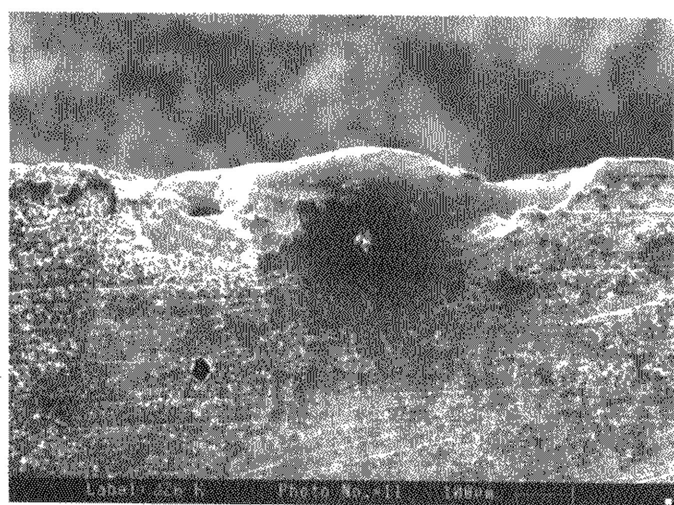
(10) 电损伤痕迹

电接触损伤痕迹包括两大类：电侵蚀（电腐蚀）和静电损伤。下面分别讨论。

1) 电侵蚀　电侵蚀的形式主要取决于电流的大小，典型的电侵蚀曲线可分为三个区：第Ⅰ区为正转移区，此时电流小于生弧极限，又称没有电弧时的阳极侵蚀；第Ⅱ区为负转移区，也称等离子区，此时电流大于生弧极限，又称存在电弧时的阴极侵蚀；第Ⅲ区不仅阳极产生迅速损耗，阴极也有较大损耗，此时电流很大，电弧产生的高温使接触材料熔化、气化或喷溅。

接触元件在断开－闭合过程的不同阶段，电侵蚀的主要因素是不同的。没有电弧时，断开的开始阶段是形成液桥；轻微拉弧时，触点表面变色，由黄灰变黑，并有积碳，有时会产生毛刺；严重拉弧时，触点烧熔，形成熔瘤熔坑，触点附近表面可出现很多微小金属颗粒。还有一种特殊的损坏形式：熔焊，就是触点在工作中突然不再断开，这是由于触点在闭合或断开时受到较大电流（或电弧）的冲击，使接触部位的金属局部熔化，致使触点黏合在一起。

电侵蚀的主要痕迹有电蚀坑、金属熔球、金属转移等，见图 5.2-50。


(a) 触点上的电侵蚀痕迹

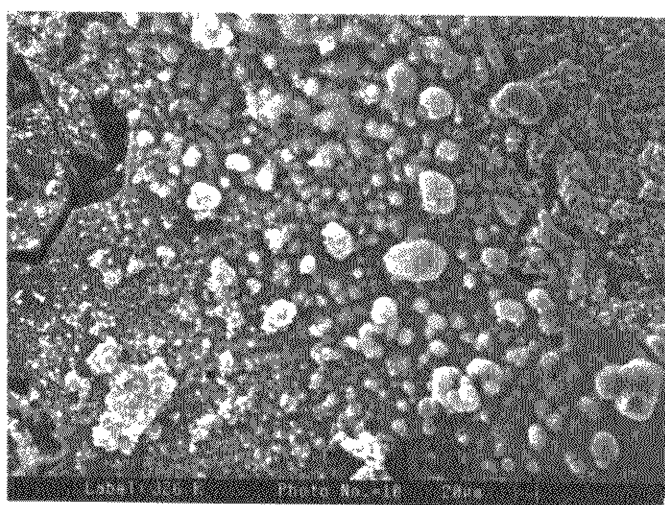

(b) 胶木盖上溅射的银颗粒

图 5.2-50　开关内部电侵蚀痕迹

电侵蚀的主要危害是使电路工作不稳定、电路元件烧毁、引起火灾。如，一飞机的左操作台起火，原因是开关内部电弧引起。

2) 静电放电痕迹　由于静电放电现象而在放电部位留下的电侵蚀痕迹称为静电放电痕迹。

静电放电痕迹主要的宏观特征是：放电过程中形成的碳及碳化物，使放电部位的表面颜色发黄、发灰或发黑。局部的高温熔融，使放电部位表面颜色变成深蓝。在放电过程中，放电体上会形成形貌类似于火山口状的高温熔融微坑，称为“火花放电微坑”，见图 5.2-51。它是静电燃爆事故残骸件上最主要的微观形貌特征。

静电火灾形成的根本原因是放电通道上所释放的静电可以成为可燃物引燃或引爆的点火源。在液化石油气燃爆事故分析中，残骸分析发现灌油枪的局部表面存在大量的“火花放电微坑”，这就证实了该部位曾发生过强烈的静电放电，燃爆起火的原因就是由于静电放电火花引起，火源就在灌油枪。

在微观分析中，需要准确区分“火花放电微坑”和“电气短路微坑”，两者的形成条件是不同的。前者是在高电压、小电流情况下发生的静电火花放电现象，而后者是在低电压、大电流情况下发生的放电形式，是大多电路开关中常见的现象。两者的主要区别如下：

①“电气短路微坑”的形状不规则，面积较大，有时用肉眼或高倍放大镜就可观察到；

②“电气短路微坑”不具有“火山口”形貌特征，而是具有明显的“贝壳”几何花样、“溅射”花样，往往存在明

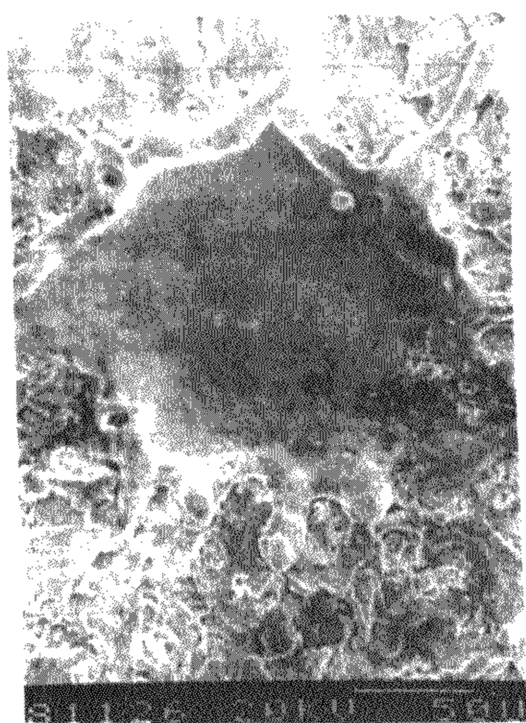


图 5.2-51 静电放电痕迹

显的金属粘连痕迹特征和大量的金属迁移现象。

(11) 与裂纹源有关的痕迹

与裂纹源有关的痕迹很多，概括起来主要有以下几种。

1) 冶金夹渣 某直升机旋翼大梁（材料 6A02-T6），使用中充气检查时已无气压，发现后段已撕裂，在后缘内表面上有一个长型凹坑（宽 1 mm，长 2 mm，深 2.6 mm），左侧内壁和坑底均有一层夹渣，它们与基体金属有明显界限，夹渣区内组织疏松。在此缺陷基础上，形成一条总长达 187 mm 的疲劳裂纹，险些造成严重事故。

2) 挤压裂口 某直升机一号旋翼大梁由于在其内腔下壁第三凸棱处有 1.5 mm 深的挤压裂口，造成该梁仅工作 355 h 28 min即疲劳破坏，导致直升机空中解体，机毁人亡。

挤压裂口表面有锌黄底漆痕迹和阳极化痕迹。微区成分分析确认为阳极化表面成分。

3) 加工痕迹 黑鹰直升机尾减输出轴在飞行中发生疲劳断裂，导致该机坠毁。裂纹起始于输出轴内壁表面的粗刀痕，在刀痕的根部，存在原始的沿晶微裂纹，从而导致使用中发生高周疲劳破坏。

一发动机叶片模锻时，在叶背形成了一折叠，使用中从折叠处萌生疲劳裂纹，导致叶片疲劳断裂，见图 5.2-52。

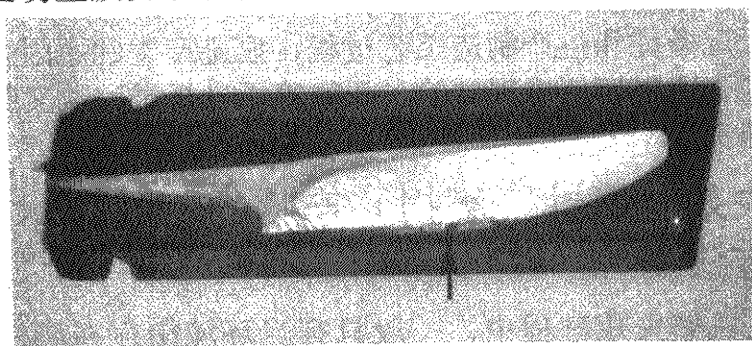


图 5.2-52 叶片裂纹起源于锻造折叠

（箭头所指处为折叠位置）

4) 电笔烙印 某活塞发动机的减速器主动齿轮（材料 12CrNi3A）试车时断裂，把疲劳源区放大观察，靠外圆表面有一白色块状物，用 X 射线能谱仪分析，发现白色块主要成分是 Al、Si、Cu、Zn 等元素。白色块状物呈铸态结构，它与钢基体的过渡组织粗大，具有高温回火状态；白块区 78HV，钢基体 420HV，两者相差 5 倍多。模拟试验证明，这是电笔烙印时火花放电造成的。

5) 钎焊沿晶氧化微裂纹 某发动机加力总管（材料 1Cr18Ni9Ti），用 HL－4 钎料（铜基，含 Ni、Si、B、Fe 等）手工焊。使用中多次发生总管疲劳失效而漏油。对裂纹源作扫描电镜观察，发现裂纹起始于外表面的沿晶分离区，晶粒比较粗大，晶间略有加宽现象，在源区管壁表层 2～3 个晶粒的深度范围内有沿晶裂纹，裂纹内有灰色氧化物。这是焊接工艺不当引起的。

6) 焊接冷裂纹 某歼击机主起落架支柱（材料 30CrMnSiNi2A）在使用中曾发生三次折断事故，均断于斜撑杆螺栓衬套焊缝处。疲劳源是焊缝底部的黑三角区，为焊接裂纹（定位点焊时造成的冷裂纹），见图 5.2-53。

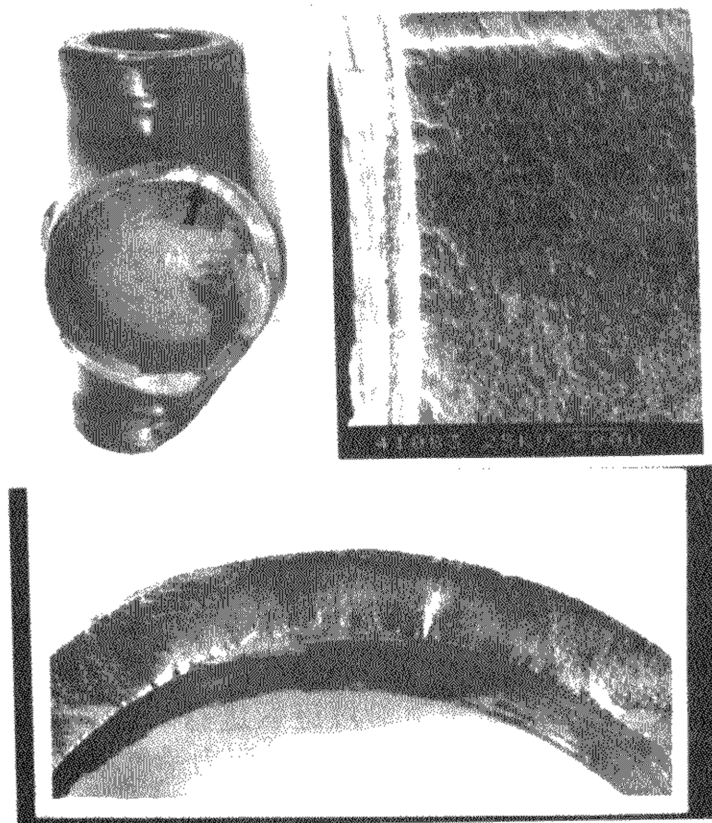


图 5.2-53 裂纹起源于焊接冷裂纹

3.4.2 痕迹的成分诊断

痕迹的成分诊断即是对留痕物上的表面附着物、金属黏结物的元素种类和含量进行分析，以确定造痕物的种类。常用的分析仪器为通过各种发射谱进行表面成分分析的各种表面分析谱仪。目前，能用来进行表面成分分析的基本激发源有：电子、离子、光子、中子、热、电场、磁场和声波 8 种，可供检测的粒子信息有电子、离子、中子和光子 4 种。常用的各种表面分析技术、仪器及其特点、适用范围、技术性能比较列于表 5.2-8。

另外，中子活化分析技术作为一种痕量分析方法在痕迹分析中得到了广泛的应用。该方法利用一定能量和流强的中子轰击样品表面，然后测定由核反应生成的放射性核素在衰变过程中放出的射线，从而对元素进行定量分析。中子活化分析具有其他分析方法所没有的突出优点：①灵敏度高，可达 10^{-6}～10^{-13} g 之间，分析样品的需求量极少；②准确度高，可用来校验其他分析方法所得结果；③样品不需要预处理，可直接检测；④可测碳以后的 77 种元素，范围广泛；⑤可同时检测三四十种元素。目前，中子活化技术可用来鉴定毛发、油漆、油脂、油墨、泥土、岩石、纤维、塑料、玻璃、弹片、弹头擦痕和电损伤等。

3.4.3 痕迹的组织结构诊断

痕迹的组织结构诊断一般采用表面微区晶体结构分析技术。所谓晶体结构分析，就是指对晶体中原子在点阵中的排列方式、点阵类型和结构（包括晶胞大小、晶胞中原子数目、原子位置等）以及点阵应变和点阵缺陷等进行分析。其基本原理是利用高能入射束（如 X 射线、中子和电子等）和晶体点阵（或称晶格）相互作用所伴随着发生的物理效应进行分析。常用的是利用电子衍射效应、电子通道效应和电子的背向散射效应进行晶体结构分析。目前，适用于厚块试样表面结构分析的有低能电子衍射技术（LEED）、反射式高能电子衍射技术（RHEED）、反射电子衍射技术（RED）、电子通道花样技术（ECP）、电子背散射花样技术（EBSP）、X 射线柯塞尔花样技术（XKP）等。

表 5.2-8　表面分析技术性能比较

项目	EPMA	AES	EELS	SIMS	ISS	ESCA	LRM	RBS	PIXE
激发源	电子	电子	电子	离子	离子	光子（X射线）	光子（激光）	离子	质子
分析信息	X射线	俄歇电子	损失电子	二次离子	散射离子	光电子	拉曼散射光子	散射离子	X射线
特长	重元素分析	轻元素分析	极微区（0.5~1.0 nm）中元素分析	全元素分析 同位素分析	最外原子层元素分析	化学状态分析	分子结构分析	轻基体中的重元素分析	表面原子密度和成分分析
分析元素范围	Na~U（EDS）B~U（EDS）	Li~U	Li~U	H~U	Li~U	Be~U	—	—	Al~U
选区尺寸	1 μm	0.1~1 μm	0.5~2 nm	1 μm	100 μm	1~3 mm	1 μm	1 000 μm	几 μm
深度分辨率	1 μm	1 nm	0.6~2 nm	1 nm	1个原子层	0.5~10 nm	1 μm	5 μm	—
检测极限	750×10^{-6}（EDS）10^{-4}（EDS）	0.1%	0.3%~0.5%	10^{-6}	10^{-6}~0.1%	—	1%	—	10^{-6}~10^{-5}
定量分析方法	ZAF	灵敏度因子	灵敏度因子	灵敏度因子	工作曲线法	—	正在发展	独立定量	ZAF
相对误差	1%~5%	10%	20%	20%	20%	—	—	—	10%~30%

3.4.4　痕迹区的性能诊断

痕迹区的性能诊断是指对痕迹区进行的各种物理性能、化学和力学性能的检测分析。包括表面力学性能检测，表面电阻、磁性等物理性能检测和耐磨性、耐蚀性等性能检测。用于这些性能检测的仪器主要有：表面应力测定仪、显微硬度计、腐蚀电位仪以及其他检测表面各种性能的测试仪器。在痕迹诊断中，可酌情选择采用。

3.5　痕迹的模拟再现

通过对痕迹的形貌、成分、组织结构等方面的分析，可以基本确定痕迹的性质、形成的条件。为了验证分析结论的准确性，一般需要进行相应的模拟试验，以使痕迹得以再现。

通常的痕迹模拟再现试验有实验室实验和现场实验两种。实验室实验可以在专门的实验机上进行，试验件可以根据试验机的要求选用同类零件或试片，但实验参数应该调整到与痕迹分析中得出的痕迹条件一致，确定无误后即可开始试验。现场试验应该是比实验室试验更准确、更可靠的模拟再现方法，其工作条件一般来说与失效时的情况更加接近，但相应的一些试验参数也应该根据前面的分析进行适当的调整。

对由模拟试验得出的痕迹，应与失效件上的痕迹进行对比分析，并判断两者是否一致。如果模拟试验得出的痕迹特征与失效痕迹特征基本一致，则基本再现了痕迹，说明对痕迹形成的条件和原因分析基本正确。如果两者不一致，则应分析导致不一致的原因，并对失效痕迹进行补充分析与完善，在此基础上调整试验参数，再次进行模拟试验，直到痕迹被再现。

对重大事故（故障），模拟再现试验一般均是必须的，尤其当分析的证据不是很充分、结论不是很可靠的时候，更需进行模拟再现试验，一方面可以验证分析结论，另一方面也可为分析开拓思路，提供依据。

3.6　痕迹的综合分析

痕迹的综合分析就是在对痕迹的形貌、成分、组织结构和性能等进行鉴定的基础上，综合考虑痕迹的模拟再现试验情况，分析确定痕迹的性质和产生条件（原因），为整个失效分析提供依据的过程。痕迹综合分析的目的就是要确定痕迹的性质和来源（产生的条件或原因）。

3.6.1　痕迹的性质

在本章的“3.4痕迹的鉴定”中，已对不同性质的痕迹的形貌特征、化学成分、组织结构和性能进行了详细的分析与讨论。由于不同性质的痕迹具有不同的特征，同时，不同的特征反应不同的痕迹形状，即痕迹性质与痕迹特征之间具有对应关系，因此，通过对痕迹进行全面的分析，即可确定痕迹的各项特征，由此基本分析出痕迹的性质。

在分析中特别需要注意的是，切不可只根据一两个特征，或只根据宏观特征或微观特征，就确定痕迹的性质，而是需要综合考虑痕迹的各方面特征后才能做出最后的结论。

3.6.2　痕迹的来源（产生）

痕迹的来源，即痕迹产生的原因或条件，一般是失效痕迹分析的重点。痕迹原因分析的基础是痕迹的性质分析，不同性质的痕迹具有不同的产生条件，也就具有不同的产生原因，即痕迹的性质与原因之间具有对应关系。

在基本确定了痕迹的性质后，综合考虑零件的工作原理、工况条件和生产制造工艺过程，痕迹产生的原因也就基本清楚，通过模拟再现试验，既完善了分析过程，又验证了分析结论。

4　失效（腐蚀、磨损）产物的诊断技术和方法

在机械失效过程中，一般只有当环境中存在腐蚀性介质、发生与腐蚀有关的失效和失效件之间相互接触、运动，发生与磨损有关的失效时，才会产生失效产物。单纯由于应力作用而导致的失效，如断裂、变形等，是不会产生失效产物的。

由于与腐蚀有关的失效是发生在有腐蚀性介质的环境中的失效，因此，在失效部位将有腐蚀产物出现；与磨损有关的失效必须在当零部件之间相互接触、并产生相对运动时才会发生，失效部位将会产生磨损，在接触界面出现磨屑。常见的与腐蚀有关的失效模式有：电化学腐蚀、化学腐蚀、氧化、应力腐蚀、腐蚀疲劳、蠕变、氢脆、低熔点金属脆性、腐蚀磨损等。常见的与磨损有关的失效模式有：微动磨损、磨粒磨损、腐蚀磨损、疲劳磨损（接触疲劳）、黏着磨损等。

失效产物与失效件的材料、环境介质、应力等密切相关，可从一个方面反映材料、失效件的失效模式、原因和机

理。因为材料不同，失效（腐蚀、磨损）产物的成分不同；不同的失效模式，可以产生不同的失效（腐蚀、磨损）产物；失效原因和机理不同，产生的失效（腐蚀、磨损）产物的特征也不同。形貌、成分、结构是失效（腐蚀、磨损）产物的主要特征量，因此也是诊断失效模式、原因和机理的主要的特征参量。在失效诊断时，应从这三个方面入手，对失效（腐蚀、磨损）产物进行分析，为整个失效诊断提供可靠的依据。

需要注意的是，失效产物只存在于失效部位或断口、裂纹的表面，有的黏附牢固、而有的与表面结合力不强，因此，在对失效件进行清理、运输、分析的过程中，应该防止对其造成破坏、假象或遗失，从而给分析造成不必要的困难。当不能确定表面的附着物是否为失效产物时，应先对其进行收集，并保管好，以便进一步分析。

4.1　失效产物的形貌诊断

失效产物主要包括附着在失效件表面的各种腐蚀产物、磨屑等。

各种金属表面的腐蚀特征见表 5.2-6。可利用各种光学和电子显微镜观察其形貌特征。

磨屑专用的动态形貌观察手段为铁谱仪。也可将润滑油系统内的磨屑过滤出来，借助各种光学和电子显微镜观察其形貌特征，并由此确定磨损状态。

4.2　失效产物的成分诊断

失效产物的成分分析诊断包括常规的、局部的、表面的和微区的四种分析。

当失效产物多时，可将这些产物单独收集起来进行分析，可采用常规的分析方法，如湿法化学分析、点滴试验分析、燃烧法分析及各种光谱分析。

当失效产物以薄层的形式分布于基体表面时，可用俄歇电子能谱仪分析五个原子以内厚度这样极薄层的元素成分；莫斯波尔光谱仪可用来识别铁合金中氧化物、硫化物、氮化物和碳化物，特别是分析稀土氧化物的鉴别率高，它分析的深度，用背散射电子测量的深度为 300 nm，用放射伽玛线测量的深度为 12.7 μm。表面微区域（0.5 μm 以上）的成分分析还有电子探针、离子探针。但离子探针虽然可作所有元素的微区分析，分析含量可低达 100×10^{-6}，但要破坏原始表面，因此分析结果不能重复，但可作逐层分析，是作氢分布图像的惟一工具。X 射线能谱仪由于可与扫描电镜同时应用，在表面微区域元素成分分析中得到了广泛的应用，但分析的误差相对较大。

对存在于润滑油系统内的磨损产物分析，一般运用铁谱分析技术，由铁谱仪进行分析。

4.3　失效产物的结构诊断

失效产物的结构诊断一般采用 X 射线结构分析技术和电子衍射分析技术。X 射线结构分析设备为 X 射线衍射仪，通过对晶体中原子排列造成的 X 射线衍射图像进行适当的变换，得出晶体结构的数据。电子衍射分析同 X 射线衍射分析一样，也是产生衍射图像，不过是通过电子束射到表面，产生的是电子衍射图像，通过演算处理，获得物质的晶体结构。电子衍射分析可对金属材料中微小区域内的特定组织、组织内部的特定相、细小的夹杂物或析出相进行选区电子衍射物相分析，判定不同显微组织的类型、组织中各相的晶体结构、它们与基体之间的位相关系，或者对金属材料中的夹杂物、析出相进行结构分析，确定其物质属性，掌握它们的形态特征及其取向关系。

5　综合诊断技术和方法

由于客观世界的复杂性，失效往往不是单一因素造成的结果，而是多种因素综合作用导致的结果。以机械构件的失效为例，失效是在应力作用下导致的，表现形式为制造构件的材料的失效，同时又有环境的加速或帮助。而失效的机理则更加复杂，从材料组织上可分为穿晶、沿晶；从应力上可分为恒载、交变、快速作用等；从环境上又可分为化学腐蚀、电化学腐蚀等。因此，失效的模式、机理、原因等各方面均极其复杂，不是从单一的一个方面就可以轻易判断的，需要从设计、使用、制造、维护、修理、原材料等方面进行综合诊断。同时在综合诊断时也要遵循从失效模式到失效原因、失效机理的诊断程序。

5.1　失效诊断的思路和程序

失效总是有一个或长或短的变化发展过程，在实质上总是表现为材料的累积损伤过程，即材料发生物理的或化学变化的过程，而且整个发展过程都是有条件的和有规律的，也就是说，失效的过程是有其内在的规律存在的，因此，在失效诊断中，应以反应失效规律与过程的思路，按照一定的程序来进行。

失效诊断思路是指在失效诊断过程中，以失效的客观规律为理论依据，把通过调查、观察和实验所获得的失效信息（失效对象、失效现象、失效环境）分别加以考虑，然后有机地结合起来作为一个整体综合分析，以获取的客观事实为证据，全面应用逻辑推理地方法，来判断失效事件的失效模式，并推断失效原因和失效机理。失效诊断思路是指导整个失效分析全过程的思维路线。

在失效分析中，我们首先面对的是失效过程的结果，如断裂、腐蚀、磨损等；但一般情况下，这些又不是我们掌握的惟一线索，如还知道系统的基本工作原理、受力状态、材料性质、环境条件等。在整个分析过程中，要充分掌握好、利用好这些信息，并以此为基础开展分析。而不能把失效分析简单地归结为从果求因逆向认识失效本质的过程。如果把失效结果比作是瓜，失效过程比作是藤，失效系统的起始状态和原因比作是根，则失效分析的思路可以概括为以下几种。

1）顺瓜摸藤　从失效的结果出发，不断由过程的结果推断其原因。

2）顺藤找根　从失效过程的中间状态出发，推断该过程形成的原因，直至过程整个失效的直接原因。

3）顺根摸藤　从系统的起始状态出发，不断由过程的原因推断其结果。

4）顺藤摸瓜　从失效过程中间状态的现象为原因，推断过程进一步发展的结果，直至失效的结果。

5）顺瓜摸藤＋顺藤找根。

6）顺根摸藤＋顺藤摸瓜。

7）顺藤摸瓜＋顺藤找根。

在失效分析的发展过程中，还有一些经典的分析思路，曾经发挥过重要的作用，有的还在广泛应用。如“撒大网”逐个排除思路。该思路认为任何失效都可以从人（操作人员）、机（系统）、料（材料）、法（制造工艺）、环（环境）和管理六个方面去寻找，可以将失效的所有原因列出来，然后逐个排除。显然，该方法的工作量大，看似十分全面、稳妥、可靠，实际难以应用，不到万不得已时，一般都不宜采用。还有“故障树”分析法。该方法来源于电子设备的可靠性工程研究，以失效结果为第一层，然后逐层列出导致上一层事件发生的可能原因。由于该方法关心的是失效的部位，而不追求失效的机理和过程，因此，不是一种完整的失效诊

断思路。

逻辑推理是从已知的知识推出未知的知识，从一个或几个已知的判断推出另一个新的判断的思维过程。只要据以推出新判断的前提是真实的，推理前提和结论之间的关系符合思维规律，则得出的结论或判断就一定是真实的、可靠的。显然，逻辑推理的思路可以作为失效分析的基本思路，以真实的失效信息事实为前提，根据已知的失效规律性知识和已知的判断，通过严密的、完整的逻辑分析，就可推理出失效的模式、原因和机理。

常用的逻辑推理方法有归纳推理、演绎推理、类比推理、选择性推理、假设性推理等五种，在失效分析中应灵活运用。

为了得到正确的失效分析结论，提高效率，少走弯路，光有好的分析思路还不够，还需要正常的分析程序，特别是在复杂的失效事件中，这一点尤其重要。在一般的失效分析中，都需遵循以下程序。

1）调查现场失效信息　这是整个失效分析工作的基础和前提。调查一般以失效现场为出发点，通过观察和现场实验等手段，全面、系统和客观地收集失效对象、失效现象和失效环境等失效信息。调查时应作好必要的记录和照相，必要时，还需反复调查。

2）初步确定肇事失效件　根据失效系统的结构特点、工作原理和相关的痕迹特征、失效件的失效特征，运用逻辑推理的思维方法，确定肇事失效件。

3）确定具体分析思路和工作程序　如果历史上曾经发生过类似的失效事件，应按类比推理的思路和程序进行分析；对首次发生的失效事件，应按逻辑推理的思路和程序进行分析。

4）初步判断肇事件的失效模式　根据失效的宏观、微观特征、痕迹特征、结构特点、材料特性、环境条件、工作原理和受力状态等信息，分析确定失效的模式。

5）查找失效原因　根据失效模式所指明的方向，围绕失效模式所涉及的原因，从内因和外因两方面查找。

6）模拟再现　根据找到的可能原因，在同样的系统和工况条件下进行现场模拟实验，应能将失效事件再现。同类系统的普查也可看作是失效事件的模拟再现。

这方面的工作可根据需要选择。

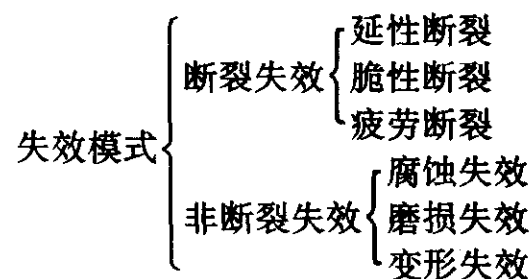
7）综合分析　在上述工作的基础上，对整个失效事件进行综合性的系统分析，从而得出失效分析的结论。

8）总结报告　这是失效分析的最后一个程序。总结时要对整个失效过程进行回顾，从总体上审视失效分析全过程，发现弥补不足，回答失效分析所赋予的使命，并提出预防再失效的建议，最后形成失效分析报告。

5.2　失效模式（性质）诊断技术和方法

失效模式是指失效的外在宏观表现形式和过程规律，一般可理解为失效的类型。失效模式诊断是失效诊断的基础和前提，因此，对失效模式的诊断应该准确、深入，才能对下一步的原因和机理的诊断真正起到指导作用。

失效模式常可分为一级失效模式、二级失效模式、三级失效模式。其中一级失效模式一般可分类列表如下：



一般一级失效模式根据不同的标准或参量，又有不同的分类。

失效模式诊断一般要求诊断到“二级”甚至“三级”。但由于越往下诊断，不同失效模式间的差别越小，使得诊断的难度越大；但失效模式诊断得越具体、越准确，对失效原因的诊断和预防措施的针对性却是十分重要。因此，失效模式诊断要求诊断得越具体、越准确越好，对失效原因诊断的准确性和预防措施制定的针对性越有指导价值。

失效模式诊断一般从失效前的工作参数（工况）、失效后的宏观特征、失效件的断口特征、失效件的表面痕迹、失效产物的形貌与成份、环境因素、制造失效件的材料、模拟试验等方面来进行分别和综合诊断。

失效前的工作参数（工况）反应了失效件的载荷情况，如果是正常工作情况下的失效，由工作参数即可基本确定可能的失效模式。但由于一般失效件的工作情况与设计状态存在一定的差异，因此，在了解设计结构特点、受力水平的基础上，还应该重点掌握工作状态与设计、试验状态之间的差异，相似设备、构件的使用情况及故障情况等。

根据失效后零件有无明显的宏观塑性变形，以及变形的方向、角度等特征，即可初步判断失效件的失效模式是延性断裂失效还是脆性断裂失效。

断口特征是诊断断裂失效件失效模式的重要依据，也是研究的重点。对一般失效件，从断口的宏观、微观特征即可诊断出失效模式，并为失效原因诊断指出方向。断口的微观特征信息一般揭示了失效的机理，因此，微观特征研究是诊断失效机理的主要手段。

应该说，零件经受的每一过程均会在零件上留下其特殊的痕迹特征，如热过程会在表面留下氧化变色痕迹，机械接触损伤会留下机械接触痕迹，环境介质作用会留下腐蚀痕迹等，因此，失效件的表面痕迹反映了失效前及失效过程中零件所经受的外界作用和时机、方向、作用物等，不但可为失效模式诊断提供依据，而且可为失效原因诊断提供证据。这是在失效综合诊断中必须充分利用的。

根据腐蚀产物的成分，可确定造成零件环境失效的介质成分，并以此进一步确定失效的模式、找到失效的原因。磨损的模式不同，磨屑的形貌也不同，由磨屑的形貌可诊断出失效的模式。因此，对腐蚀和磨损等失效模式，由失效产物的形貌与成份可为失效模式诊断提供参考，为失效原因诊断提供依据，因此，具有极为重要的作用。

各种失效，归根结底都是材料的失效，因此，在失效诊断时，应该充分考虑材料的特性、失效的特征和规律，只有了解了材料，才能做到诊断的准确无误。

对复杂的、难以确定的失效，通过模拟诊断，可确定失效件；对已确定的失效件，其失效模式、原因和机理目前还没有认识，或与现有认识有差异时，模拟诊断是必须的。通过已知条件下的模拟试验，对相关特征的对比分析，可为诊断结论提供重要的依据。

在以上分析研究工作的基础上，进行综合诊断是必须的，也是所有失效诊断中都进行的一项重要工作。当所有方面的分析结论反映的是同一问题的不同方面时，我们的诊断结论就会显得更加充分；当一些分析结论不一致、有矛盾时，就应该特别注意，往往这一不一致就是失效的原因所在，甚至隐含有不为我们所认识的新的机理存在。如断口反映的是实际工作条件的情况，由断口特征反映出的失效模式与设计工况不一致时，说明设计考虑不周，没有包括全部的实际工作情况，这可能就是失效的原因或原因之一。由18世纪火车轮轴的断裂就引出了材料的疲劳失效问题。

5.2.1　断裂失效模式诊断技术和方法

断裂失效模式是在实际中经常遇到的一类失效模式，特别是机电装备的灾难性事故，往往都与断裂失效密切相关。断裂是指一个整体的分开或分离，因此，断裂总有一个或长

或短的过程。断裂失效模式诊断是从断口的宏观、微观形貌特征人手，重点分析断裂过程的形貌特征与应力状态、环境之间的关系，从而诊断出断裂失效的表现形式和过程规律。

断裂失效模式根据需要或其表现特征可以有很多的分类方法，其中主要的分类方法有以下几种。

1）按断裂性质分

① 延性断裂　单调加载，断裂前有明显的塑性变形（宏观塑性变形），一般在5%以上。由于延性断裂在断裂之前均有一定的变形，容易被发现，一般不会造成较大的危害。

② 脆性断裂　单调加载，断裂前无明显的塑性变形（宏观塑性变形很小）。一般塑性变形量小于5%的断裂均可称为脆性断裂。脆性断裂由于断裂前一般没有明显的征兆，往往容易引起突发性事故。

③ 疲劳断裂　在交变载荷反复作用下引起的断裂。

④ 环境断裂　在侵蚀性环境条件下，单调加载的慢速断裂。属于这种断裂的有应力腐蚀开裂、氢脆断裂、液态金属致脆断裂、固态金属脆化和辐射损伤等。

2）按断裂路径分类

① 穿晶断裂　大多数金属材料在常温下的断裂均为穿晶断裂。包括解理断裂、滑移和延伸断裂、韧窝断裂等。

② 沿晶断裂　金属沿着晶界面分离断裂的模式。沿晶断裂大多为脆性断裂，如回火脆性断裂、氢脆断裂、应力腐蚀断裂、淬火开裂、液态金属致脆断裂等。

3）按断裂过程的微观机制分类

① 解理断裂　金属在正应力作用下，由于原子结合键的破坏而沿一定的晶体学平面（即解理面）快速分离的过程。

② 准解理断裂　介于解理断裂与韧窝断裂之间的一种过渡断裂形式。

③ 韧窝断裂　金属在微区范围内塑性变形产生显微空洞，经形核、长大、聚集，最后相互连接而导致的断裂形式。

④ 滑移断裂　金属以位错滑移变形为机理的断裂形式。

从断裂失效诊断的角度来说，断裂的分类应是失效的受力、原因、表现形式等方面的一个综合反映，不宜太深奥、太专业。因此，断裂失效模式的一级分类可分为延性断裂、脆性断裂、疲劳断裂和环境断裂四种一级失效模式。下面将对它们的诊断技术进行介绍。

（1）延性断裂模式诊断

1）宏观变形诊断　延性断裂和脆性断裂是根据断裂件的宏观塑性变形来分类，延性断裂的前提是塑性变形，即延性断裂在断裂位置附近有明显的变形，而脆性断裂的变形不明显，因此失效件的失效模式是否为延性断裂，主要应该从宏观上来判断其有无明显的塑性变形。常见的拉伸断裂试件，不但有明显的伸长变形，而且断裂处还有明显的缩颈。

2）表面颜色诊断　当断裂后没有介质和温度作用的情况下，断口应该是干净的，没有附着物，并具有一定的金属光泽。

3）断口的宏观特征诊断　典型延性断裂件的断口宏观上呈“杯－锥状”断口特征，断口上可分为三个区，即裂纹起始的纤维区、裂纹扩展的放射线区和最后断裂的剪切唇区。纤维区粗糙，一般位于断口心部，当失效件表面存在缺口时，也可能位于表面附近；放射线区放射线的发散方向指向裂纹的扩展方向；纤微区和放射区均基本与应力轴线垂直，而剪切唇区一般与应力轴呈45°斜角。韧性材料的过载断裂均具有此种典型的断裂特征，但构件的形状、材质、受力状态及表面有无缺口等对断口的宏观特征有很大的影响。图5.2-11中示出了板状试样的延性断裂特征，尽管也有上述断口三要素存在，但放射区的形状已有很大的改变，表现为一种“人字纹”花样，“人字纹”的尖端指向断裂源的位置－纤微区。冲击载荷作用下的延性断裂也是常见的一种断裂模式，因为断裂件可对其他机件进行撞击，或失效时，如失事飞机撞击地面时，大多的断裂件均是撞击载荷作用下形成的延性断裂。这类延性断裂的断口也具有三要素特征，但由于零件在受冲击时，受冲击一侧受压，而另一侧受拉，就使得受拉的一侧表现为正断型，受压的一侧裂纹扩展受抑制而出现纤微区，剪切唇区在零件断口的周围不完整，典型冲击断口三个区的分布见示意图5.2-11。

纤微区、放射区、剪切唇区的相对大小受材料塑性、受力状态、构件形状的影响很大，有时甚至只表现出两个区、一个区；而温度、形变速率等因素的影响相对较小。材料的塑性越好，形变速率越小、构件的尺寸越小，一般纤微区、剪切唇区就越大，放射区就越小，有时甚至不出现。对矩形板状零件，受板厚的影响很大，试样的厚度减小时，“人字纹”区缩小，剪切唇区扩大，当板厚小到一定尺寸以下时，其断口就只能见到剪切斜断口，表现为全剪切型的切断断口。

在实际零件中，零件的形状、特别是表面有无缺口对断口三要素的位置影响很大，对一般光滑拉伸断口，纤微区一般在断口中心位置，当表面有缺口、台阶、螺纹时，由于裂纹首先从缺口根部的应力集中处产生，因此，纤微区就不再位于断口中心位置，而是沿外表面分布，放射区放射线收敛的方向指向裂纹扩展方向，见图5.2-9。

4）断口微观特征诊断　在宏观上，延性断裂的前提是塑性变形；在微观上，金属的塑性变形是位错在滑移面上增殖运动和挤出滑移面的结果。在塑性变形过程中，因为位错的塞积或交替作用会造成微裂口，在外力的作用下，微裂口逐渐长大并连接起来，最后导致断裂。这些微裂口的微观形貌即为韧窝，即金属的延性断裂的断口微观形貌均表现为韧窝。其中纤微区属于正断型断裂，为等轴韧窝，而剪切唇区为拉长韧窝，韧窝拉长的方向与匹配断口上韧窝拉长的方向正好相反。在碳化物强化的钢件上，韧窝底部还有碳化物颗粒。

5）参数诊断　根据应力－强度干涉模型，在结构设计中，明确要求材料的屈服强度大于许用应力，因此，从设计上来说，正常使用中是不应该出现延性断裂的，但当工况超过使用限制时，如转动件的转速超过规定转速、飞机飞行科目超过飞行包线时，即工作参数超过使用限制时，则会导致零件的延性断裂。但材料的屈服强度达不到要求时，结构尺寸小于规定尺寸时，相当于增大了工作应力，也可能导致延性断裂。同时，从应力变化上来说，延性断裂的应力应是一次性的，或变化缓慢，且次数很少，而不是其他类型的应力。

最后需要说明的是，即使是脆性断裂的零件，在最后断裂区也同时包含有局部的延性断裂方式，如疲劳断裂的瞬断区。

综上所述，判别金属零件延性断裂模式的判据见表5.2-9。

表5.2-9　金属零件延性断裂模式诊断判据

序号	内容	特　征
1	材料	延性金属材料
2	温度	在材料的韧－脆转变温度以上
3	应力状态	静应力，大于材料的屈服强度
4	宏观特征	断裂位置附近有明显的宏观塑性变形
5	断口宏观形貌	粗糙、色泽灰暗、呈纤维状；边缘有与零件表面呈45°角的剪切唇

续表 5.2-9

序号	内容	特　征
6	断口微观形貌	韧窝花样
7	组织	断口附近表面金相组织有明显的变形层
8	表面状态	断口附近表面脆性的镀层、涂层等表面覆盖膜破裂

由于影响材料的失效模式和断裂特征的因素很多，一般的延性断裂不一定具备以上全部特征，在失效分析中也不一定需要确定以上所有特征才能做出诊断。一般从断口的宏观、微观特征即可以做出判断。

(2) 脆性断裂模式诊断

根据定义，脆性断裂是断口部位无宏观塑性变形痕迹、断口表面相对较平齐的一类断裂失效模式。由于其在断裂前一般没有明显的征兆，因此该类断裂很危险，常可造成很大的损失。

脆性断裂包括以下两种类型：低应力脆断和环境介质作用下的脆断。疲劳断裂也属于脆性断裂的一种，由于这类断裂失效模式在实际生产、生活中发生的较多，约占整个断裂失效的70%～80%，同时其断裂过程又有其本身的特殊性，工程界对其研究较多，因此，常把疲劳断裂模式作为一类单独的失效模式进行研究。环境介质作用下的断裂失效模式近年来越来越普遍，已引起了广泛的关注，将在下文中将其单独作为一种断裂类型来进行研究。因此，这里探讨的脆性断裂模式主要是指低应力脆性断裂失效模式。

低应力脆性断裂通常是指在弹性应力范畴内，在许用应力条件下，一次加载引起的脆性断裂事故。零件受载时的加载速率、环境温度、几何形状、应力集中等外部因素以及材料内部的宏观缺陷，如裂纹、孔洞、大块夹杂、疏松等以及材料本身存在的低温脆性、蓝脆、回火脆、不锈钢中的475℃脆、σ相脆性等因素对零件是否发生脆性断裂的影响很大。

1) 参数（工况）诊断　由于低应力脆性断裂具有突然性，在实际使用中是不允许的，也是不正常的现象，因此其使用条件往往超出了设计限制范畴，故在使用参数上往往可以找到依据。工作在韧－脆转变温度以下的零件在较低的应力下就可导致脆性断裂。载荷作用速度大、应变速率快的条件下也容易出现，如冲击载荷作用时。

2) 宏观变形诊断　脆性断裂与延性断裂的区分是宏观变形量的大小或有无。根据脆性断裂模式的定义可知，脆性断裂的零件，断裂部位在宏观上塑性变形很小或几乎看不到。

3) 断口宏观特征诊断　由于断裂具有突然性，断裂部位没有明显的宏观塑性变形，因此，断口的两部分的匹配性良好；同时断裂没有经过时间过程，因此断口也应该很干净，没有其他损伤和污染（不包括断裂后的二次污染）。

常见的脆性断裂断口宏观形貌有如下几种。

① 结晶状断口　断面平整，在阳光下转动可观察到反光的小平面—穿晶小平面。这是典型的脆性断裂断口。

② 岩石状断口　这也是一种典型的脆性断裂断口形貌，其表面粗糙，颜色灰暗，属于沿晶断裂。

③ 放射状断口　从一处或多处发射的放射状条纹断口，放射条纹的收敛处为裂纹源。在板材断口上，出现的放射状断口，其放射状条纹形成“人字纹”特征，“人字纹”头部指向裂纹源；两侧有缺口的板材，其“人字纹”方向与无缺口板材的相反。

4) 断口微观形貌诊断　严格的脆性断裂，其微观形貌应只有沿晶断裂、穿晶解理和穿晶准解理三类，其中沿晶断裂形貌与宏观上的岩石状断口对应，解理断裂形貌与宏观上的结晶状断口对应。通常见到的脆性断裂的微观形貌主要为准解理断裂形貌。以上三种微观断裂形貌在断口的微观形貌分析中已有详细的叙述，在此不再重述。

5) 材质诊断　晶粒度粗大、夹杂物多、脆性相沿晶界分布以及脆性大的材料容易出现脆性断裂。此外，含有裂缝、孔洞、疏松等缺陷的材料也容易发生脆性断裂。

综上所述，脆性断裂失效模式的判据如表 5.2-10。

表 5.2-10　脆性断裂失效模式判据

序号	内容	特　征
1	材料	晶粒度粗大、夹杂物多、脆性相沿晶界分布以及脆性大的材料，含有裂缝、孔洞、疏松等缺陷的材料均容易出现
2	温度	在材料的韧－脆转变温度以下
3	应力状态	断裂发生时常有动载荷存在，或有冲击载荷作用
4	宏观特征	断裂位置及其附近均无明显（或很小）的宏观塑性变形
5	断口宏观形貌	断口匹配面吻合好，断面粗糙
6	断口微观形貌	通常可见解理、准解理、沿晶等特征
7	起裂部位	在应力集中部位或有表面缺陷、内部缺陷处
8	其他	断裂过程具有突然性

(3) 疲劳断裂模式诊断

1) 疲劳断裂的定义　疲劳断裂是材料（或构件）在交变应力反复作用下发生的断裂。

所谓交变应力是指应力的大小、方向或大小和方向同时都随时间做周期性改变的应力。这种改变可以是规律性的或不完全规律性的。

2) 疲劳断裂的危害性

① 多数机件承受的应力是周期性变动的（称为循环交变应力）。这些零件的失效，据统计60%～80%是属于疲劳断裂失效；

② 疲劳破坏表现为突然断裂，断裂前无显著变形。不用特殊的探伤设备，无法预察损坏迹象。除定期检查外，很难防范偶发性事故；

③ 造成疲劳破坏时，循环交变应力中的最高应力一般远低于静载荷下材料的强度极限；有时也低于屈服强度极限，甚至低于最精密测定的弹性极限；

④ 零件的疲劳断裂不仅取决于材质，而且对零件的形状、尺寸、表面状态、使用条件、外界环境等非常敏感。加工过程也对疲劳抗力有很大影响。材料内部宏观、微观的不均匀性对材料抗疲劳性能的影响也远较静负载下为大；

⑤ 很大一部分机件承受弯曲扭转应力，这种机件的应力分布都是表面应力最大。而表面状况，如切口、刀痕、粗糙度、氧化、腐蚀、脱碳等都对疲劳抗力有极大的影响，增加了疲劳损坏的机会。

3) 疲劳断裂过程　疲劳断裂过程可分为疲劳裂纹的萌生、稳定扩展以及失稳断裂三个阶段（图 5.2-54）。

疲劳裂纹的萌生（形核）：在交变载荷作用下，材料发生局部滑移。随着循环次数的增加，滑移线在某些局部区域内变粗，并形成滑移带，其中一部分滑移带为驻留带。进一步增加循环次数，驻留滑移带上可以形成挤出峰、挤入槽现象，这就是疲劳裂纹的萌生。表面缺陷或材料内部缺陷起着尖缺口的作用，使应力集中，促进疲劳裂纹的形成。实际工程构件的疲劳裂纹大都在零件表面缺陷、晶界或第二相粒子处萌生。

疲劳裂纹稳定扩展的两个阶段。

① 疲劳裂纹稳定扩展第一阶段　疲劳裂纹扩展的第一阶段是在裂纹萌生后，在局部载荷作用下立即沿着滑移带的主滑移面向金属内部伸展。此滑移面的取向大致与正应力成45°角，这时裂纹的扩展主要由于切应力的作用。对于大多数合金来说，第一阶段裂纹扩展的深度很浅，大约在2～5

个晶粒之间。镍基高温合金在高温下的疲劳断裂，这一阶段发展得比较充分，有如下两种类型的断裂形貌特征：a）类解理断裂小平面；b）平行锯齿状断面。

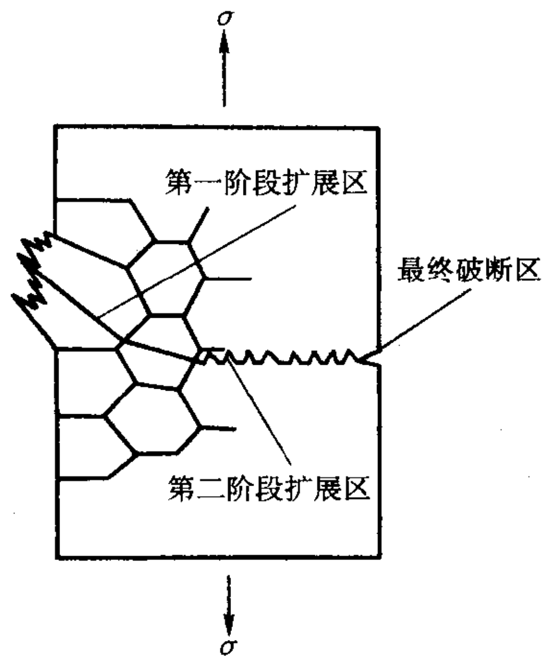


图 5.2-54　疲劳断裂过程示意图

② 疲劳裂纹稳定扩展第二阶段　疲劳裂纹按第一阶段发生扩展一定距离后，将改变方向，沿着与正应力相垂直的方向扩展。此时正应力对裂纹的扩展产生重大影响。这就是疲劳裂纹稳定扩展的第二阶段。疲劳裂纹扩展第二阶段断面上最重要的显微特征是疲劳条带，又称疲劳条纹或辉纹。

每一条条带一般代表一次应力循环，每条条带代表该循环下裂纹前端的位置，疲劳条带在数量上与循环次数基本相等。但在实际断口上，疲劳条带的数量不一定与循环次数完全相等，因为它受应力状态、环境条件、材质等因素的影响很大。同时，疲劳条带有时也容易与其他花样混淆。疲劳条带是疲劳断裂的基本特征，凡在断口上发现了疲劳条带，即可判定此断口为疲劳断裂。但如果在断口上未发现疲劳条带，并不能判断此断口为非疲劳断裂，还要根据宏观、微观的其他特征和其它分析工作的结果加以综合分析判断。一般来说，延性材料容易形成疲劳条带，脆性材料则比较困难。

通常将疲劳条带分成延性疲劳条带与脆性疲劳条带。脆性疲劳的特征是断裂路径呈放射状扇形，疲劳条带被放射台阶分割成短而且平坦的小段。

4）疲劳断裂失效分析　断裂失效分析的内容包括分析判断零件的断裂失效是否属于疲劳断裂，疲劳断裂的类型，引起疲劳断裂的载荷类型与大小，疲劳断裂的起源等。疲劳断裂失效分析的目的在于找出引起疲劳断裂的确切原因，从而为防止同类疲劳断裂失效再次出现所要采取的措施提供依据。

① 疲劳断裂失效的分类　根据零件在服役过程中所受载荷的类型与大小，加载频率的高低及环境条件等的不同，可将疲劳断裂按表 5.2-11 进行“二级”失效模式分类。

由于各类疲劳断裂寿命均是以循环周次计算，一般分为高周与低周疲劳。

a）高周疲劳，又称应力疲劳，是指零件在较低的交变应力作用下产生的疲劳现象。

b）低周疲劳，又称应变疲劳，是指零件在较高的交变应力作用下产生的疲劳现象。

② 疲劳断口的宏观分析　典型的疲劳断口按照断裂过程的先后有三个明显的特征区，即疲劳源区、扩展区和瞬断区（图 5.2-55、图 5.2-56）。

在一般情况下，通过宏观分析即可大致判明该断口是否属于疲劳断裂，断裂源区的位置，裂纹的扩展方向以及载荷的类型与大小。

表 5.2-11　疲劳失效的“二级”失效模式分类

	分类依据	类型	细分
疲劳断裂失效	根据频率不同	高频疲劳断裂	
		低频疲劳断裂	
	根据循环周次不同	高周疲劳断裂	
		低周疲劳断裂	
	根据控制参量不同	应力疲劳断裂（控制应力幅）	
		应变疲劳断裂（控制应变幅）	
	根据环境介质不同	疲劳断裂	
		腐蚀疲劳断裂	
		应力腐蚀疲劳断裂	
	根据温度不同	低温疲劳断裂	
		室温疲劳断裂	
		高温疲劳断裂	
	根据应力来源不同	机械疲劳断裂	拉－拉（压）疲劳断裂
			弯曲疲劳断裂
			扭转疲劳断裂
			接触疲劳断裂
		冷热疲劳断裂	
	根据宏观裂纹走向不同	前切型疲劳裂纹（如疲劳第一阶段）	
		正断型疲劳裂纹（如疲劳第二阶段）	
	根据微观裂纹走向不同	穿晶疲劳断裂	晶格型
			非晶格型
		沿晶疲劳断裂	

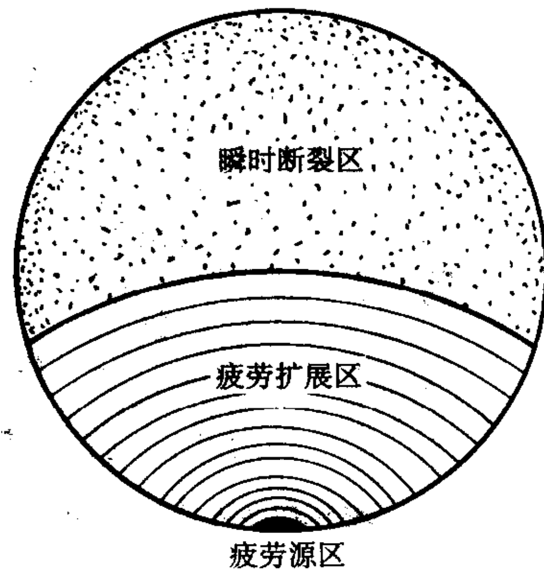


图 5.2-55　疲劳破坏断口上三个特征区示意图

Ⅰ）疲劳断裂源区宏观特征及位置的判别　通常宏观上所说的疲劳源区是指疲劳裂纹的萌生和疲劳裂纹扩展第一阶段区域。疲劳源区一般位于零件表面或亚表面的应力集中处，也可能在表面或内部的缺陷、损伤处，如表面加工刀痕、打伤、划伤、腐蚀坑处，内部的夹杂、脆性相、气孔、疏松等。疲劳源区的形成时期早，暴露于环境中的时间长，一般均有一定的氧化或腐蚀，同时，由于这阶段的裂纹扩展速率慢，在交变应力的作用下，裂纹反复张开和闭合，断面反复接触、分开，会造成不同程度的磨损；在大应力作用下，疲劳裂纹还会从多处起始，这些从不同高度起始的裂纹在扩展中将相遇、汇合，从而在汇合处形成辐射状台阶或条纹。因此，疲劳源区在断口上一般具有如下宏观特征：

a）相对于断口上的其他区域，其氧化或腐蚀较重，颜色较深；

b）一般位于零件应力集中处的表面或亚表面，位于零件内表面时，源区一定存在有缺陷；

c）断面平坦、光滑、细密，有些断口可见到闪光的小刻面；

d）有向外辐射的放射台阶和放射状条纹；

e）可能存在具有不同高度差的台阶；

f）在源区虽看不到疲劳弧线，但它看上去像向外发射疲劳弧线的中心。

有时疲劳源区不只一个，在存在多个源区的情况下，需要找出疲劳断裂的主源区。

Ⅱ）疲劳断裂扩展区的宏观特征　疲劳扩展区是疲劳裂

纹的稳定扩展阶段。该区断面较平坦，与主应力相垂直，颜色介于源区与瞬断区之间，一般最基本的特征是疲劳弧线(又称海滩花样或贝壳花样)。疲劳弧线是由于外载荷大小、方向发生变化，应力松弛或者材质不均，使得裂纹扩展不断改变速度和方向的结果。疲劳弧线的形状和疲劳弧线的间距变化等与受力状态、材质及环境介质等有关。



图 5.2-56 疲劳断口的宏观形貌

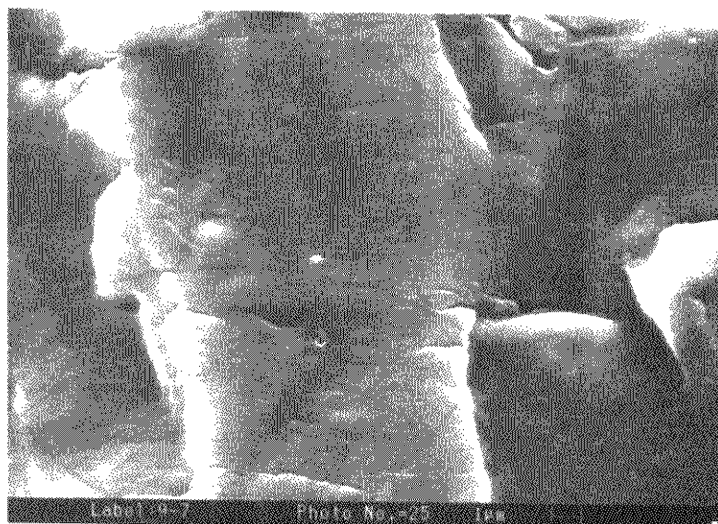
Ⅲ）最终断裂区的宏观特征 当疲劳裂纹的扩展达到临界尺寸（即零件剩余截面不足以承受外载荷时的尺寸）后发生失稳快速断裂，称为瞬时断裂。它的宏观特征与静载拉伸断口相近，即由纤维区、放射区和剪切唇三部分组成。

瞬断区面积的大小取决于载荷的大小、材料的性质、环境介质等因素。通常情况下，瞬断区面积较大，则表示所受载荷较大或材料较脆；反则反之。当疲劳裂纹扩展到应力处于平面应变状态以及由平面应变过渡到平面应力状态时，其断口宏观形貌呈现人字纹或放射条纹，当裂纹扩展到使应力处于平面应力状态时，断口呈现剪切唇状态。

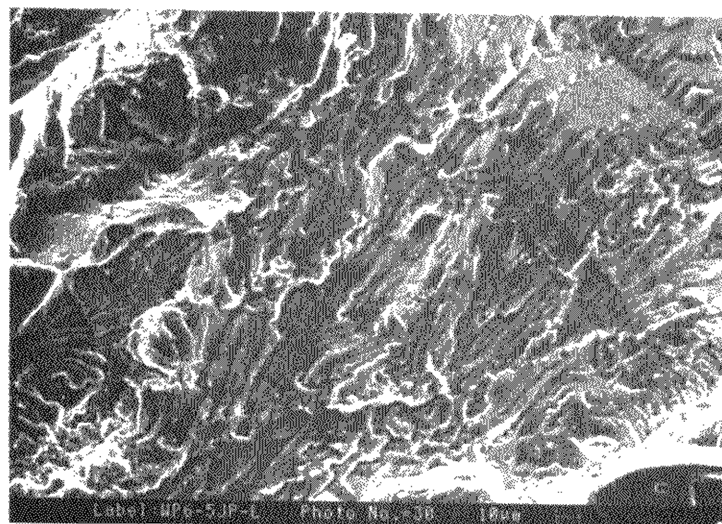
③ 疲劳断口的微观分析 某些情况下，宏观分析难以判明断裂的性质，找出准确的断裂源位置等，这就需要对断口进行微观分析。

Ⅰ）疲劳源区的微观分析。首先确定疲劳源区的具体位置是表面还是亚表面，判明主源与次源。其次要分析源区的微观形貌特征，包括裂纹萌生处有无外物损伤痕迹、加工刀痕、磨损痕迹、腐蚀损伤及腐蚀产物、材质缺陷等。疲劳源区的微观分析能为判断疲劳断裂的原因提供十分重要的信息与数据，是分析的重点。

Ⅱ）对于疲劳扩展的第一阶段，要仔细观察其上有无疲劳条带、韧窝、台阶、二次裂纹以及断裂小刻面的微观形貌。对第二阶段的微观分析主要是观察有无疲劳条带，疲劳条带的性质，条带间距变化的规律等（图 5.2-57）。这些特征对于分析疲劳断裂机制、裂纹扩展速率、载荷的性质等将起重要作用。



(a) 延性疲劳条带



(b) 脆性疲劳条带

图 5.2-57 微观疲劳条带

Ⅲ）瞬断区微观特征分析主要是观察韧窝的形态是等轴韧窝、撕裂韧窝还是剪切韧窝。这有助于判断引起疲劳断裂的载荷类型。

④ 疲劳载荷类型的判断 各种类型的疲劳断裂失效均是在交变载荷作用下造成的，因此，在分析疲劳断裂失效时，首要的是要以断口的特征形貌来分析判断所受载荷的类型。

Ⅰ）反复弯曲载荷引起的疲劳断裂 构件承受弯曲载荷时，其应力在表面最大、中心最小。所以疲劳核心总是在表面形成，然后沿着与最大正应力相垂直的方向扩展。当裂纹达到临界尺寸时，构件迅速断裂，因此，弯曲疲劳断口一般与其轴线成90°角。

弯曲疲劳可分为单向弯曲疲劳、双向弯曲疲劳（图 5.2-58)、旋转弯曲疲劳（图 5.2-59）等。

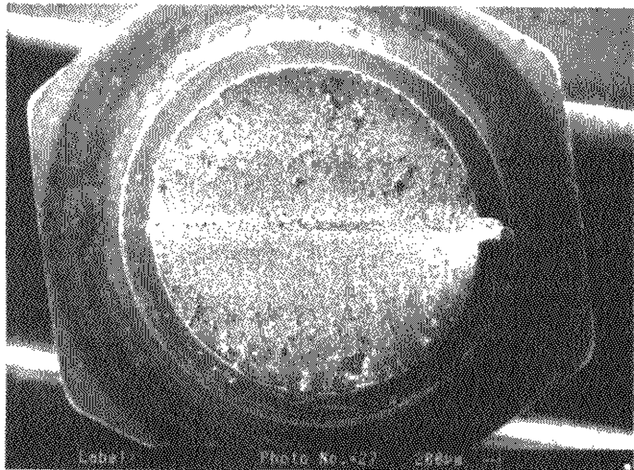
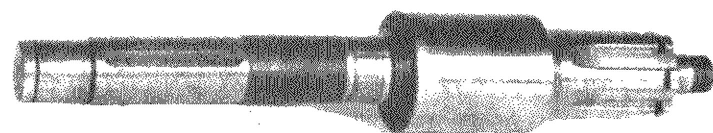
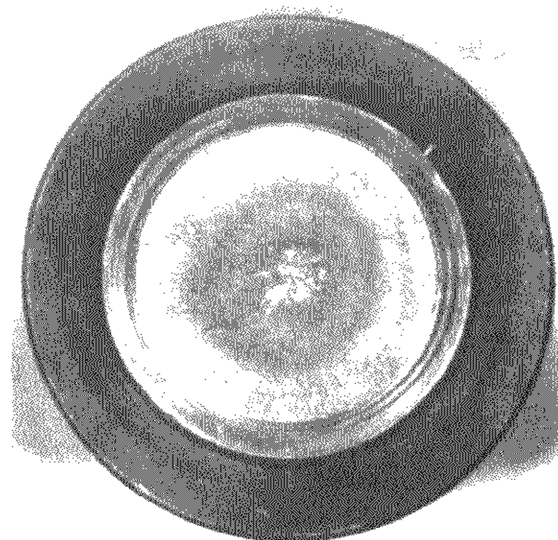


图 5.2-58 双向弯曲疲劳断口



(a) 轴



(b) 旋转弯曲疲劳断口

图 5.2-59 旋转弯曲疲劳断裂的轴断口

Ⅱ）拉——拉（压）载荷引起的疲劳断裂 当材料受到拉——拉（压）交变载荷时，其应力分布是沿着整个零件的横截面均匀分布，因此，疲劳源可以产生在零件的外表面，

也可以产生在零件的内部，这主要取决于各种缺陷在零件中的分布状态及环境因素的影响。

Ⅲ）扭转载荷引起的疲劳断裂 轴在交变扭转应力作用下，可能产生一种特殊的扭转疲劳断口——锯齿状断口。

⑤ 载荷大小与试样形状对断口形态的影响 载荷大小与试样形状对断口形态的影响见图 5.2-60。

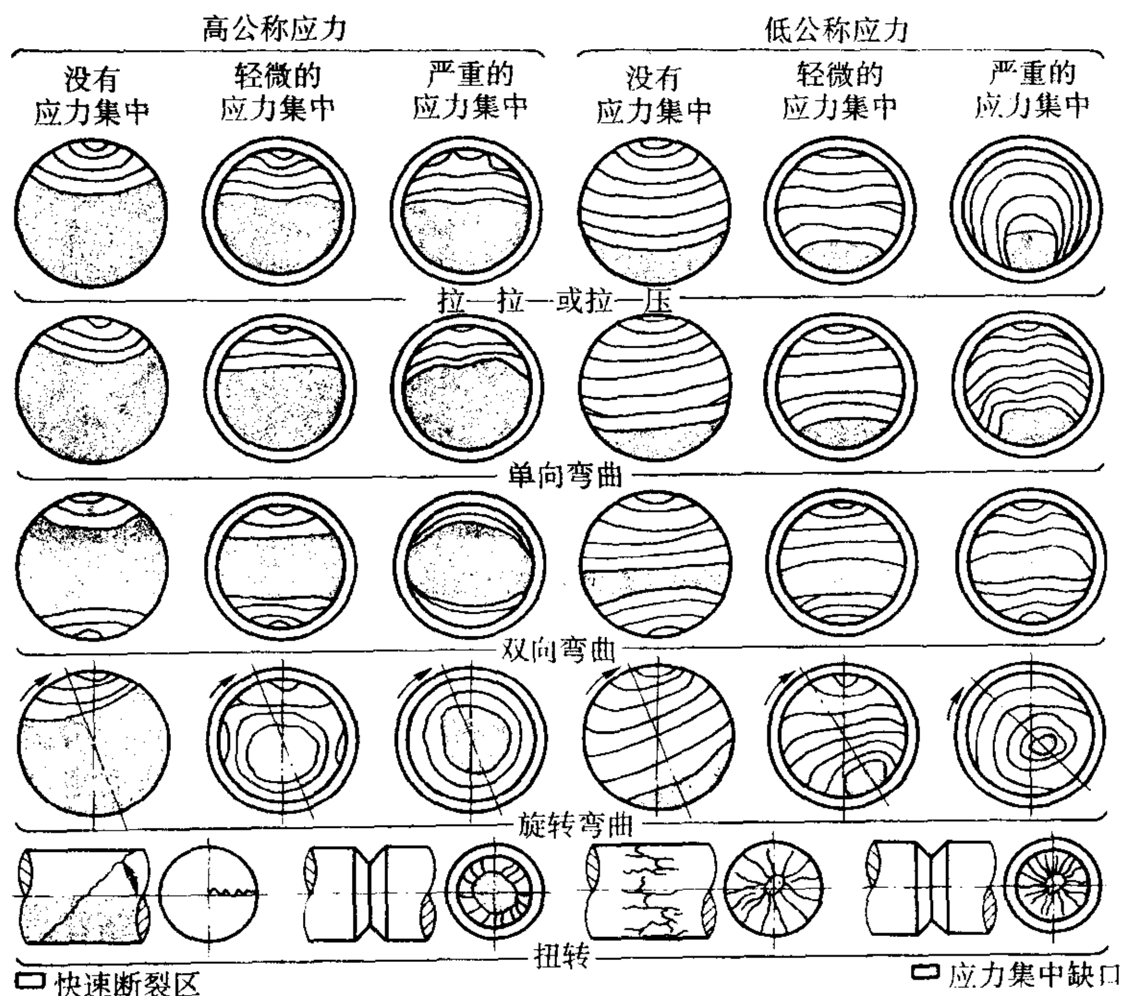


图 5.2-60 载荷大小与试样形状对断口形态的影响

⑥ 低周疲劳的判断

Ⅰ）宏观特征 低周疲劳断裂宏观断口除具有疲劳断裂宏观断口的一般特征之外，还有如下几点。

a）具有多个疲劳源点，且往往成为线状。源区间的放射状棱线（疲劳一次台阶）多而且台阶的高度差大。

b）瞬断区的面积所占比例大，甚至远大于疲劳裂纹稳定扩展区面积。

c）疲劳弧线间距加大，稳定扩展区的棱线（疲劳二次台阶）粗而短。

d）与高周疲劳断口相比，整个断口高低不平。随着断裂循环数的降低，断口形貌愈来愈接近静拉伸断裂断口。

Ⅱ）微观特征 低周疲劳断裂微观断口的变化是由于宏观塑性变形较大，静载断裂机理就会出现在疲劳断裂过程中，在断口上出现各种静载断裂所产生的断口形态。在一般情况下，当疲劳寿命小于 90 次时，断口上为细小的韧窝，没有疲劳条带出现；当疲劳寿命大于 300 次时，出现轮胎花样；当疲劳寿命大于 10 000 次时，才出现疲劳条带，此时的条带间距较宽，可达 2～3 μm/周。如果使用温度超过等强温度，还会出现沿晶断裂。

⑦ 腐蚀疲劳断裂分析 腐蚀疲劳断裂是在腐蚀环境与交变载荷协同、交互作用下发生的一种失效模式。在航空装备的实际运行中，因腐蚀疲劳而导致早期断裂失效的事例屡见不鲜。例如：起落架、机翼大梁、刹车轮毂、涡轮盘、叶片等关键部件，均曾发生过腐蚀疲劳断裂失效，有的酿成灾难性事故。

腐蚀疲劳断裂失效既不同于应力腐蚀开裂，也不同于一般的机械疲劳断裂，腐蚀疲劳对环境介质没有特定的限制，不像应力腐蚀那样，需要金属材料与腐蚀介质构成特定的结合关系。

影响腐蚀疲劳断裂过程（包括裂纹的萌生与裂纹的扩展过程）的相关因素。

a）环境因素 包括环境介质的成分、浓度、介质的酸度（pH 值）、介质中的含氧量、介质的电极电位以及环境温度等。

b）力学因素 包括加载方式、平均应力、应力比、载荷波形、频率以及应力循环周数。

c）材质冶金因素 包括材料的成分、强度、热处理状态、组织结构、冶金缺陷、夹杂物等。

机械疲劳、腐蚀疲劳和应力腐蚀三者之间的关系是相对的（图 5.2-61），要严格地将这三者加以区分是困难的。当应力比 $R=1$，且频率 f 很低时易产生应力腐蚀；当 $R=0$，f 为中等程度时，易产生腐蚀疲劳；随着 f 的升高，腐蚀的作用愈来愈小，趋于纯机械疲劳。

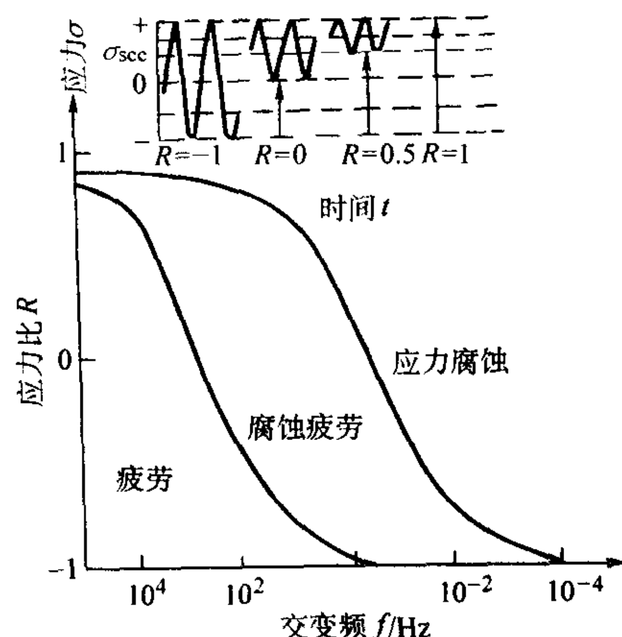


图 5.2-61 机械疲劳、腐蚀疲劳和应力腐蚀三者之间的关系

腐蚀疲劳的断裂特征和一般疲劳断裂一样，腐蚀疲劳的断口上也有源区、扩展区和瞬断区，但在细节上，腐蚀疲劳断口有其独特的特征，主要表现在如下几方面。

a）断口低倍形貌呈现出明显的疲劳弧线。

b）腐蚀疲劳断口的源区与疲劳扩展区一般均有腐蚀产物，通过微区成分分析，可以测定出腐蚀介质的组分及相对含量。但应当指出，疲劳断口上覆盖有腐蚀产物，并不一定

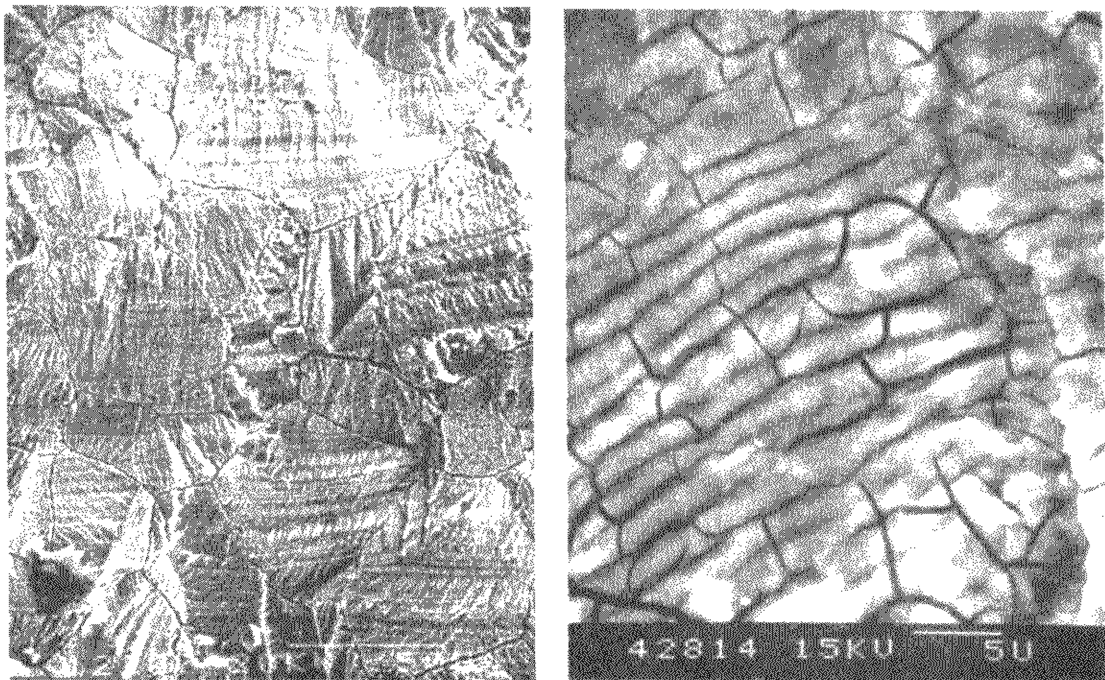


图 5.2-62　腐蚀疲劳断口微观形貌

就是腐蚀疲劳断裂。因为常规疲劳断裂后的断面上，亦有可能产生锈蚀（图 5.2-62）。

c）腐蚀疲劳断裂一般均起源于表面腐蚀损伤处（包括点腐蚀、晶间腐蚀、应力腐蚀等），因此，大多数腐蚀疲劳断裂的源区可见到腐蚀损伤特征。

d）腐蚀疲劳断裂扩展区有某些较明显腐蚀特征，如腐蚀坑、泥纹花样等。

e）腐蚀疲劳断裂的重要微观特征是穿晶解理脆性疲劳条带。

f）在腐蚀疲劳断裂过程中，当腐蚀损伤占主导地位时，腐蚀疲劳断口呈现穿晶与沿晶混合型。

g）当 $K_{max} > K_{Iscc}$，在频率很低的情况下，腐蚀疲劳断口呈现出穿晶解理与韧窝混合特征。

上述断裂特征并非在每一具体腐蚀疲劳断裂失效件上全部具备，对某一具体失效件是究竟具备上述特征的哪几项，随力学因素、环境因素、材质冶金因素而定。

腐蚀疲劳断裂模式的分析判据

在实际的失效分析中，判断腐蚀疲劳断裂失效的主要判据有如下几方面：

a）构件是在交变应力和腐蚀条件下工作，交变应力的频率和应力比一般处在腐蚀疲劳区内，在液态、气态和潮湿空气中有腐蚀性元素；

b）断裂表面颜色灰暗，无金属光泽，通常可见到较明显的疲劳弧线；

c）断裂表面上或多或少存在有腐蚀产物和腐蚀损伤痕迹；

d）疲劳条带多呈解理脆性特征，断裂路径一般为穿晶，有时出现穿晶与沿晶混合型甚至沿晶型。

腐蚀产物是分析、判断失效零件工作环境和工作时间的重要依据。可以采用能谱仪、电子探针以及其他化学分析方法确定腐蚀产物的化学元素及量的分布规律。

以上是对几种疲劳断裂“二级”失效模式的具体讨论，下面论述疲劳断裂模式的判据。

1）宏观变形诊断　疲劳断裂属于脆性断裂的一种，因此，同其他脆性断裂模式一样，在断裂位置的疲劳源区和疲劳扩展区附近也没有明显的宏观塑性变形，但在瞬断区附近则可能观察到明显的宏观塑性变形特征。

2）断口宏观形貌特征诊断　疲劳断口的最显著的宏观形貌特征是断口齐平，存在疲劳弧线（也有人称之为海滩标记、贝壳状花纹）。一般的疲劳断口（包括疲劳裂纹源区和扩展区）均与主应力轴垂直，断面齐平、细腻，附近没有明显的塑性变形；断口上有疲劳弧线和从疲劳源区向外发散的放射状棱线，有的源区还有疲劳台阶。

3）断口微观形貌特征诊断　疲劳断口的微观特征主要为疲劳条痕，包括疲劳条带、疲劳二次裂纹、疲劳韧窝带、轮胎花样等形貌。在有的疲劳试验件、实际失效件上，当循环周次少、交变应力较大时也可能在宏观上和微观上均观察不到疲劳特征。因此，在具体的失效分析中，即使没有观察到宏观的疲劳弧线、微观的疲劳条痕特征时，也轻易不要排除疲劳断裂的可能，需要结合其他特征来进一步进行诊断。

4）参数诊断　由于疲劳断裂是在交变载荷作用下导致的断裂失效，因此，其力学参数应该具有交替变化的特点，且交变的应力大小应该大于材料的疲劳极限。

5）断口颜色　由于疲劳区是裂纹在萌生和扩展过程中形成的，经历了一定的时间过程，而瞬断区只是最后断裂时形成的，因此，疲劳区和扩展区在氧化程度上一般都略有差异。疲劳区是裂纹缓慢扩展过程中形成的，而最后断裂区则是瞬时形成的，具有延性断裂的特征，因此，两者的光亮程度还有明显差异，且这种差异表示由于材质（包括状态）不同造成的。

综上所述，诊断金属零件疲劳断裂失效模式的判据如下（表 5.2-12）。

表 5.2-12　疲劳断裂失效模式的判据

序号	内容	特　征
1	宏观特征	断裂位置附近没有明显的宏观塑性变形
2	应力状态	交变动载荷，大于材料的疲劳极限
3	断口宏观形貌	断口齐平、光滑，具有宏观疲劳弧线和放射棱线、有的有疲劳台阶特征。断口可分为疲劳源区、扩展区和瞬断区三部分。疲劳源区一般位于零件表面应力集中处或缺陷处、内部缺陷处
4	断口微观形貌	疲劳条痕特征，如疲劳条带、平行的二次裂纹带、韧窝带、轮胎花样等
5	断口颜色	疲劳区颜色相对于瞬断区较暗，氧化较重，较光亮
6	组织	断口附近表面金相组织有明显的变形层
7	表面状态	断口附近表面脆性的镀层、涂层等表面覆盖膜破裂

需要说明的是，在实际失效事件中，疲劳失效的特征千变万化，各种因素对其的影响很大，具备以上所有疲劳失效特征的失效件很少见。一般具备以上特征中的断口宏观形貌特征、微观形貌特征和参数特征三者之一者，即可诊断其为疲劳断裂失效。

非金属材料也存在疲劳断裂失效，其失效特征随材料不

同而差异很大。以航空橡胶材料为例，一般难以观察到微观的疲劳条带特征，而只有疲劳弧线和放射棱线特征。图 5.2-63 是国产 5080 航空橡胶的疲劳实验断口，其宏观疲劳弧线和放射棱线特征典型，但观察不到微观疲劳条带特征。

疲劳失效模式的断口诊断技术方法：疲劳断口的典型宏观特征是疲劳弧线，典型微观特征是疲劳条带。因此，从断口进行诊断时，首先应该从宏观上判断其是否存在平坦、光滑区，是否存在贝壳状海滩花样。微观分析时，寻找微观疲劳特征应从瞬断区与疲劳区的交界开始，往疲劳区方向，即与疲劳裂纹扩展方向相反的方向来寻找。即应该在扩展区中后期寻找疲劳条带特征，而不要在源区附近去寻找。因为在疲劳裂纹扩展早期，裂纹扩展速度慢，条带间距小，当设备的分辨率较低时有时难以准确判断，而且，由于应力的交变作用，两侧断口还会反复接触、碰磨，使早期的疲劳特征损坏，难以观察到。而在疲劳扩展的中后期，特别是疲劳区与瞬断区的交界区附近，裂纹扩展速度较快，条带间距大，与瞬断区的特征差别也大，容易快速区分出来。对大应力疲劳断裂断口，有的疲劳区和瞬断区没有明显的分界时，则应将寻找的重点放在疲劳扩展区的中期。

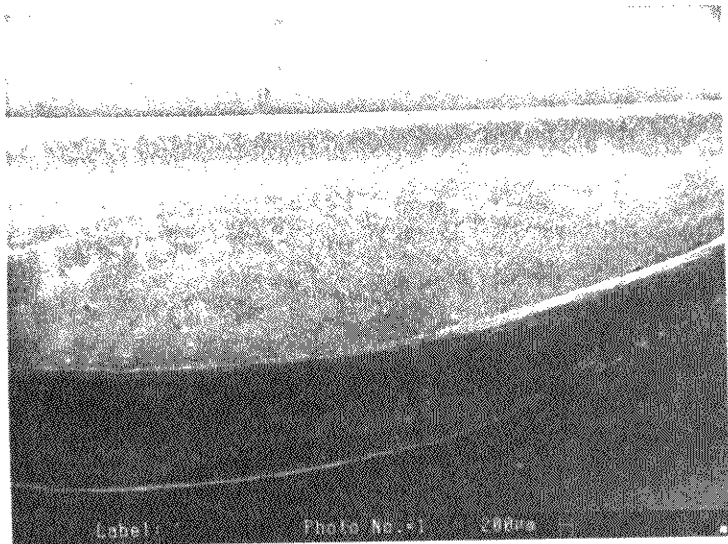


图 5.2-63　5080 航空橡胶疲劳断口形貌

总之，从断口诊断疲劳断裂失效模式时，应重点从疲劳扩展区寻找特征来确定。而诊断失效原因应该将重点放在疲劳源区。导致疲劳断裂失效的应力类型、大小诊断应该从源区形貌特征、扩展区条带特征和间距、瞬断区面积大小等来判断。

(4) 环境断裂模式诊断

机械产品的失效与环境有关。金属构件或整个机械产品的环境失效的主要方式是腐蚀及"环境"与应力共同作用下的破坏，如应力腐蚀、氢脆、腐蚀疲劳及液态金属致脆等。

1) 应力腐蚀断裂　金属构件在静应力和特定的腐蚀环境共同作用下所导致的脆性断裂为应力腐蚀断裂。

① 应力腐蚀的条件和特点

a) 引起应力腐蚀的应力一般是拉应力。

b) 纯金属不发生应力腐蚀破坏。但几乎所有的合金在特定的腐蚀环境中，都会引起应力腐蚀裂纹。

c) 金属材料只有在特定的活性介质中才发生应力腐蚀开裂。

d) 应力腐蚀断裂属脆性断裂。断口齐平，与主应力垂直，没有明显的塑性变形痕迹，断口形态呈颗粒状。

e) 应力腐蚀是一种局部腐蚀，而且腐蚀裂纹常常被腐蚀产物所覆盖，从外表很难观察到。

上述两个特征使应力腐蚀断裂成为断裂之前没有预兆的突然性断裂，不易预防，危害性极大。

f) 应力腐蚀试样持续加载时，应力腐蚀裂纹扩展大体上由三个阶段组成。在第一和第三阶段，da/dt 与 K_I 值有强烈的关系，而在第二阶段 da/dt 与 K_I 值无明显的关系。

g) 焊接、冷加工产生的残余应力和组织变化很容易成为应力腐蚀的力学原因。应力腐蚀的应力源可以由外加载荷引起，也可以由部件加工成型过程中，如铸造、锻造、轧制、挤压、机加工、焊接、热处理及磨削等工序中产生的残余应力引起。

h) 应力腐蚀断裂的速度比机械快速脆断慢得多。

i) 金属材料在腐蚀环境中所经历的过程也很重要。

② 应力腐蚀的断口特征　应力腐蚀断裂（图 5.2-64、图 5.2-65）是脆性断裂，因此应力腐蚀断裂断口的宏观特征

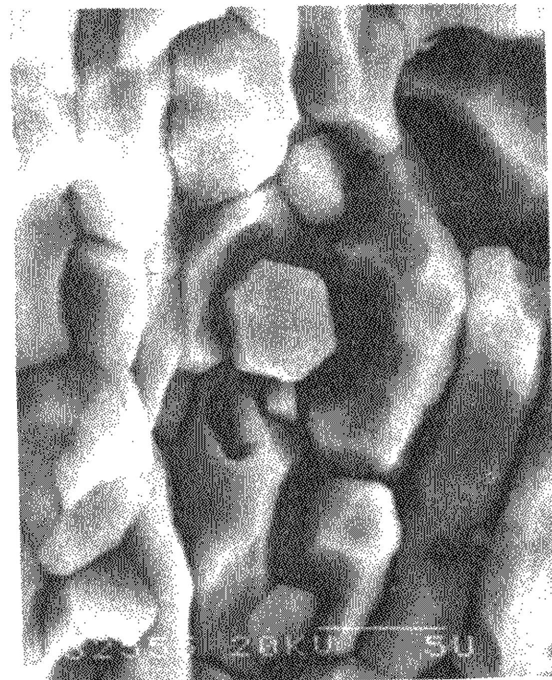


图 5.2-64　铝合金应力腐蚀断口形貌

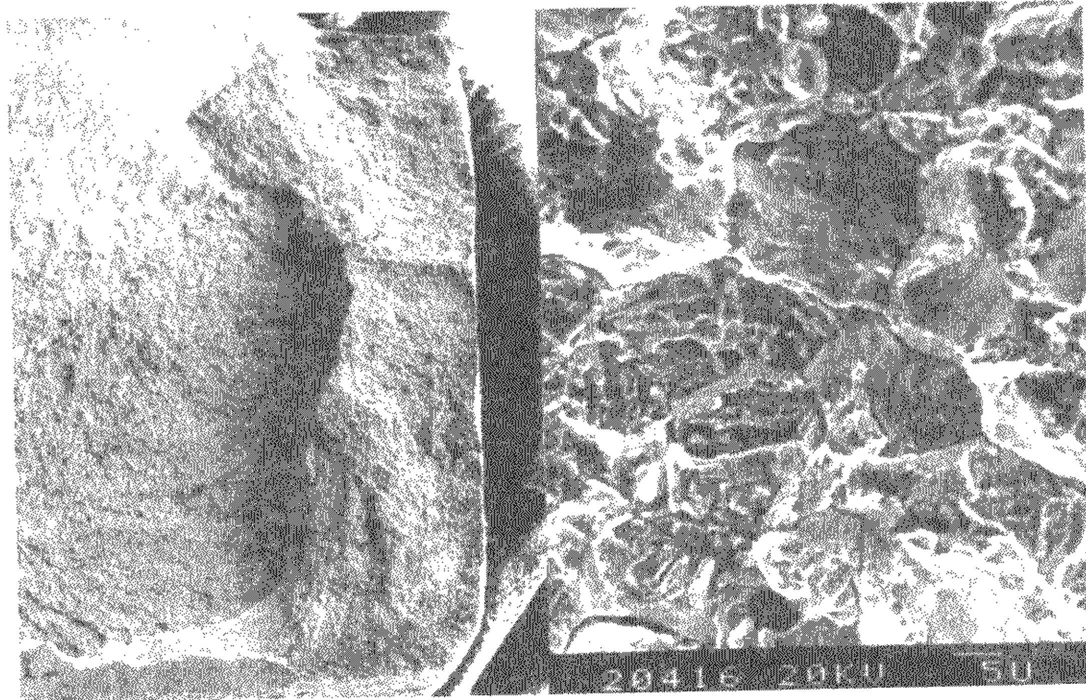


图 5.2-65　某机翼大量的应力腐蚀断口

为脆性断裂的特征：断口平直，并与主应力垂直，没有剪切唇，没有明显的塑性变形，断口表面有时比较灰暗，这通常是由于有一层腐蚀产物覆盖着断口的结果。同时应力腐蚀断裂起源于表面，且为多源，起源处表面一般存在腐蚀坑，且存在有腐蚀产物，离源区越近，腐蚀产物越多。应力腐蚀断裂断口上一般没有放射形花样。

应力腐蚀断口的微观形态可以是解理或准解理（河流花样、解理扇形）、沿晶断裂或混合型断口。

应力腐蚀的微观断口上还常见二次裂纹，沿晶界面上一般存在腐蚀沟槽，棱边不太平直。

应力腐蚀裂纹扩展过程中会发生裂纹分叉现象。用应力腐蚀的这一特征来区分实际断裂构件是应力腐蚀还是腐蚀疲劳、晶间腐蚀或其他断裂方式。

2）氢脆断裂　由于氢渗入金属内部导致损伤，从而使金属零件在低于材料屈服极限的静应力作用下导致的失效称为氢致破断失效，俗称氢脆。

① 氢的来源　金属材料在加工、制造过程中以及在使用环境下很容易受到氢的侵入。

a）在钢的冶炼、焊接及热处理过程中进入的氢。

b）在电镀、酸洗及放氢型腐蚀环境中产生的氢。

c）在使用环境下氢的渗入：金属材料处于在高温的氢气氛中。

② 氢脆断口的特征　氢脆断口宏观形貌主要特征是：断口附近无宏观塑性变形，断口平齐，结构粗糙，氢脆断裂区呈结晶颗粒状，色泽为亮灰色，断面干净，无腐蚀产物。非氢脆断裂区呈暗灰色纤维状，并伴有剪切唇边。

金属氢脆断口微观形貌一般显示沿晶分离，也可能是穿晶的，沿晶分离系沿晶界发生的沿晶脆性断裂，呈冰糖块状。断口的晶面平坦，没有附着物并存在鸡爪形的撕裂棱（见图 5.2-66）。



图 5.2-66　氢脆断口微观形貌

3）液态金属致脆　液态金属致脆指的是塑性金属或合金与液态金属接触后导致塑性降低而发生脆断。

液态金属致脆裂纹一般沿晶扩展，仅在少数情况下发生穿晶扩展。导致开裂的表面通常覆盖着一层液态金属，对该层表面膜进行化学分析是判断液态金属致脆的重要途径。

镉脆断口形貌见图 5.2-67。

5.2.2　磨损失效模式诊断技术和方法

磨损是机械零部件的三种主要失效模式（磨损、腐蚀和断裂）之一，是相互接触的一对金属表面相对运动时，表面金属不断发生损耗或产生残余塑性变形，使金属表面状态和尺寸改变的现象。任何零件之间在接触状态下发生相对运动（滑动、滚动），都会产生摩擦，进而导致磨损。磨损轻则可使零件部分失去其规定的功能，重则可使零件无法使用。因此，分析磨损及其表面损伤的规律和原因，对提高零件的耐磨性、延长机件寿命具有重要的意义。



图 5.2-67　镉脆断口微观形貌

根据定义，磨损是在使用中产生的，是一种逐步损坏的过程，一般是一种表面现象。磨损的分类形式有很多种，其中有些还不乏争议。比较常用的分类方法是按照磨损的破坏机理来分类，将磨损分为黏着磨损、磨粒磨损、表面疲劳磨损和腐蚀磨损等四类。

（1）黏着磨损

接触并相对运动的表面发生了固相黏着，使材料从一个表面转移到另一表面的现象，称为黏着磨损。黏着磨损是一种严重的磨损方式，也称为磨伤、擦伤、咬住和结疤，是两个金属表面在压力下相互滑动时发生的，有时可使摩擦副咬死。

1）黏着磨损的机理　不管经过怎样仔细的抛光，金属表面实际上均是高低不平的，故当两物体接触时，总是只有局部显微凸起或凹凸不平处接触，其真实接触面积很小，从而使接触处即使在较小压力下也会产生非常高的局部压力而塑性变形，导致表面的润滑油膜、氧化膜破裂，从而使两物体的金属面直接接触而发生黏着，随后在相对滑动时黏着点又被剪切而断裂，从一个表面上撕下金属，并把它转移到另一个表面上。这样，在一个表面上形成微小的凹坑，在另一个表面上形成微小的凸起，从而又造成进一步的损伤。黏着点的反复形成与断裂就造成了黏着磨损。

黏着磨损过程也可以产生松散的磨损粒子，造成磨料磨损。

2）黏着磨损的特点与分类　黏着又称为冷焊。当摩擦副相对运动时，将产生大量的摩擦热并产生高温。黏着点被剪切的部分实际上受到局部高温和应变强化的作用，产生的黏着块的强度一般高于摩擦副的强度。

黏着磨损的典型特征是接触点局部的高温使摩擦副材料发生相互转移，因此，对整个摩擦副来说，在一定程度上可保持摩擦副材料的质量总和不变，这一现象已被许多实验所证实。

按照零件表面的损坏程度，通常把黏着磨损又分为五类，其各自的破坏现象和原因见表 5.2-13。

3）影响黏着磨损的因素

① 材料特性的影响　脆性材料的抗黏结能力比韧性材料高，韧性材料的黏着破坏常常发生在离表面一定的深度部位，磨损下来的颗粒较大；脆性材料的黏着磨损的破坏深度较浅，磨损产物呈金属磨屑碎片状。

互溶性大的材料所组成的摩擦副，黏着倾向大；反之，黏着倾向小。

多相金属比单相金属黏着倾向小；金属与非金属材料组成的摩擦副比金属组成的摩擦副黏着倾向小。

元素周期表中的 B 族元素与 Fe 不相溶或能形成化合物，

它们的黏着倾向小；而 Fe 与 A 族元素组成的摩擦副偶件黏着倾向大。

表 5.2-13　黏着磨损类别及现象、原因

类型	破坏现象	破坏原因
轻微磨损	剪切破坏发生在黏着结合面上，表面转移的材料数量极少	黏着结合强度比摩擦副的基体金属都弱
涂抹	剪切破坏发生在离黏着结合面不远的软金属层内，轻金属涂抹在硬金属表面	黏着结合强度大于较软金属的抗剪强度
擦伤	剪切发生在较软金属的亚表层内，有时硬金属表面也有划痕	黏着结合强度比两基体金属都高，转移到硬面上的金属物又拉削软金属表面
撕脱	剪切破坏发生在摩擦副一方或两方金属较深处	黏着结合强度比两基体金属都高，切应力高于黏着结合强度
咬死	摩擦副之间咬死，不能相对运动	黏着结合强度高于任一基体金属的抗剪强度，黏着区域大，切应力低于黏着强度

② 接触压力与滑动速度的影响　黏着磨损量的大小与接触应力、摩擦速度密切相关。在摩擦速度一定时，黏着磨损量随接触压力的增大而增大，当接触压力大于材料硬度的 1/3 时，黏着磨损量急剧增大，严重时摩擦副间可咬死。当接触应力一定时，黏着磨损量随滑动速度的增加而增加，达到某一极限速度后，又随滑动速度的增加而减少。

随着滑动速度的增加，磨损类型可能会发生变化，由一种形式变为另一种形式，如由黏着磨损变为氧化磨损。

③ 其他因素的影响　摩擦偶件表面的粗糙度、温度和润滑状态对黏着磨损也有较大的影响。减小表面粗糙度，可提高零件的抗黏结磨损能力；但当粗糙度减小到一定程度后，由于摩擦偶件表面不能储存足够的润滑剂，摩擦副之间润滑不良，从而促进黏着磨损。

温度升高，同样条件下的黏着磨损量增加。

润滑不良或没有润滑，将增加黏着磨损量。

(2) 磨粒磨损

一个表面同它的匹配表面上的尖硬凸起物或同相对磨损表面运动的硬粒子接触时造成的材料转移称为磨粒磨损，也称为磨料磨损或研磨磨损。当摩擦副之间存在硬粒子时，硬粒子可存在于两个滑动表面之间，磨削其中的一个或两个表面，或者被镶嵌在其中的一个表面上去磨削另一个表面。

1) 形成机理　磨粒磨损的机理是在一定的接触压力下，硬材料的凸出部分压入软材料中至一定深度，摩擦副之间相对滑动时，软材料将被切削出与硬材料突出部分对应形状的切屑，切屑断裂形成磨屑；磨屑的形成又加剧磨损。

2) 磨粒磨损失效的过程与特点　由磨粒磨损的形成机理可知，磨粒磨损失效的主要过程为显微切屑、金属堆挤、辗压抹平、断裂剥落。因此，磨粒磨损的最显著特点是在接触面上有明显的磨削痕迹。同时，磨粒磨损还具有自加速特性。

3) 影响磨粒磨损失效的主要影响因素　在对磨粒磨损机理研究的基础上，人们得出磨损量 W 的定性表达式如下：

$$W \propto Pl\tan\theta/\mathrm{HV}$$

式中，W 为磨损量；P 为接触压力；l 为滑动距离；θ 为硬材料上突出部分与软材料平面的接触夹角；HV 为软材料的显微硬度。

由上式可知，磨粒磨损的磨损量与接触应力、摩擦距离成正比，与材料的显微硬度成反比，与硬材料表面凸出部分尖端的形状有关。

1) 材料硬度的影响　材料的显微硬度 HV 越高，材料的磨粒磨损相对耐磨性 ε 越好。

2) 磨粒特性的影响　从磨粒磨损的过程可以看出，滑动磨粒磨损的主要机理是显微切削，即磨粒像金属切削刀具一样切削金属材料而产生磨屑。因此，磨粒的形状、硬度及大小等对磨损过程均有重要的影响。

3) 其他因素的影响　接触应力大小、润滑条件、材料显微组织、滑动速度、加工硬化等对磨粒磨损的磨损量均有重要影响。接触应力、滑动速度加大，均会增加磨损量；良好的润滑可减少磨损量；材料加工硬化程度高，再次切削的量就少，磨损量就会减少。

(3) 疲劳磨损

两接触表面作滚动或滑动，或两者复合的摩擦状态，在交变接触应力作用下，使材料表面疲劳而产生物质流失的过程称为表面疲劳磨损，也称为接触疲劳磨损。齿轮表面、轴承表面和凸轮表面经常容易产生表面接触疲劳。

1) 金属表面疲劳磨损的特点和形貌特征　金属表面疲劳多发生在表面缺陷处或浅层表面缺陷处，如冷加工表面划伤处、刀痕、冶金缺陷、热加工缺陷、组织局部不均匀、流线不均匀等部位，也可在表面晶界处或界面处。

金属表面疲劳过程也包括疲劳裂纹萌生、扩展和瞬断的过程。表面疲劳磨损的直接原因是金属表面的凹凸不平及表面油膜破坏，从而使得两个物体的表面接触不连续，每次循环中都接触的表面在接触正应力和摩擦力的作用下，产生局部变形，并使表面塑性区及其周围的温度升高，当表面塑性流动达到一定程度时产生裂纹，随后裂纹逐渐扩大，达到临界值时，表面与裂纹之间的材料被剪断，产生薄片状磨屑。这种表面损伤过程称为脱层过程。

由于脱层过程受塑性变形、空洞形成和裂纹扩展所控制，所以材料的显微组织参量，如材料的硬度、第二相硬质点的大小和形状、杂质的数量等都影响脱层的形成。材料的力学性能对脱层的形成也有密切关系。

表面疲劳引起表面金属小片脱落后，在金属表面形成一个深度在几微米到几百微米之间的麻坑。麻坑较小时，在以后的多次应力循环时可以被磨平；当尺寸较大时，麻坑成为凹下的舌头状，成椭圆形。麻坑附近有明显的塑性变形痕迹，塑性变形中金属的流动方向与摩擦力的方向一致。在麻坑的前缘和坑的底部有多处微裂纹，它们是没有明显发展的表面疲劳裂纹和二次裂纹。

表面疲劳磨损是一种介于疲劳和磨损之间的破坏形式，相当于周期脉动压缩加载情况，有疲劳裂纹起源、扩展和剥落的过程；也有疲劳极限等，这时它类似于一般疲劳的地方。但它还存在表面摩擦现象，表面有塑性变形、氧化磨损和润滑介质作用等情况，这是它不同于疲劳而与磨损相似的地方。

表面疲劳磨损失效的另一个重要特点是疲劳起源于相互接触的表层下一定深度（最大剪切应力处），产生表层破碎，最后引起剥落。

2) 影响疲劳磨损的因素　凡是对磨损、疲劳裂纹形核、扩展有影响的因素，均可对表面疲劳磨损产生影响。

① 材质的影响　钢中的非金属夹杂物，特别是脆性的带有棱角的氧化物、硅酸盐及其他各种复杂成分的点状和球状夹杂物，他们破坏了基体的连续性，对疲劳磨损有严重影响。

材料的组织状态也有重要影响。粗大晶粒、带状组织、大块碳化物等均可降低材料的抗疲劳磨损能力。

② 表面硬度的影响　轴承钢抗表面疲劳磨损的最佳表面硬度为62HRC，此时，材料的抗疲劳磨损能力最大。随着硬度降低或增加，抗疲劳磨损能力均会大幅下降。对齿轮材料，具最佳抗疲劳磨损能力的硬度为58～62HRC，一般均要求小齿轮的硬度大于大齿轮的硬度。

③ 表面粗糙度的影响　表面粗糙度越低，抗疲劳磨损的能力越强。

④ 残余内应力的影响　当表面层在一定深度范围内存在残余压应力时，可减少疲劳磨损。

⑤ 其他影响因素　提高装配精度，可有效降低接触应力，从而降低疲劳磨损。润滑油的黏度越高，接触部位的面积越大，接触压力越接近平均分布，对防止疲劳磨损越有力。

(4) 腐蚀磨损

腐蚀磨损是由于外界环境引起金属表层的腐蚀产物（主要是氧化物）剥落，与金属磨面之间的机械磨损（磨粒磨损与黏着磨损）相结合而出现的损伤，又称为腐蚀机械磨损。包括氧化磨损、微动磨损两种形式。

1) 氧化磨损

① 定义及形成机理　在摩擦副中，金属表面被氧化、氧化膜被切削剥落、新露出的金属表面又被氧化、新形成的氧化膜再被切削剥落，如此反复的过程称为氧化磨损。氧化磨损是最广泛的一种磨损状态，不管在何种摩擦条件下均会发生，因此，氧化磨损也是一种无法避免的磨损，但它是各类磨损中磨损速率最小的。

② 影响因素　氧化磨损速率取决于所形成的氧化膜的性质和氧化膜与基体的结合能力，以及金属表面的塑性变形抗力。其中，氧化膜的性质主要是指其脆性程度。致密而非脆性的氧化膜能显著地提高零件的磨损抗力。如广泛采用的发蓝、磷化、蒸汽处理、渗硫及有色金属的氧化处理等，均对降低氧化磨损速率有较好的效果。氧化膜与基体的结合强度主要取决于它们之间的硬度差，硬度差越小、结合力越强，氧化膜越不易剥落，则抗氧化磨损能力越好。提高基体表层硬度，可增加表层塑性变形能力，从而减轻氧化磨损。

2) 微动磨损　两个名义上属于静配合的表面之间由于一微小振幅而不断往复滑动引起的磨损称为微动磨损。

① 微动磨损的特点　微动磨损与一般磨损的区别在于引起磨损的往复滑动的距离不同，即微动磨损的滑动距离很小，一般仅在几十微米以下。因此，它有以下特点。

a) 往复滑动的速度小，磨损缓慢。由于往复滑动的振幅小，滑动的相对速度低，因此，微动磨损的构件属于高频、小振幅，其微动时的运动速度和方向不断改变，其速度始终在0与最大速度之间变化，但最大速度不会很大，因此，基本上属于慢速运动，外观上一般难以发现往复式滑动，所以其磨损过程很缓慢。

b) 由于振幅小，且属于往复性的相对摩擦运动，因此，磨屑很难逸出；摩擦面多为三体磨损，磨粒与金属表面产生的接触应力大，常超过磨粒的压溃强度，使延性金属的摩擦表面产生塑性变形或疲劳，使脆性金属的摩擦表面产生脆裂或剥落。

c) 微动磨损损伤属于表面损伤，损伤涉及的范围（深度）很小，基本与微动的幅度相当。

d) 磨损产物：铁基金属的微动磨损产物为红棕色粉末，主要成分为$\alpha-Fe_2O_3$和FeO；铝和铝合金的微动磨损产物为黑色粉末，主要成分为金属铝和Al_2O_3；铜、镁、镍等金属的磨屑多为黑色氧化物粉末。

② 微动疲劳的特征　工程中一构件与其他构件接触面间发生微动磨损的条件下受交变载荷作用而发生的疲劳损伤过程称为微动疲劳。它是微动磨损、氧化及腐蚀、交变应力综合作用的结果。微动疲劳也包括裂纹的萌生和扩展过程，微动疲劳裂纹一般萌生于微动磨损造成的表面损伤的边界处，如皿状浅坑的边缘或微动磨损深坑的边缘。如没有微动磨损存在，零件承受的交变载荷根本不足以使疲劳裂纹萌生，因此，微动磨损是微动疲劳产生的根本原因。

微动磨损的初期可产生多个疲劳裂纹，这些微裂纹同时扩展，在扩展过程中可合并为一主裂纹并垂直于外加交变正应力而进一步扩展。在扩展过程中，疲劳裂纹的扩展速率受外加应力强度因子幅和微动磨损过程中形成的腐蚀产物及腐蚀性介质（空气、水、润滑剂）进入裂纹内部产生的附加应力强度因子幅叠加作用的影响，而大大加速。同时，化学作用的影响也会加速疲劳裂纹的扩展。

影响微动疲劳寿命的主要因素是微动磨损过程中配合表面之间的法向夹紧压应力、相对运动幅度、摩擦力、内应力、周围介质、相匹配面的材料等。一般，随夹紧压应力的增加，微动疲劳寿命会降低，但达到一定值后，再增加夹紧应力对微动疲劳寿命的影响已不大。

③ 微动疲劳的断口特征　微动疲劳断口的宏观、微观形貌与纯机械疲劳断口的完全一致，整个断口在宏观上也可分为疲劳源区、裂纹扩展区和瞬断区，但瞬断区的面积相对较小；在微观上，可以看到典型的疲劳条带。

在裂纹源区和扩展区的前期，往往可以看到腐蚀产物，对这些产物的分析可为诊断微动疲劳损伤失效的模式和原因提供大量信息。

微动磨损失效的最明显特征是在断口的侧表面，即微动磨损面上有大量的微裂纹、表面金属掉块、不均匀磨损擦伤，色泽发生明显改变且有腐蚀坑。微动产生的微裂纹大多集中于微动区的边缘，大多与表面呈45°角，断口常呈杯锥状。

微动损伤表面还常常可以看到层状及山丘状的塑性变形，同时还可看到由于辗压形成的微裂纹。

④ 微动磨损失效模式的判据　微动磨损失效的失效模式可根据以下判据对其进行诊断：

a) 工况条件存在引起紧配合表面间滑动的振动或交变应力；

b) 接触表面存在麻点坑或小划痕，以及残留的磨屑；

c) 在与接触表面垂直的面上，可观察到接触亚表层有微小裂纹，这些微裂纹可与表面平行，也可与表面呈一定角度；

d) 磨屑的颜色，钢的磨屑比普通铁锈红得多，极易团聚；铝的磨屑是黑色，而氧化铝通常是白色的。

⑤ 微动磨损的形成机理　从微动磨损发展的过程可知，磨损初期为黏着并形成粘着磨屑、磨屑的研磨和磨屑的氧化；在稳定磨损阶段，以疲劳磨损为主的磨损。当存在应力循环时，微动磨损不但会造成零件的装配松动，还会造成微动疲劳而引起断裂。

5.2.3　腐蚀失效模式诊断技术和方法

金属表面与周围介质发生化学及电化学作用而遭受破坏的现象称为金属腐蚀。在国民经济的各个方面均有大量的金属构件和装备因为腐蚀问题而报废。据国外统计，每年由于金属腐蚀而报废的金属材料约占金属年产量的20%～40%，损失巨大。当腐蚀与其他因素，如力学因素共同作用时，往往还会造成灾难性的事故，损失更加惊人。

腐蚀类型、腐蚀速率及其发展程度受金属表面的性质、成分及均匀性（或不均匀性）和环境温度、湿度、金属零件或装配件之间的缝隙、环境和金属零件之间的相对运动，以及在导电环境中存在的异种金属的影响。加工工艺过程和表

面质量因可使金属表面产生局部或全面的变化而对表面腐蚀敏感性有影响。

按照腐蚀的原理，可将腐蚀引起的失效分为化学腐蚀失效和电化学腐蚀失效两大类。

本节主要讨论单纯的化学腐蚀和电化学腐蚀的失效问题，腐蚀与其他因素共同作用、腐蚀只起促进作用的失效已在环境断裂模式诊断一节中论述，本节不再赘述。

(1) 腐蚀的分类及其破坏形式

1) 化学腐蚀失效　化学腐蚀是金属表面与周围介质发生化学作用而引起的损伤。其特点是在腐蚀进行过程中没有电流产生。化学腐蚀又可分为气体腐蚀和在非电解质溶液中的腐蚀两小类。

气体腐蚀是指金属在干燥气体中（表面没有湿气冷凝）发生的腐蚀，一般指金属在高温时的氧化或腐蚀。在非电解质溶液中的腐蚀是指金属在不导电的液体中发生的腐蚀，如金属在有机液体（如酒精、石油等）中的腐蚀。

由于几乎所有的金属以及许多非金属材料在空气中均是不稳定的，有通过反应降低自由能的趋势，因此，它们与空气作用而形成氧化物，在有些情况下则形成硫化物和氮化物。因此，金属化学腐蚀的特点之一是能在金属表面形成一层腐蚀产物——氧化膜。这层氧化膜有时可对金属表面起到保护作用，但保护作用的程度与氧化膜的完整性、致密性和氧化的动力学密切相关。氧化膜越完整、越致密，对氧原子的扩散阻碍作用越强烈，则氧化膜的保护作用越好。由于光线照射到氧化膜时，将在氧化膜的外表面和内表面均发生反射，反射出来的光线之间发生干扰，从而使氧化膜呈现出颜色，金属基体的种类不同、氧化的温度和时间不同，导致的氧化膜的结构和厚度不同，氧化膜表现出不同的颜色。

2) 电化学腐蚀失效　电化学腐蚀是金属表面与周围介质发生电化学作用而引起的损伤，其与化学腐蚀的区别是电化学腐蚀在进行过程中有电流产生。按照所接触的环境介质的不同，电化学腐蚀可分为如下几种。

① 大气腐蚀　在潮湿气体中进行的腐蚀。

② 土壤腐蚀　埋设在地下的金属结构件的腐蚀。

③ 在电解质溶液中的腐蚀　这是电化学腐蚀中极为广泛的一类腐蚀，天然气及大部分水溶液对金属结构的腐蚀（如在海水和酸、碱、盐的水溶液中所发生的腐蚀）都属于这一类腐蚀。

④ 在熔融盐中的腐蚀　如在热处理中，熔盐加热炉中的盐炉电极和所处理的金属发生的腐蚀。

⑤ 接触腐蚀　两种电极电位不同的金属相互接触时发生的腐蚀。

⑥ 缝隙腐蚀　在两个零件或构件的连接缝隙内产生的腐蚀。

⑦ 其他条件下发生的腐蚀　如生物腐蚀、杂散电流的腐蚀、摩擦腐蚀、液态金属中的腐蚀等。

按照腐蚀破坏的形式，可将腐蚀分为三大类：均匀腐蚀、局部腐蚀和腐蚀断裂。均匀腐蚀是均匀地发生在整个金属表面上的腐蚀；局部腐蚀是其腐蚀作用仅局限在一定区域内的腐蚀；腐蚀断裂失效则是金属材料在受到腐蚀物质作用的同时又受到应力的作用，最后导致零件或构件的断裂。

局部腐蚀又可分为如下几种。

① 斑点腐蚀　腐蚀象斑点一样分布在金属的表面，面积大，但深度不大，如轻微汽蚀。

② 脓疮腐蚀　金属腐蚀破坏的情况类似人身上长出的脓疮，被破坏的地方较深、较大。

③ 点蚀　在金属某些部位被腐蚀成为一些小而深的圆孔，有时甚至发生穿孔。

④ 晶间腐蚀　沿金属晶体界面腐蚀的腐蚀。金属遭受晶间腐蚀后，其晶粒间的结合力显著减小，从而使力学性能大大降低。

⑤ 穿晶腐蚀　这是一种沿最大张应力线发生的局部腐蚀，特征是腐蚀是穿晶的。

⑥ 选择腐蚀　在多元合金中，某一组份溶解到腐蚀介质中，从而造成另一组份富集在合金的表面上。如黄铜脱锌，锌被溶解到介质中去，金属表面上富集着铜而呈红色。

⑦ 剥蚀　这是变形铝合金中常见的一类腐蚀，其本质是晶间腐蚀。由于顺变形方向的晶粒被拉长，腐蚀沿晶界进行时，因腐蚀产物的体积效应，从而使金属呈层状剥离特征。

(2) 金属腐蚀失效分析

1) 均匀腐蚀　金属由于均匀的化学浸蚀而产生的腐蚀为均匀腐蚀，是最简单而又常见的一种腐蚀形式。如钢的锈蚀、银的变暗就是典型例子。

发生均匀腐蚀的金属表面一般均有均匀的化学成分和显微组织，造成腐蚀的环境介质与金属表面的接触也是均匀的，没有限制。在钢等金属中，会因均匀腐蚀而有大量的金属损失，使金属表面变得粗糙，表面附着有疏松的、多孔的腐蚀产物层。而银在空气中变暗和铅在含有硫酸的环境中腐蚀这些反应中，产生薄而致密附着的保护膜，金属表面保持光滑状态。

2) 点蚀　金属材料或构件与环境中的游离物质（如氯离子）之间发生电化学作用而导致零件的损伤称为点蚀。当金属的一个表面区域相对于其余表面为阳极时，或在与金属接触的腐蚀介质浓度发生大的局部变化时，都可产生点蚀。点蚀是一种非常集中的腐蚀，其危害性相当大。点蚀一般都产生极其尖锐的小孔，而孔内壁的腐蚀常常是相当均匀的，但在特定的条件下会引起次生的晶间腐蚀的地方，也可能是不规则的。工程上常用的金属及其合金对点蚀都比较敏感。金属表面结构的不均匀性，特别是表面的夹杂物、表面氧化膜的不完整性是点蚀发生的重要原因。

金属零件的点蚀坑边缘比较平滑，因腐蚀产物覆盖，坑底呈深灰色。垂直于蚀坑磨片观察，蚀坑多呈半圆形或多边形。点蚀并不一定择优扩展。菊花形点蚀坑往往外小内大，类似蚁巢。

引起金属点蚀失效的腐蚀介质一般均与氯化物或氯离子有关。腐蚀介质中的氯离子浓度越高，合金越容易产生点蚀。如氯化物中含有铜、汞及铁等金属离子时，则点蚀的倾向性更大。

3) 晶间腐蚀　晶间腐蚀是指金属材料或构件沿晶界产生并沿晶界扩展的腐蚀损伤，其实质是晶界上的相或直接靠近晶界的区域择优溶解。晶间腐蚀不仅降低材料的机械性能，而且由于不易发现，可能造成突然失效，危害性极大。晶间腐蚀严重时，可导致整个晶粒因其晶界完全被破坏而脱落，俗称掉晶。

金属晶界是排列整齐而又具有稳定的晶体点阵结构的晶粒间紊乱失配的地方，同时也是各种溶质元素偏析或金属化合物（如碳化物）沉淀析出的有利区域，因此，大多数的金属或因一种合金元素在晶界处富集，或因碳化物分布不均匀或过饱和固溶体分解不均匀，而引起电化学不均匀，从而促使晶界成为阳极区，在一定的腐蚀介质中发生晶间择优溶解而导致腐蚀损伤。晶间腐蚀起源于表面，沿晶界扩展。

晶间腐蚀敏感性通常与热加工有关，如焊接热处理、去应力热处理等。如奥氏体不锈钢在 550~850℃温度范围内加热时，由于碳化物沿晶界析出，使得晶界处贫铬，而对晶间腐蚀敏感。沉淀强化的镍基合金，在某些环境中，如热的苛性溶液、沸腾的 75% 硝酸及含低浓度的卤化物或其他盐类的高温水溶液中，对晶间腐蚀都敏感。固溶镍基合金，如果

在热处理过程中产生晶界碳化物沉淀，也会对晶间腐蚀敏感。铝及其合金也有晶间腐蚀问题，高强铝合金的沉淀相会对晶间腐蚀敏感，对含有沿晶界或滑移线而沉淀出 $CuAl_2$ 的合金最为显著，而对沉淀出 $FeAl_3$、Mg_5Al_8、Mg_2Si、$MgZn_2$ 及 $MnAl_6$ 的合金，这种影响稍小。在室温下，钛和某些钛合金在发烟硝酸中会产生晶间腐蚀。各种钛合金，在空气中加热时，只要在260℃以上的温度下同潮湿的氯化钠接触，通常就会沿着晶界产生晶间腐蚀或应力腐蚀开裂。

4）大气腐蚀　金属表面与大气接触产生的腐蚀称为大气腐蚀。在大气环境中，由于雨水的直接沉淀或湿度、温度的变化引起凝聚，金属表面往往附着有一层极薄的水膜，当其厚度达到20～30个分子厚度时，水膜中可溶解大量的杂质和盐类，而成为电化学腐蚀所需的电解液膜，就可能发生电化学腐蚀。同时，由于金属本身的电化学不均匀性，形成了腐蚀微电池，促使其表面受到电化学腐蚀损伤。在污染严重的大气中，由于含有大量的腐蚀性强的气体，可使这种腐蚀表现得很严重。海洋大气中因含有大量的 NaCl 盐分，溶于水膜中后，可大大提高水膜的电导性，使腐蚀加速。

5）电偶腐蚀　两种具有不同电位的金属在电解液中处于接触状态时，即可产生电化学腐蚀，电位较负的金属（阳极）被腐蚀，且程度较单独腐蚀时严重，这就是电偶腐蚀，也称为接触腐蚀。机械设备中，不同类别的金属装配到一起时，在潮湿环境中，常常会发生接触腐蚀。如飞机铝合金长桁上的腐蚀往往集中在搭铁线处就是一个例证。

两种金属间的电位差越大，电偶腐蚀就会越快、越严重。按照金属的化学活性排序，金属的电动势（电位）由低到高的顺序如下：Na、Mg、Al、Ti、Mn、Cr、Zn、Fe、Co、Ni、Cu、Ag、Au。但在实际中，金属的溶解电位还受金属上存在的钝化膜或其他防护膜、极化效应、充气程度、络合剂及温度的影响。在海水中，金属的电偶序列发生了相应的变化，电位由低到高的顺序为：镁、镁合金、锌、镀锌钢（或铸铁）、铝合金、镉、低碳钢、铸铁、不锈钢、铜合金、镍基合金、银、钛、石墨、金、铂。在大多数的电解液中，靠近电偶序列表活性端的金属将起到阳极的作用而被腐蚀，靠近惰性端的金属则起到阴极端的作用而被保护；两金属在表中分开的程度则粗略地反映了电位差的大小。同一类金属，由于成分不同，其电位差可能存在差异，也可发生电偶腐蚀。

在实际产品中，由于没有对在电解液中处于电接触的金属进行选择，而让具有不同电位的金属接触，经常导致了金属的腐蚀失效。因此，在产品设计中，应引起重视。

6）缝隙腐蚀　金属零件缝隙腐蚀是指金属材料由于腐蚀介质进入缝隙并滞留产生电化学腐蚀作用而导致的损伤。显然，作为可导致电化学腐蚀的缝隙，其宽窄程度必须能使腐蚀介质进入并滞留其中，缝隙太宽，腐蚀介质无法滞留，太窄，腐蚀介质无法进入，均不会发生缝隙腐蚀。一般发生缝隙腐蚀的缝隙宽度均在几微米至几百微米。

金属表面的缝隙、两块金属表面之间的缝隙、或一块金属与一块非金属之间连接处的缝隙、或在金属表面上某一固体物质颗粒的下面的缝隙均可发生缝隙腐蚀。金属表面上，能从表面除氧、或是在沉积物的底下由于扩散受阻而使金属离子集聚的任何固态物质层都是缝隙腐蚀发生的可能部位。

缝隙腐蚀有时可发展得很快。

7）微生物腐蚀　微生物腐蚀是近年来引起广泛关注的一类腐蚀。微生物有机体通过对阳极或阴极反应的直接影响、对金属上的防护性表面膜的影响、产生腐蚀性物质以及产生固体沉积物等方法影响金属的腐蚀过程。这些有机体包括微观型的菌类及宏观型的海藻、贝壳等。

在潮湿的泥土、沼泽中存在的厌氧微生物可制造硫化物，从而对埋设于地下的钢材产生腐蚀作用。微生物污染也使飞机整体油箱在铝合金制成的下限储油区结构表面上产生腐蚀。

8）金属热腐蚀　金属的热腐蚀是在高温下，金属基体与沉积在表面的盐类（主要指 Na_2SO_4）及金属同气体综合作用产生的腐蚀现象，也有人称之为"硫化腐蚀"。这是一种发生在热燃气通道部件上的一种非常严重的表面腐蚀。

热腐蚀与正常的氧化不同，主要与燃料中的杂质（如硫、钠、钒及碳离子等）及随空气一起摄入的海盐、灰尘及水蒸气等复杂的化学作用、热力学作用及流体动力学作用有关。在600～1 000℃温度范围内，热腐蚀剧烈。以镍基高温合金为例，正常氧化在合金表面形成一层薄而致密的 Al_2O_3 保护层，阻止氧化的进一步进行。而热腐蚀形成的氧化膜疏松、无黏附性，其下面还有内部氧化物和硫化物，加速氧化过程的进行；或形成大量的夹杂着金属颗粒和硫化颗粒的疏松而无黏附性的氧化物层，形成灾难性氧化。某些碱类与硫、钒结合特别具有促进腐蚀的作用。

9）金属表面腐蚀的形貌特征　金属材料由于其成分和组织状态及腐蚀介质、腐蚀形式的不同而具有不同的腐蚀形貌特征。按腐蚀后表面形貌特征的不同，一般可分为两种。

① 均匀腐蚀　腐蚀分布在整个金属表面，可能均匀分布，也可能不均匀分布。这类腐蚀的腐蚀程度可用平均腐蚀速度来评定，如天平均腐蚀深度、年平均腐蚀深度来表示，或以每小时（或每天）、每平方米（或平方厘米）金属损失量来表示。

均匀腐蚀容易发现，危害性较小，在设计中容易预防。

② 局部腐蚀　腐蚀只发生在金属表面的个别部位或合金的某一组织上。这类腐蚀包括：

选择性腐蚀：合金中的某一组份或组成物被优先腐蚀。如黄铜发生电化学腐蚀时，锌被优先腐蚀进入溶液，金属表面逐渐成为低锌黄铜。

点状腐蚀：腐蚀表面呈点坑状，腐蚀点多，但深度浅。

穴状腐蚀：腐蚀发生在表面有限的面积上，但深度大，呈蚁巢状。

晶间腐蚀：腐蚀沿晶界发生并扩展。

所有腐蚀表面均可找到腐蚀产物，对腐蚀产物进行分布规律、形态，尤其是成分和数量的分析，有利于找到腐蚀产生的原因、过程。腐蚀产物分析一般通过X射线能谱仪、电子或离子探针先定性分析出组成腐蚀产物的主要元素，然后可用X射线或透射电镜对腐蚀产物的结构进行鉴别分析，为进一步分析腐蚀发生的过程和原因提供依据。

5.3 失效原因诊断技术和方法

对一个一般的失效事件来说，失效原因的诊断往往是失效研究的核心和关键，它对于失效预防的针对性和有效性是重要的前提和基础。

失效原因是指造成失效事故（或事件）的直接根据因素。失效原因也可分为"一级"失效原因和"二级"失效原因。其中，"一级"失效原因的诊断一般是指造成该事故（或事件）的直接根据因素处于该事故肇事件投付使用整个过程中的哪个阶段或工序的分析诊断。一级失效原因常可归纳为：

失效原因
- 设计原因
- 制造原因
- 使用原因
- 环境原因

一般在"一级"失效原因诊断的基础上，应进一步明确其"二级"失效原因，例如，设计原因引起的失效又可分为

设计准则和思想原因、结构和受力原因、选材和性能原因等。

失效原因的诊断一般按以下程序进行：

1）首先进行"一级"和"二级"失效模式的诊断；

2）然后探讨和分析确定是什么原因引起的"二级"失效模式的失效形式，并综合分析全部或主要的失效现象或"规律"；

3）最后进行失效模拟或失效加速试验加以验证。

失效原因的诊断一般是在失效模式诊断的基础上进行的，应将失效模式和失效原因相结合对其失效特征判据进行研究，以达到正确的失效原因诊断的目的。因此，失效原因的诊断在思想方法上也应该与失效模式的诊断一样，从宏观到微观，从定性到定量，从"一级"到"二级"，从诊断（失效-原因）到模拟（原因-失效）。

5.3.1 断裂失效原因诊断技术和方法

按照应力-强度干涉模型，任何断裂失效都是零件所承受的外力超过零件本身所具有的抗力干涉的结果。其中外力是指应力或载荷、腐蚀动力等对材料造成损伤的因素，包括外加载荷、残余应力、腐蚀、磨损等；而抗力是指材料本身所具有的抵抗断裂、磨损和腐蚀的能力，包括材料强度、腐蚀抗力、磨损抗力等指标。

（1）由外力超过抗力引起的断裂失效

由于设计、生产制造、使用条件复杂多样以及人们对客观事物的认识和科学技术水平的限制，出现因设计考虑不周全、分析不透彻、计算不准确的现象在所难免，从而使零件承受的外力大于本身具有的抗力（局部或整体），导致零件断裂失效。由外力超过抗力引起断裂失效一般有以下四种情况。

1）对几何形状复杂的零件，复杂形状部位的应力分布（主要是应力集中）分析不透彻、计算不准确，正常零件局部应力过大而引起断裂失效。

2）对工作条件恶劣、复杂的零件，外载荷的种类有遗漏、大小不准确，导致实际承受的外力超过设计载荷，引起零件断裂失效。

3）对转动零部件，缺乏深入而全面的系统分析，使得零件的自振频率与系统的某一振动频率相耦合，引起共振而造成超载断裂失效。

4）对零件承受的主要载荷类型及大小与选用的材料所具有的主要抗力指标不匹配，造成超载失效。

（2）零件具有的抗力不足而引起的断裂失效

零件本身所具有的抗力不足是实际失效事件中经常遇到的。造成零件失效抗力下降的主要原因有以下几方面。

1）材质冶金方面

① 材料的主要化学成分不符合标准的要求。

② 存在微量有害元素，如 O、N、H、Sb、As、Sn、Bi 等。

③ 显微组织异常或超标，包括基体组织、第二相的数量、大小和分布，析出相的组分、大小和分布，晶粒度及残余奥氏体等。

④ 非金属夹杂物的种类、数量、大小及分布等超标。

⑤ 存在超标冶金缺陷，包括疏松、偏析、气孔、夹杂等。

⑥ 流线分布不合理。

⑦ 表面或内部存在宏观裂纹或显微裂纹。

2）表面完整性方面　表面完整性不符合要求，或在使用中被破坏，造成零件的力学性能、物理性能与化学性能（即失效抗力）下降，从而诱发裂纹在这些部位萌生。表面完整性包括表面粗糙度、表面防护层的致密性、完整性及外界因素造成的机械损伤等。

3）表面残余应力的类型、大小与分布　残余应力往往与外加应力叠加而提高零件局部的应力水平，诱发裂纹萌生，促进断裂失效。一般残余拉应力是有害的，促进裂纹产生，因此应该尽量避免；残余压应力能提高疲劳断裂寿命，降低应力腐蚀敏感性，是有利的。对于化学、电化学腐蚀，无论是残余拉应力还是残余压应力均促进腐蚀的发展，都是有害的。

4）零件几何形状设计不当　由于零件的几何形状设计不当，导致零件的应力分布不均匀，局部区域应力集中严重，应力水平大幅升高，是零件的承受实际载荷的能力降低。疲劳裂纹大多起源于零件的尖角、倒角、孔边、圆弧过渡处等部位就是例证。

5）零件的加工质量不符合要求　主要包括几何尺寸不符合要求和表面完整性不符合要求，局部区域承受大的应力集中，导致零件承受实际载荷的能力下降。

6）温度的影响　温度升高，材料的屈服强度、断裂强度，疲劳抗力、蠕变抗力等抗力指标下降，降低材料本身的失效抗力。温度的急剧变化还会引起零件的抗热疲劳能力下降。

在一定的温度，具有韧-脆转变现象的材料会发生延性与脆性之间的相互转变，低于此温度，会引起材料的低温脆性断裂，材料的断裂失效抗力大幅度下降。

7）环境介质的影响　环境介质会促进材料的吸氢、加速腐蚀及与腐蚀有关的失效行为，从而使零件的氢脆、腐蚀、应力腐蚀、腐蚀疲劳等失效抗力下降。

总之，断裂失效的原因是错综复杂的，诊断起来难度较大，上面叙述的仅是一般的原则和方法。在实际的失效分析中，失效的原因在大的原则方面还是应该从动力和抗力的影响因素方面来查找。事实上，影响零件失效的动力和抗力的因素很多，难以穷尽，但从零件的全寿命使用过程来看，不外乎设计、制造、使用和环境影响等几方面。

失效原因的诊断牵涉到很多方面，一般都相当复杂，因此，在诊断时，应按以下程序和要求进行。

1）失效模式诊断要准确，且越细越好。

失效模式诊断是整个失效诊断的基础和前提，对失效原因诊断具有导向作用。不同的失效模式对应着不同的失效原因，因此，只有失效模式诊断准确，失效原因的诊断才不会走弯路。失效模式诊断越细致、具体，为失效原因诊断的方向就指的越明确、具体。表 5.2-14 列出了各种常见的失效模式对应的可能失效原因。

2）调查掌握使用工况，并根据使用工况进行肯定或排除失效原因。

3）调查设计资料，掌握设计思路和要点。

4）调查制造工艺过程，排除或肯定失效原因。

5）综合分析，确定失效原因。

6）提出改进措施和建议。

（1）设计原因诊断

即使经过仔细构思和周密计算、试验过的设计，有时也不可避免地存在设计缺陷。由于设计缺陷导致的失效一般都属于早期失效。

导致失效的设计原因一般有以下几种。

1）设计载荷不准确。主要表现为载荷考虑不全（应力估算不足），载荷变动分析不够，计算假定中出现误差，致使实际工作载荷超过设计载荷，材料在正常工况下也无法承受工作载荷而失效。如发动机叶片设计时没有考虑其在某一转速下发生共振，致使叶片使用中发生共振疲劳失效。

表 5.2-14　外力与抗力分析示意表

<table>
<tr><th>失效模式</th><th>抗力或外力原因</th><th colspan="3">具体原因</th></tr>
<tr><td rowspan="19">断裂失效的模式（属性）</td><td rowspan="11">零件具有的失效抗力低</td><td rowspan="6">冶金因素</td><td colspan="2">化学成分不合格</td></tr>
<tr><td colspan="2">组织结构不合格</td></tr>
<tr><td colspan="2">力学性能不合格</td></tr>
<tr><td colspan="2">物理性能不合格</td></tr>
<tr><td colspan="2">化学性能不合格</td></tr>
<tr><td colspan="2">内部存在缺陷</td></tr>
<tr><td rowspan="3">表面因素</td><td colspan="2">表面完整性不符合要求</td></tr>
<tr><td colspan="2">表面存在残余应力</td></tr>
<tr><td colspan="2">零件几何形状、尺寸不符合要求</td></tr>
<tr><td rowspan="2">环境因素</td><td colspan="2">环境温度超过使用限制（过高或过低）</td></tr>
<tr><td colspan="2">环境介质具有腐蚀性</td></tr>
<tr><td rowspan="8">零件承受的外力超过限制</td><td rowspan="2">正常载荷</td><td colspan="2">类型与大小</td></tr>
<tr><td colspan="2">频率与振幅</td></tr>
<tr><td rowspan="6">非正常载荷</td><td rowspan="2">一次性过载</td><td>动载荷</td></tr>
<tr><td>静载荷</td></tr>
<tr><td rowspan="4">非正常振动</td><td>共振</td></tr>
<tr><td>颤振</td></tr>
<tr><td>喘振</td></tr>
<tr><td>随机振动</td></tr>
</table>

对航空发动机转子零部件来说，设计载荷不准确可能是导致其失效的设计原因中的最主要原因。具体可表现为没有考虑振动应力或考虑振动应力过小；没有考虑气动应力，或将其考虑太小；温度场分布不均导致的热应力没有考虑或考虑太小。另外，应力分析中对模型的相关简化不合理也是一个主要的方面，需要引起重视。

2）设计结构不合理。主要表现为零件几何形状设计不当，出现剖面突变或尖角面，导致这些部位承受较大的应力集中，超过材料的强度极限，从而过早地萌生裂纹而失效。如飞机3A21O防锈铝液压导管使用仅几个小时就在扩口处产生对称破裂漏油失效，其根本原因就是结构不合理，使扩口处局部处于应变状态下工作，加上导管的振动，导致了扩口处在短期内即疲劳破裂失效。

3）设计选材不当。主要表现在材料性能不足，状态不对或不相容材料相互接触、不相容材料状态相互接触、匹配零件材料线胀系数相差太大等方面。如汽车转向机齿条在北方寒冷冬天的一次轻微撞击中发生脆性断裂，其根本原因是选用的材料的韧脆转变温度偏高，不适应寒冷天气条件，只要受到轻微的撞击力就会断裂。

在实际失效分析中，设计选材不当导致断裂失效，经常遇到的是所选材料的主要抗力指标与实际损伤模式不符合，致使材料的性能指标在该强的方面不强，不能满足使用的需求。如某型航空发动机涡轮叶片选用具有良好高温性能和蠕变抗力的定向凝固合金，但中温疲劳性能相对较弱，断裂韧性也较低，且分散性大，设计中选用该材料是希望发挥其良好的高温性能和抗蠕变性能，在使用中，该叶片多次发生疲劳断裂失效。尽管存在其他一些设计、制造原因，但究其原因，在设计选材上是有缺陷的。因为尽管该材料具有良好的高温强度和抗蠕变性能，但实际该叶片的使用温度不高，属于中高温范围，蠕变失效不是其主要的失效模式，因此，并不需要其具有很高的高温性能和抗蠕变性能。实践证明，该叶片的主要失效模式为疲劳断裂，要求材料具有良好的中高温疲劳抗力，而这恰恰是该材料的弱点，因此，在使用中多次发生疲劳断裂失效就在所难免了。这就是典型的所选材料的抗力指标与实际的失效模式不符合的例子。

设计原因引起的失效，失效原因的诊断往往比较困难，特别是在该类零件首次失效时经常容易被忽视或没有考虑到，而且当同类零件已有大批量使用时，人们往往难以接受，有时甚至需同一失效模式发生多次以后，人们才会怀疑其在设计上是否存在问题。这主要是由于设计上的问题分析起来较复杂，工作量大，时间长，而失效原因诊断往往又要求尽快做出结论，从时间上不允许；其次是设计上的问题处理起来难度大，一般在管理上也不愿意接受；更重要的是，设计原因不像制造原因，在失效件上一般难以找到导致失效的直接证据，相反，在由设计原因导致的首次失效事件中，失效件上恰恰还经常可能存在有制造的缺陷。鉴于设计原因诊断的复杂性和困难性，在失效诊断中，对设计上是否存在问题，我们既不能随便怀疑，更不能轻易否定，否则，将可能导致错误的结论。但一般来说，由设计原因引起的断裂失效虽然难以在失效件上找到直接证据，但在失效模式上也还是会有所反映的，如叶片等转动部件，高周疲劳失效就意味着存在共振的可能。因此，在失效原因分析中应紧紧依据失效件表现出的失效模式，充分发挥失效模式的“导向”作用，这样可以少走弯路。

（2）制造原因诊断

在零件的批量生产中，由于人员的变动、设备的老化、偶然的失误，不可避免地将带来一些制造缺陷，并最终影响到零件的失效。

在断裂失效中，导致失效的制造原因主要表现在材质缺陷和制造工艺缺陷两方面。它们通过降低材料的抗力指标或引起零件局部应力集中，来为失效作贡献。

表5.2-15列出了材料的主要抗失效指标。凡是降低材料抗失效指标的缺陷均可对零件的失效做出贡献。

表 5.2-15　机械材料抗失效性能指标

抗失效指标类别	抗失效指标名称	符号	单位
抗弹性变形失效指标	拉伸弹性模量	E	GPa
	弹性极限	σ_e	MPa
	比例极限	σ_p	MPa
	泊松比	μ	
	切变模量	G	GPa
抗塑性变形失效指标	屈服点或屈服强度	σ_s 或 $\sigma_{0.2}$	MPa
	剪切屈服极限	τ_s	MPa
	扭转屈服极限	τ_s	MPa
抗断裂失效指标	抗拉强度	σ_b	MPa
	抗扭强度	τ_b	MPa
	冲击韧度	a_k	kJ/m^2
	抗剪强度	τ_b	MPa
抗疲劳失效指标	对称循环疲劳极限	σ_{-1}	MPa
	疲劳裂纹扩展速率	da/dN	
	扭转疲劳强度极限	τ_{-1}	MPa
抗环境失效指标	蠕变极限	σ_ε^t/T	MPa
	持久强度极限	σ_T^t	MPa
抗温度失效指标	无延性转变温度	NDT	℃
	断口形貌转变温度	FATT	℃
	弹性断裂转变温度	FTE	℃
	塑性断裂转变温度	FTP	℃
	熔点	T_m	K
抗失效断裂韧性指标	弹塑性断裂韧性	J_{Ic}	kJ/m^2
	平面应变断裂韧性	K_{Ic}	$MPa\cdot m^{1/2}$
抗失效缺口效应指标	材料缺口敏感系数	Q	
	高温缺口敏感系数	β_δ	
	缺口对称循环疲劳极限	σ_{-1n}	
	应力集中系数	K_t	

1）材料缺陷 材料纯洁度不符合要求，是评定金属材料中夹杂物含量和分布的质量指标。金属中的夹杂物是金属基体的均匀性、连续性受到破坏，其形态、性质、含量和分布都不同程度地影响材料的各项抗失效性能指标。

组织稳定性不好。长期使用组织发生变化，将直接影响到力学性能的衰减。塑性、冲击韧性、缺口敏感性、热疲劳抗力等，都是对长时间使用组织稳定性较为敏感的性能。

表5.2-16列出了常见的一些材质缺陷。

2）工艺缺陷 与制造工艺有关的常见缺陷见表5.2-17。

制造原因导致的断裂，一般在失效件上均可找到相应的证据，但需进行仔细的相关分析。另外，材料的性能分散性大，也可能使部分零部件的相关性能落入与外力的干涉区，从而导致失效发生。

（3）使用原因诊断

表5.2-16 常见材质缺陷及产生原因

缺陷名称	缺陷特征	产生原因
裂纹	一般呈直线状，有时呈Y形	坯料上有裂纹或皮下气泡、夹杂物
气泡	表面无规律地分布、呈圆形的大大小小的凸包，其外缘比较圆滑；大部分气泡是鼓起的，也有的不鼓起；经酸洗平整后表面发亮，其剪切断面有分层	炼钢时沸腾不好，出气不良，使钢锭、钢坯的内部产生严重气泡，经多次压力加工没有焊合。沸腾钢浇铸温度过低，浇铸速度太快，使气体没有排出的机会，形成气泡的数量过多，尺寸过大，轧制时未能消除
分层	钢材截面上有局部的明显的金属结构分离，严重时则分成2~3层，层与层之间有肉眼可见的夹杂物	钢锭开坯时，缩孔未切净，使坯上带有缩孔残余 钢锭中的气泡在轧制中未被焊合 钢锭中有集中的夹杂物 化学成分严重偏析也能形成分层
发纹	深度甚浅、宽度极小的发状细纹，一般沿轧制方向排列，长短不一，一般在20~30 μm以下，个别可达100~150 mm，呈分散和链状排列，或成簇分布	坯料上的皮下气泡、非金属夹杂物在轧制中未焊合，多见于钢锭下部锻轧的钢材上，当切削或侵蚀后暴露于表面
麻点（麻面）	表面呈现局部的或连续的成片粗糙面，分布着形状不一、大小不同的凹坑，严重时有类似橘皮状、比麻点大而深的麻斑	解热过程中钢材表面氧化严重，轧制时，氧化铁皮成片状或块状压入。在轧制过程中或酸洗后脱落，形成细坑，常称为氧化麻点 在轧制或热处理的加热过程中，由于煤气中的焦油喷到板面上所腐蚀的小坑称为焦油麻点 加热过程中被某种气体腐蚀形成气体腐蚀麻点；酸洗过度产生酸洗麻点；轧辊磨损严重造成麻点
辊印	表面有带状或片状的周期性轧辊印，其压印部位较亮，且没有明显的凸凹感觉	轧辊材质不良，硬度和粗糙度不够 轧件强度高，轧制中造成搭铁，使辊子产生烙印 平整机调整不当
开裂	钢管表面出现呈穿透管壁的纵向裂开，一般发生在全长，有时发生在一端	退火不当，温度不均，延伸不一 压下量过大，加工硬化严重 拔后未及时退火，应力未消除，含碳量较高的钢管易开裂
疏松	钢的不致密性表现。多出现在钢锭的上部与中心，这些部位因为集中了较多的杂质和孔隙，切片经过酸洗侵蚀以后，扩大成许多洞穴。根据其分布可分为一般疏松和中心疏松	主要是钢中的杂质和气体造成。当钢液凝固时，由于体积的收缩，形成树枝状的晶间空隙，以及凝固过程中气体上浮而构成的显微空隙没有为钢液填充，构成了组织的不致密性
偏析	根据其表现形式可分为树枝状、方框形、点状、偏析和反偏析等	钢锭结晶过程的产物，主要由钢的成分不均匀及其杂质构成
缩孔残余	在横向酸浸试片的中心部位，呈现不规则的空洞或裂纹，空洞或裂纹中往往残留有外来杂质	钢锭切头不够而残留 缩孔深及锭身产生二次缩孔
非金属夹杂	在横向酸性试片上见到的一些无金属光泽，呈灰白、米黄和暗灰色的物质	冶炼时钢中气体与煤渣互相作用而形成的各种氧化物 冶炼时混入钢中的耐火材料
金属夹杂物	横向低倍试片上呈现的具有金属光泽、与基体金属显然不同的金属盐类	出钢与浇注时，杂物混入钢液中 炉渣不能及时浮出
白点	钢的内部破裂的一种。在钢的纵向断口上呈圆形或椭圆形的银白色斑点，经过磨光和酸蚀后的横向切片上，表现为细长的发裂；有时呈辐射状分布，有时则平行于变形方向或无规则分布	钢中含有氢气 钢中有应力 冷却时缓冷不好等
晶粒粗大	酸浸试片断口上有强烈金属光泽	主要是热加工时加热温度与时间控制不当造成的，如过热、过烧等 加热温度越高，高温停留时间越长，晶粒越粗大，对钢的力学性能有很大影响，降低塑性和韧性

续表 5.2-16

缺陷名称	缺陷特征	产生原因
脱碳	钢的表层碳份较内层碳分降低的现象 全脱碳层是指钢的表面因脱碳而全部为铁素体组织部分 部分脱碳是指全脱碳层到钢的含碳量正常部分之间的组织处	主要反映氢和氧对金属的作用，使钢中的碳与氢或氧作用生成甲烷或一氧化碳；其次是氧化作用，氮化铁先氧化产生脱碳
折叠	沿轧制方向呈直线状或曲线状，有的呈锯齿状或翘起，深浅不等，有连续的或不连续的，外形与裂纹相似。在横断面上一般呈折角	坯料表面带有折叠和夹杂物 坯料表面带有严重擦伤和裂纹，经拔制、轧制后延伸互相折合 坯料修磨处带有棱角，再拔制时产生折叠
软硬不均	制品上各处硬度不一	主要是加热时制品受热不均

表 5.2-17 常见制造工艺缺陷

工艺过程	缺陷名称	常存在的零件
铸造	冷隔或浇注不足、缩陷、裂纹、表面粗糙、气泡、变形、夹杂、偏芯或断芯	熔模铸件
	化学黏砂、机械黏砂、热力黏砂、表面粗糙、夹砂、裂纹、缩孔（疏松）、渣孔、砂眼、变形、气孔、冷隔和欠铸、氧化斑疤、金属豆、皱纹、氧化和脱碳	熔模精密铸件
	疏松和缩孔、冷隔和欠铸、裂纹	金属型铸件
	气孔、疏松和缩孔、裂纹、花纹（网纹、条纹）、冷隔、夹渣、飞边、流痕	压铸件
	铸瘤、鼓包、披缝、表面粗糙	石膏型铸件
	披缝毛刺、黏砂、气孔、疏松和缩孔、表面皱纹、跑火、夹渣、变形	壳型铸件
	气孔（气穴、气孔、气窗）、缩孔（内缩孔、外缩孔）、疏松、夹渣（渣孔、氧化夹渣）、黏砂、表面结疤、表面麻点、冷隔	低压铸件
	裂痕、气孔	陶瓷型铸件
	气孔、缩孔、心部疏松带过宽	真空吸铸
	疏松、缩孔、内部气孔、表面裂纹、跑火、鼓胀、表面针孔、夹砂（起皮）、夹杂、变形、冷隔、流痕、毛刺	合金铸件
	疏松、氧化夹渣、熔剂夹渣、裂纹、气孔、冷隔、欠铸、铸件燃烧、针孔、比重偏析、共晶偏析	铝镁合金铸件
	气孔、缩孔（疏松）、夹渣、浇不足和冷隔、裂纹、偏析	铜合金铸件
	气孔、缩孔、欠铸、凹坑、裂纹、变形、剥落、毛面皮	铸造永磁合金
	浇不到与冷隔、缩孔或缩松、夹渣、硬点、变形、砂眼、气孔、裂纹、夹砂、黏砂、抬箱、错箱、冷豆（铁豆）、成分组织及性能不合格	灰口铸铁
	缩孔或缩松、夹渣、石墨飘浮、皮下气孔、球化不良	球墨铸铁
	气孔、夹砂、结疤、冲砂、掉砂、砂眼、渣孔、黏砂、冷隔、皱纹、缩孔、热裂、冷裂、偏析	铸钢
	柱状件生长不均匀、柱状件粗大、横向晶界、柱状晶偏离、断晶	定向凝固叶片
锻造	毛细裂纹、折叠、结疤、层状断口、亮线（亮区）、非金属夹杂、碳化物偏析、铝合金氧化膜、异金属夹杂物、白点、粗晶环、缩管残余	原材料缺陷引起
	切斜、坯料端部弯曲并带毛刺、端部金属被拉掉、端部裂纹、气割裂纹、凸芯开裂	坯料剪切和切割时产生
	过热、过烧、铜脆、加热裂纹、石状断口、萘状断口、脱碳、增碳、加热不足引起心部开裂、铝合金锻件表面气泡	加热不当产生
	大晶粒、晶粒不均匀、冷硬现象、脱碳层堆积、十字裂纹、龟裂、飞边裂纹、分模面裂纹、孔边龟裂、裂纹、锻造折叠、穿流、锻件流线分布不当、带状组织、剪切带、碳化物偏析级别不符合要求、铸造组织残留、铜合金锻件应力腐蚀开裂、冷却裂纹、网状碳化物	锻造工艺不当产生
切削加工	尺寸超差、表面粗糙度差、毛刺、划伤、啃刀、表面烧伤、裂纹、刀痕	切削加工各过程
冷加工	破裂、拉穿、波浪形、折皱、横向裂口	冷冲拉加工过程
	划痕、锈蚀、球化退火不足、带状组织、晶粒粗大或粗细不均、成分不合格、性能不合格、冲模错位	冷冲压加工过程
	裂纹、压痕、端面鼓起或不平、存在挠度、曲度	冷弯加工过程
热处理	硬度过高、过共析钢网状碳化物、石墨化、压共析钢魏氏组织、铁素体晶粒粗大	退火和正火缺陷

续表 5.2-17

工艺过程	缺陷名称	常存在的零件
热处理	硬度不足、过热或过烧、淬火开裂、淬火软点、淬火不完全、表面脱碳、表面腐蚀	淬火缺陷
	回火脆性、表面腐蚀	回火缺陷
	过热、过烧、氧化、脱碳、机械性能不合格、软点、变形与裂纹	钢的热处理缺陷
	表面硬度不足、表面网状或粗大块状、爪状碳化物、氮化物、渗层过深、不足或不均匀、表面网状屈氏体组织、黑色斑点状组织、芯部硬度超差、零件变形与开裂、渗层脆性和剥落	渗氮（碳氮共渗）缺陷
	石墨碳、球化组织不良、网状碳化物、淬火脱碳、硬度不足、裂纹、淬火过热或过烧、腐蚀、软点、茶断口	工模具钢热处理
	硬度不足、弹性低、脆性大、变形、淬火开裂、表面脱碳或元素贫化、表面腐蚀	弹簧钢
	硬度偏差、裂纹、点蚀、晶间腐蚀、应力腐蚀	不锈钢
	过热、氢脆、氧化、热裂、晶粒大小不均、硬度不足、气泡、表面斑点、腐蚀	铜合金
	过热、氢脆、氧化	钛合金
	力学性能不合格、过烧、气泡、变形与裂纹、腐蚀、粗晶、腐蚀	变形铝合金
	氧化、力学性能不合格、晶粒长大、过烧、变形与裂纹	镁合金
	固溶不完全、不均匀，过烧，翘曲、裂纹、严重氧化、烧点，性能不合格	铝镁合金铸件
	晶间氧化、表面成分变化、腐蚀点和腐蚀坑、氧化剥落、翘曲变形、裂纹、粗晶或混合晶粒、过烧和过热、硬度不合格	高温合金
	变形、氧化、粘结、污损、磁性能受损	软磁合金
	表面不光亮、合金元素表面贫化、淬火硬度不足、瞬时渗碳、表面晶粒粗大、高速钢表面白层、微渗氮、粘连	真空热处理
表面处理	镀层发脆、麻点、起泡、发黑、无光泽	氰化镀锌
	镀层色暗、粗糙无光泽	无氰化镀锌
	镀层结合力差、起泡、发脆、脱落、发暗、麻点	氰化镀镉
	镀层粗糙、疏松、结合力差、起泡、脱落、呈暗红并析氢、毛刺	镀铜
	镀层有针孔、毛刺、晶粒粗大、黑色条纹、烧焦、粗糙、麻点、起皮或剥落	镀镍
	镀层厚度不足、龟裂、剥落、麻点	镀铬
	镀层粗糙、呈海绵状结晶、树枝状结晶、脱落、疏松多孔、不均匀	镀锡
	镀层结合力差、起泡、粗糙、厚度不均匀、斑点或黑色条纹、锈点	镀银
	镀层发黑、结合力差、脱皮、粗糙	镀铂
	表面烧伤、氧化膜疏松粉化、泡沫状或网状花纹、黑斑、腐蚀	硫酸阳极化
	表面烧伤、腐蚀坑、氧化膜粉化、发红、发黑	铬酸阳极化
	磷化膜不完整、结晶过细、粗大、不均匀、脱落	黑色金属磷化
焊接装配	焊缝尺寸偏差、咬边、焊瘤、弧坑、烧穿、焊漏、气孔夹渣、未焊透、裂纹	熔焊件
	焊点及焊缝位置不正、形状不正、压痕过深、过热、烧伤、裂纹、未焊透或核心小、缩孔、内部飞溅、熔透过大、接头变脆、接头不气密	点焊、缝焊
	间隙未填满、气孔、钎缝表面粗糙、夹杂物、裂纹、钎料侵蚀金属表面	钎焊

导致断裂失效的使用原因一般表现为超过设计限制使用和维护不当、造成损伤。

1）超设计限制使用原因　按照应力－强度干涉模型，失效是零件所受的外力超过材料的相应强度的结果。使用限制是根据材料的强度、一定的安全储备及零件工况可导致的应力制定的。如果超过设计限制使用，必将使零件所承受的应力升高，就有可能使零件承受的应力超过材料的强度而使零件失效。若存在应力集中、有害残余应力，这种现象会更加明显。

超过设计限制使用原因导致的失效，一般可通过工作参数分析、应力校核途径来分析诊断。现在已有研究试图通过对断口特征的应力反推来确定失效零件实际所受的应力水平，并据此来推断是否超载使用。该方法应该说比较客观，但由于目前的断口应力反推还存在误差大、实际操作困难的问题，因此，在实际分析中还仅能作为参考。

与超限制使用原因有关的失效模式主要有：一次性延性断裂、低周大应力疲劳断裂、蠕变断裂等。

2）使用维护原因　使用维护的目的是保证零件能够正常工作，防止发生意外损伤。如转动部件的润滑油加注、损伤部位的检查修复等。由于维护一般需要对零件进行拆装，就有可能在重新装配时，其状态与制造状态不一致，并带来额外的损伤、附加的应力等，从而导致零件的失效。如飞机导管在维护安装时，固定卡子没拧紧，致使导管振动，产生疲劳断裂失效就是一个典型的例子。

转动部件维护中没有及时加注润滑油、润滑油中有杂质、冷却系统中缺少冷却水（液）、零件表面黏附有腐蚀介质等都是常见的维护不当的原因。

由于使用中维护不当导致的失效，一般在失效件上或失效系统中均可找到一定的证据，如转动件缺油导致磨损失效，零件上会有高温变色；润滑油中带进外来物，会在油滤上见到外来物，在磨损件表面发现相应的划痕等。

(4) 环境原因诊断

机械失效中的环境原因主要有两方面，一是腐蚀介质作用，二是温度效应。相应地，这些原因在失效模式中都会有所反映，因此，在原因诊断时应根据失效模式指明的方向，根据现有理论与实际经验，有所侧重地进行。

1) 腐蚀介质原因　腐蚀介质可使材料与环境之间发生有害的化学或电化学作用，引起表面腐蚀损伤或材质脆化，裂纹在损伤处萌生或加速扩展，导致失效。

金属零件常见的腐蚀介质引起的断裂失效主要有：腐蚀疲劳失效、氢致脆断失效、应力腐蚀断裂失效、低熔点金属侵蚀致脆失效。在这些失效中，有的裂纹直接起源于腐蚀损伤处，如应力腐蚀断裂失效、腐蚀疲劳失效；有的是腐蚀介质加速裂纹的扩展，如腐蚀疲劳失效、应力腐蚀失效；有的是腐蚀介质使材料脆化，如氢致脆断失效、低熔点金属侵蚀致脆失效。在失效原因诊断中，一般可根据失效模式及失效件上相关的失效信息，如腐蚀产物的结构、成分、使用环境条件等诊断出失效原因，如腐蚀介质的种类和来源等。

2) 环境温度作用　环境温度的作用主要是降低材料的强度性能或增加材料的脆性。

环境温度导致的失效类型主要有：高温延性断裂失效、高温疲劳失效、蠕变破断失效和低温致脆失效。其中，温度增加一般使材料流变强度、断裂强度和弹性模量等与强度、刚度抗力有关的指标下降，韧性增加，可导致材料延性断裂和蠕变断裂失效（拉伸强度下降引起）、疲劳失效（疲劳强度下降）。相反，温度降低，材料的韧性一般会降低，当温度降低至材料的韧－脆转变温度以下，可使材料脆化，导致低应力脆断失效。

5.3.2　磨损失效原因诊断技术和方法

磨损是造成材料失效的三大失效模式之一。材料的耐磨性与力学性能、物理性能和化学性能不同，不属于材料固有的性能，而是受到摩擦学系统一系列条件、工况、环境、介质等多方面的影响，而成为一个系统的性能。研究磨损失效的原因，对采取有效的、合适的方法和工艺措施来防止和减少磨损具有重要的意义。

导致材料磨损失效的原因很多，但归纳起来不外乎设计、制造、使用和环境四大方面。

(1) 设计原因诊断

对磨损系统没有设计适当的冷却与润滑、冷却与润滑不良、摩擦副的匹配性不好、受力过大等是磨损失效原因的重要表现。有摩擦就有热量的产生，就会导致摩擦副温度升高，材料性能下降，润滑剂变质失效，加速磨损，因此都要进行适当的冷却，使摩擦副的温度不会升高或升高不大。其次，对摩擦副进行适当的润滑也是必须的，在零件表面形成的润滑油膜，可显著降低零件的磨损程度。但有时由于设计结构上的不合理，使磨损部位得不到合适的冷却与润滑，必将导致零件的磨损失效。摩擦副材料的匹配性不好，相互磨擦的两个零件的性能相差较大，可使磨损集中在耐磨性差的一侧，并有加速的可能，使其快速磨损失效。摩擦副设计承受的纵向压应力过大，可使摩擦表面无法形成完整的润滑膜或使已形成的润滑膜破裂、材料产生塑性变形等因素，不利于减少磨损。

(2) 制造原因诊断

磨损失效的制造原因主要表现在零件的表面粗糙度大、材料的强－韧性匹配不好、导致耐磨性差以及结构尺寸不符合要求。

磨损是一种表面损伤，磨损面粗糙度大，可使磨损很容易发生。由于对一般金属来说，其强度与韧性具有相反的变化规律，随着强度的升高，韧性会下降，反之，适当降低材料的强度，可使韧性得到提高。如果不考虑其他因素，单纯从材料性能来说，合适的强度与韧性的匹配，可使材料具有良好的耐磨性能。过高的强度或过低的强度都会使材料的耐磨性下降。零件的结构尺寸不合格，可使局部承受的摩擦应力增大，导致磨损加剧。

(3) 使用原因诊断

磨损零件的使用条件主要是指其运转条件，包括载荷、速度、润滑及运转周期（连续、间歇）。

超过使用限制的运转方式，如承受的外在载荷过大、超转等，都可加剧磨损程度，促进磨损失效。使用中的润滑油加错、加注不足等都可能导致摩擦副的润滑不良，促进磨损的发生。另外，运转周期是连续运转还是间歇运转对磨损也有较大的影响。

(4) 环境原因诊断

影响磨损失效的环境条件主要有潮湿、腐蚀性气体、粉尘、高温、低温、辐照和真空等。水分及腐蚀气氛会促使零件腐蚀、润滑剂变质，从而加剧磨损；高温及辐照会促进润滑剂老化分解而具有酸性或使材料的性能改变；真空条件影响表面润滑膜的形成，也对减少磨损不利。粉尘等进入摩擦系统，会在摩擦副中起到磨料的作用，从而加速磨损。低温环境会使材料变脆，容易使材料产生剥落，加速磨损失效。

5.3.3　腐蚀失效原因诊断技术和方法

腐蚀是造成材料失效的三大失效模式之一。由于腐蚀是材料与周围介质发生化学及电化学作用的结果，因此，腐蚀原因都与腐蚀介质及其接触反应有关。应该说影响材料发生腐蚀的原因很多，难以一一例举，但与断裂、磨损等失效模式一样，同样可以从设计、制造、使用和环境四个方面来进行分析。金属的腐蚀一般可分为化学腐蚀和电化学腐蚀，但常见的主要是电化学腐蚀，因此，下面重点对电化学腐蚀的原因进行讨论。

(1) 设计原因诊断

从电化学腐蚀的机理可以知道，金属零件发生电化学腐蚀必须具备三个基本条件：

① 两种金属或合金中的两个不同区域甚至两相之间的电极电位不同；

② 使这两种金属或合金中的两个不同区域相互接触或用金属导线将其连接起来；

③ 存在并都在同一电解液中。

显然，条件①、②是难以避免和相当容易满足的，即使没有异金属接触，同一金属和合金内部的相间也存在电位差；同时，在潮湿的大气环境中，电解液也是自然存在的，因此预防腐蚀的关键是将腐蚀介质与金属基体隔离开，不使其接触，从而避免腐蚀的产生。

在零件的设计中，影响腐蚀的原因主要表现在零件的防腐蚀品质不高、结构上存在积水部位和异种金属接触等。

1) 零件的防腐蚀品质原因　零件的防腐蚀品质包括材料本身及其表面防护层的防腐蚀能力。对于在自然环境中，特别是在高温潮湿环境中使用的零件，不但要选用对相应条件下的腐蚀不敏感的材料（但一般选材首先考虑的是强度上的要求），而且要采用相应的防护涂层，如表面阳极化、涂漆、涂胶等，从而将腐蚀介质与基体材料分隔开，以防止腐蚀的产生。如果零件的防腐蚀品质不高，防护层容易破坏，则腐蚀将不可避免地会发生。

2）结构设计上不合理，存在有积水部位　如果腐蚀介质在零件的局部区域残留、积聚，一旦防护层损坏，腐蚀将会迅速产生和发展。因此，在结构上应设计相应的排水措施，不让介质积聚。

3）在结构上，异类金属存在接触　应该说在现代设计中，大的结构件，一般不会发生异类金属接触问题，但对一些细节问题，这样的事情就经常发生，如螺钉孔处、飞机结构件上的搭铁线处、铆接处等，经常发生不同电位金属间的接触，并且没有采用相应的防护措施，从而导致严重的腐蚀。

（2）制造原因诊断

腐蚀失效的制造原因主要表现在防护层不符合要求及装配时破坏了防护层。防护层不符合要求，降低了零件的防腐蚀品质，可使腐蚀过早发生。如果防护层遭到破坏，则使零件失去应有的保护，只要腐蚀介质存在，就会发生腐蚀。

（3）使用原因诊断

腐蚀失效的使用原因主要是零件表面的防护层在使用中遭到破坏和意外的腐蚀介质侵入。前者同制造时防护层破坏一样，可使零件失去应有的防护，引起腐蚀；后者可能超过零件防护层的防护能力，导致防护层失效，基体材料被腐蚀。

（4）环境原因诊断

一般情况下，导致腐蚀失效的腐蚀介质均来自使用环境。由于环境的污染，大气中都不可避免地存在如 CO、SO_2、HCl 等具有较强腐蚀性的气氛，当零件表面存在水膜时，这些有害气氛就会溶入其中，形成电解液，腐蚀基体金属。在高温、潮湿的环境、污染严重的地区，环境介质的腐蚀性越强，这种腐蚀越严重。

另外，工作在特殊环境条件下的零件，由于工作环境条件的恶劣，也极容易产生腐蚀。如工作在高温燃气中的发动机涡轮叶片，经常会发生热腐蚀（高温氧化）；工作在酸、碱性环境中的材料，也经常容易发生腐蚀。

5.4　失效机理诊断技术和方法

一般来说，一个失效事故是外因和内因共同作用的结果。常把应力、环境（温度、大气）、时间、缺陷等作为产生失效的外因，而失效的内因 - 导致发生失效的零件或材料的物理、化学或机械过程等称为失效机理。拿生病来比喻，失效模式相当于病症，失效机理则相当于病理。失效机理研究是对失效的深层次的内因或内在本质，即导致失效的必然性和规律性的研究，是对失效性质认识的理论提高和升华。

由于失效有必然性和偶然性，失效的必然性寓于偶然性之中。失效的必然性是研究失效的热力学、动力学和机构学，即研究失效的内在规律性，或研究导致失效事故的诱发因素或事故链。对一个完整的失效机理研究来说，就是要在失效必然性认识的基础上，揭示失效的特殊性和偶然性，因此，对具体失效事件特殊性的分析往往是一个失效分析诊断的关键和核心所在。只有找到了失效的特殊性，才能使失效分析具有针对性、失效预防才能做到有效性，但如果只停留在失效的特殊性、偶然性的认识上，就不能揭示失效的必然性，就不能真正做到“举一反三”和“亡羊补牢”，因此，二者缺一不可。

尽管失效机理诊断对揭示失效的内在本质规律以及从根本上预防同类事故的再次发生具有重要的作用，但在实际的失效事件研究中，失效机理的诊断往往不为大家所重视，经常被忽视，表现在失效研究中，就是一般均不进行失效机理的诊断，而大都只停留在失效的特殊性和偶然性上，特别是当某一产品的某一故障首次发生时，由于一般均存在有一定的缺陷，因此，失效就停留在对缺陷存在这一偶然性上，而对失效的深层次原因（必然性）没有注意到，从而使失效机理诊断的工作没有很好地进行下去，这样的事例可以列出很多。这也是有的事故多次重复发生的主要原因。

失效机理的研究方法一般分为定性和定量两大类。在定量的研究方法中又可分为确定型的定量方法和概率型的定量方法。在以往的某些重大失效事故的分析中，曾发生缺乏定量性，而定性又存在片面性，造成误判或错判的问题。一般来说，一个成型产品的小概率的或长寿命的偶然失效事故的原因或机理往往与工艺或使用中的“偶然”差错或因素有关；而一个试制品的、或重复出现的大概率的失效事故，总是与设计结构等因素有关，因此，在失效机理诊断中应根据失效概率的大小而有所侧重。

在目前的研究中，失效机理的诊断定性研究的多，而定量和数理统计模型方法应用的少；在定性研究中，大多也只是根据断口上观察到的一些微观特征，然后结合材料的一些规律，从微观上进行定性的分析研究，缺乏应有的深度和广度。

5.4.1　微观形貌诊断技术

断裂失效机理的微观形貌诊断一般是通过对失效件的形貌特征进行微观分析来进行的。金属材料断裂失效的微观机理主要有韧窝断裂、滑移分离、解理和准解理断裂、疲劳断裂、沿晶断裂等，磨损失效机理有黏着磨损、磨粒磨损、表面疲劳磨损和磨蚀磨损等，腐蚀机理有化学腐蚀和电化学腐蚀，它们都有各自的形貌特征。因此，通过对断口的微观形貌特征观察，可定性地判断零件的断裂失效机理。对断口微观形貌特征的诊断技术和方法已在本章的前面几节中进行过详细的讨论，在此不再重复。

5.4.2　物理数学模型诊断

失效机理的定量研究的一个主要内容是失效物理和数学模型的研究。失效物理和数学模型大致上可以分为理化模型（物理、化学、材料、力学等方面的模型）和概率（统计）模型。但是由于材料微观物理和化学因素导致的失效在时间和空间上原本都是随机发生的，因而，上述两类失效物理数学模型之间必然存在其内在的联系。失效的物理数学模型大致有：界限模型与耐久模型、应力强度干涉模型、反应论模型、失效率模型、最弱环模型与串联模型、绳子模型与并联模型、比例效应模型、退化模型或损伤累积模型等。目前用得较多、较好的是应力 - 强度干涉模型和失效率模型。

按照应力 - 强度干涉模型，零件所受的应力和材料的强度（抗力）都服从各自的分布，零件失效是材料的强度小于零件所受应力的部分，即应力和强度发生干涉的部分，因此，可运用数理统计的方法求出零件失效的概率分布函数和密度函数，从而求出零件的失效概率，并从概率数值的大小来诊断出失效的本质，并提出相应的预防措施。

失效率模型则是从使用中的失效概率大小来判断失效的内在本质。统计学告诉我们，小概率事件代表偶然性，大概率事件代表必然性。若同类零件的失效率很小，则与系统性失误原因导致的可能性不大，而更可能是制造、使用维护和环境中的偶然失误造成的。相反，若零件的失效为大概率事件，则偶然性失误不会是失效的根本原因，顶多起到促进作用，而更根本的原因是设计、制造中的普遍现象。

前面对失效诊断技术和方法从断口、裂纹、痕迹、失效产物和综合诊断五个方面进行了专门的和综合的分析与讨论。表 5.2-18 列出了五大类、16 项共 52 种失效信息以及由这些失效信息综合诊断出失效的模式、原因和机理的思路。反之，也可根据特定的失效模式、原因和机理推断出可能表现出的相关失效信息（特征）。表 5.2-19 则归纳总结了从断口、裂纹、痕迹和参数四个方面的特征来分别诊断失效模式、原因和机理的技术、方法和思路。在失效诊断实际中，可以借鉴这两个表对失效的模式、原因和机理进行分别诊断和综合诊断。

表 5.2-18　失效信息与失效模式、原因和机理的诊断系统

信息	项目			失效信息名称	脆性（解理断裂）	脆性（沿晶）断裂	脆性（沿夹杂物）断裂	塑性断裂	疲劳断裂	腐蚀疲劳断裂	应力腐蚀（沿晶）断裂	
裂纹信息	裂纹起源位置和扩展方向	宏观	与零件应力、焊缝的关系	①在应力集中处（R 处）				●	●		●	●
				②在焊缝区								
				③非应力集中区				●			●	●
			与接触介质的关系	①在介质接触的表面							●	●
				②在点蚀（腐蚀）坑处						●		
				③与接触介质无关	●							
			与主正应力或主切应力方向的关系	①与主正应力方向垂直	●				●	●	●	●
				②与主切应力方向平行				●				
				③与主正应力、主切应力无关								
		微观	与显微组织的关系	①在夹杂物处			●		●			
				②在某一相组织处		●						
				③与显微组织无关								
			与晶粒边界的关系	①在晶粒边界或相界处		●				●		
				②在晶内或相内				●	●	●		●
				③既在晶界也在晶内处	●							
			与晶面晶向的关系	①在某一特定的晶面或晶向								
				②非特定的晶面、晶向				●		●		
				③混合情况								
	裂纹分布和形貌	宏观	按点、线分布情况	①以点放射分布	●				●	●		
				②沿线分布	●							
				③不规则分布	●				●		●	●
			数量和平直情况	①单条、“平直”状								
				②分叉或台阶或锯齿状			●		●	●		●
				③网状或龟裂形貌	●	●			●		●	●
			啮合和间隙情况	①啮合好、间隙小、裂尖尖锐								
				②啮合差、间隙大、裂尖圆钝				●				
		微观	与显微组织的关系									
			与晶粒边界的关系									
			与晶面晶向的关系									
断口信息	断口的形貌和特征	宏观	断口附近的残留塑性变形	①塑性变形大、断口凹凸不平				●				
				②塑性变形小、断口比较平直	●				●	●	●	●
			断口的颜色	①断口为本色	●				●			
				②断口上有氧化或腐蚀产物的颜色		●			●	●	●	●
			断口的形貌特征	①纤维状（鹅毛绒状）				●				
				②结晶状（颗粒状）		●	●				●	
				③放射状（人字纹）	●							
				④贝壳状（弧线状）					●	●		
		微观	断口的形貌特征	①韧窝（撕裂棱、微孔等）				●				●
				②解理（河流、扇形、台阶等）	●							
				③沿晶（岩石、冰糖块等）		●	●				●	
				④条带（辉纹、“轮胎”痕迹等）					●	●		
				⑤“鸡爪状”形貌							●	
				⑥腐蚀产物（氧化物、腐蚀产物、泥纹等）						●		
				⑦其他形貌								
腐蚀信息	材料与腐蚀介质和腐蚀特征形貌	材料与腐蚀介质		①碳钢及合金钢在 HCl、碱、硝酸盐、HNO_3、海岸大气、工业大气、H_2S、$H_2SO_4+HNO_3$ 等中							●	
				②铬不锈钢在 NaCl、氧化物、氟化物、溴化物、碘化物、HCl、海岸大气、工业大气、水及蒸汽、H_2S 等中							●	
				③奥氏体不锈钢在氟化物、氯化物、碱、海岸大气等中								●
				④奥氏体不锈钢在碱、$FeCl_2$、$FeCl_3$、H_2SO_4 等中						●		
				⑤材料上有腐蚀性的介质								
		腐蚀形貌特征		①腐蚀产物附着、剥离						●		
				②表面颜色发暗、敲击声音频率降低							●	●
				③腐蚀坑						●		
应力信息	应力性质和特征			①静载荷（$\sigma \geqslant \sigma_b$ 或 $k_1 \geqslant k_{1c}$）		●	●	●			●	●
				②交变载荷（$\sigma \geqslant \sigma_{-1}$）					●	●		
				③冲击载荷（$a \geqslant a_c$ 或 $k_1 \geqslant k_{1d}$）	●							

表 5.2-19　机械使用失效的主要模式、原因和机理的主要形成条件和主要诊断思路

失效模式 \ 内容	失效分类和形成主要条件（或初步特征）					诊断的主要依据			
	断裂与否	宏观塑性变形大小	载荷循环与否	有无腐蚀介质	腐蚀介质的种类或环境	断口诊断依据	裂纹诊断依据	痕迹诊断依据	参数诊断依据
失效	断裂失效	脆性断裂（变形小）	一次性脆断（不循环）	无腐蚀介质	脆性材料	断口有夹杂物	宏观平直，微观沿夹杂物	有时有撞击痕迹	$\sigma_f^{①} > \sigma_b$，$\sigma_{0.2} \approx \sigma_b$
				材质脆性	低温脆性	解理或准解理，有时为少量沿晶	宏观平直，微观弯折		$T_{工作} < T_c^{②}$
				温度脆性	低应力脆性	$a > a_c^{③}$（a很小）	宏观平直		$K_I > K_{Ic}$ 或 $a > a_c$
				腐蚀脆性（有腐蚀介质）	蠕变（高温）	沿晶，晶界有空洞或楔形裂纹，氧化	沿晶，啮合性差	各种腐蚀痕迹	环境参数
					氢脆（氢介质）	冰糖块状断口，有鸡爪形花样	宏观平直		
					碱脆（碱介质）	沿晶，有碱金属残留	沿晶裂纹		
					低熔点金属脆性（低熔点金属）	沿晶或准解理 有低熔点金属痕迹	沿晶或穿晶		
失效	断裂失效	脆性断裂（变形小）	一次性脆断（不循环）	腐蚀脆性（有腐蚀介质）	辐照脆性（辐照射线）	沿晶或准解理	沿晶或穿晶	各种腐蚀痕迹	环境参数
					液体浸蚀脆性（腐蚀液体）	沿晶 有浸蚀液残留	沿晶		
					晶间腐蚀（晶间贫Cr+电化学腐蚀环境）	沿晶 晶间附近贫Cr	宏观平直，微观曲折		
					应力腐蚀（介质与材料匹配）	沿晶或准解理 有时有腐蚀产物	裂纹分叉，沿晶或穿晶		
					氢腐蚀（$4H + Fe_3C \rightarrow CH_4 + 3Fe$）	沿晶 晶界上有CH_4气泡	沿晶，有CH_4气泡		
			疲劳断裂（循环）	一般疲劳断裂（无腐蚀介质）	机械疲劳断裂（机械循环应力引起）	疲劳弧线或疲劳条带或台阶	一般平直，穿晶		$\sigma_a > \sigma_R^{④}$
					热疲劳断裂（温度交变引起）	疲劳条带间隔大，二次裂纹多，氧化严重	一般平直，穿晶，有时沿晶	氧化腐蚀痕迹	有温度的交变
					室温疲劳断裂	疲劳弧线、条带、台阶	一般平直，穿晶，啮合性好		$\sigma_a^{⑤} > \sigma_R$
失效	断裂失效	脆性断裂（变形小）	疲劳断裂（循环）	一般疲劳断裂（无腐蚀介质）	高温疲劳断裂	疲劳弧线、条带、台阶	裂纹内有氧化或腐蚀	高温腐蚀氧化痕迹	$\sigma_a > \sigma_R$
					高周疲劳断裂	疲劳弧线、条带、台阶	一般平直		$\sigma_a < \sigma_{0.2}$，$N_f^{⑥} > 10^5$
					低周疲劳断裂	疲劳条带间隔大，有时有成排韧窝	一般平直	塑性变形痕迹	$\sigma_a > \sigma_{0.2}$，$N_f < 10^5$
				腐蚀疲劳断裂（有腐蚀介质）	凡是对材料有腐蚀作用的介质均会引起腐蚀疲劳断裂	疲劳弧线、条带不清晰 有腐蚀产物	有沿晶，有穿晶，啮合，有时裂纹内有腐蚀产物	腐蚀痕迹	腐蚀参数
		塑性断裂（变形大）		一次性断裂	应力大，或强度低	鹅毛绒状或纤维状韧窝	穿晶，啮合差	塑性变形大	$\sigma_f > \sigma_b$
	非断裂失效	腐蚀失效	无循环	有腐蚀介质环境	非电化学腐蚀（高温氧化）	表面氧化膜，有的材料氧化膜剥落	一般穿晶，等强温度以上沿晶	腐蚀痕迹	氧化或腐蚀参数
					电化学腐蚀：一般电化学腐蚀	一般为沿晶	沿晶裂纹		
					电化学腐蚀：液滴下的氧电池腐蚀	一般为沿晶	沿晶裂纹		
					电化学腐蚀：缝隙腐蚀	一般为沿晶	沿晶裂纹		
					电化学腐蚀：显微选择性腐蚀	异相遭腐蚀，或相界扩展（孔洞）	沿相界		
					电化学腐蚀：其他电化学腐蚀				

续表 5.2-19

失效模式＼内容	失效分类和形成主要条件（或初步特征）					诊断的主要依据			
	断裂与否	宏观塑性变形大小	载荷循环与否	有无腐蚀介质	腐蚀介质的种类或环境	断口诊断依据	裂纹诊断依据	痕迹诊断依据	参数诊断依据
失效	非断裂失效	磨损失效	有时有循环应力	可以有腐蚀介质环境	微动磨损	磨蚀坑、痕、微裂纹⑦	磨擦痕与相对运动方向垂直，裂纹扩展方向与主应力方向垂直，次表面疲劳生核	磨擦痕迹和磨削形貌	相对运动、润滑及应力参数
					磨粒磨损	磨损坑、划痕⑦			
					腐蚀磨损				
					疲劳磨损（接触疲劳）	磨蚀坑、“鱼眼”断口⑦			
					黏着磨损	有黏着、脱落⑦			
		变形失效		无腐蚀	弹性失稳失效			弹性和塑性变形痕迹	$\sigma > \sigma_{失稳}$
					塑性变形失效	各种形式的塑性变形痕迹			$\sigma > \sigma_{0.2}$
					形状和尺寸变形失效（显微组织稳定性和内部应力联合作用）	形状和尺寸的变化痕迹			尺寸稳定性参数

①工作应力；②韧脆转变温度；③临界裂纹长度；④交变应力；⑤平均应力；⑥疲劳寿命；⑦为表面特征。

6　智能诊断技术和方法

随着科学技术的不断进步和发展，尤其是测试技术、计算机技术的迅速发展和普及，设备状态监测和故障诊断技术得到了大量的应用，在国际上成为一大热门学科。

6.1　状态监测和故障诊断技术

设备状态监测与故障诊断技术的实质是了解和掌握设备在运行过程中的状态，评价、预测设备的可靠性，早期发现故障，并对其原因、部位、危险程度等进行识别，预报故障的发展趋势，并针对具体情况做出决策。

设备状态监测与故障诊断技术包括识别设备状态和预测趋势两方面的内容。具体过程分为状态监测、分析诊断和治理预防三个基本环节（图 5.2-68）。

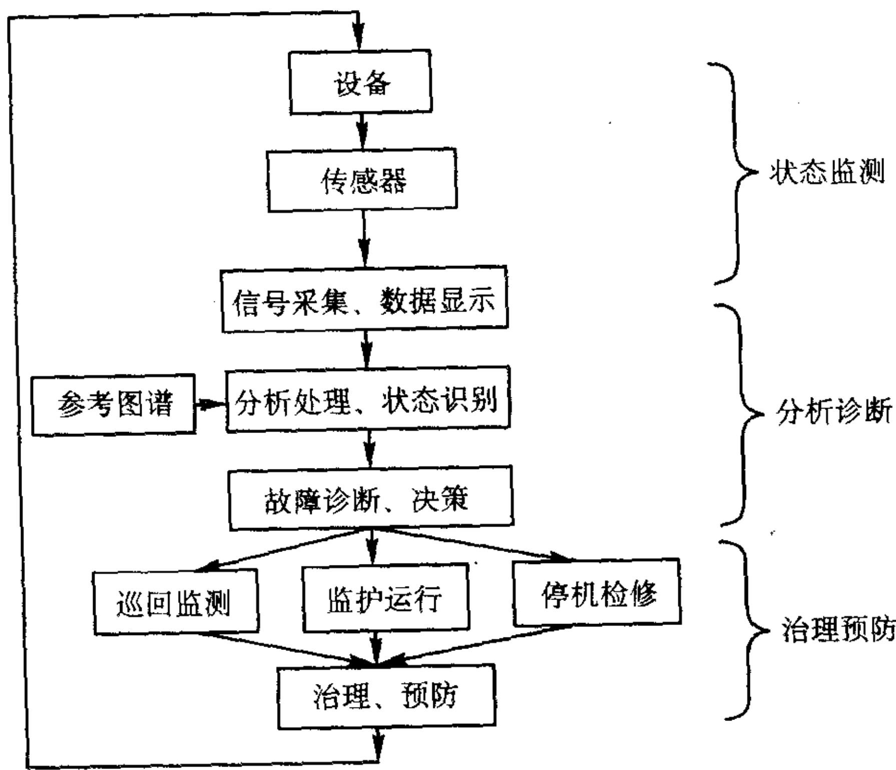


图 5.2-68　状态监测与故障诊断的三个阶段

状态监测是设备运行中，对特定的特征信号进行检测、变换、记录、分析处理并显示、记录，是对设备进行故障诊断的基础工作。检测的信号主要是机组或零部件在运行中的各种信息，通过传感器把这些信息转换为电信号或其他物理量信号，送入信号处理系统中进行处理，以便得到能反映设备运行状态的特征参数，从而实现对设备运行状态的监测和下一步诊断工作。

从传感器或人的感官所获取的信息需要记录下来，供分析和日后对比、查阅，这就需要进行数据采集和简单的处理、显示。大多数传感器输出的信号属于模拟信号，在磁带记录仪上记录时不需要进行数字转换，而输入计算机或进行分析处理时，还需要转换为数字信号（即进行 A/D 转换），这些是数据采集的主要工作。

在这些信息和信号中，有些是有用的，能反映设备故障部位的症状，这种信息称为症兆或故障症兆；有些并不是诊断所需要的信号，因此，需要处理和排除。为了提取症兆信号，需要做一些特征信号的提取工作，这是由信号处理系统来完成的。

分析诊断实际上包括两方面的内容：信号分析处理、故障诊断。

由传感器或人的感官所获取的信息往往特征不明显、不直观，很难直接进行故障诊断。信号分析处理的目的是把获得的信息通过一定的方法进行变换处理，从不同的角度提取最直观、最敏感、最有用的特征信息。分析处理可用专门的分析仪器或计算机进行，一般情况下要从多重分析域、多个角度来分析观察这些信息。分析处理方法的选择、处理过程的准确性以及表达的直观性都会对诊断结果产生较大影响。

故障诊断是在状态监测与信号分析处理的基础上进行的。进行故障诊断需要根据状态监测与信号分析处理提供的能反映设备运行状态的症兆或特征参数的变化情况，有时还需要进一步与某些故障特征参数（模式）进行比较，以识别设备是运转正常还是存在故障。如果存在故障，要诊断故障的性质和程度、产生原因或发生部位，并预测设备的性能和故障发展趋势。这是设备诊断的第二阶段。

设备状态监测与故障诊断既有区别又有联系，在生产实际中，有时又将二者统称为设备故障诊断。实际上，没有监测就没有诊断，诊断是目的，监测是手段；监测是诊断的基础和前提，诊断是监测的最终结果。

状态监测通常是通过监视和测量设备或部件运行信息和

特征参数，例如振动、温度、压力等，并以此来判断其状态是否正常。当特征参数小于允许值时便认为是正常，否则是不正常。还可以用超过允许值的多少来表示故障严重程度，当它达到某一设定值时就应停机检修。这个过程的前半部分就是状态监测。

有些情况下，监测结果不需要作更进一步的处理和分析，仅以有限的几个指标就可以确定设备的状态，这就是以监测为主的简易诊断，也属于诊断的范围。通常情况下，当简易诊断发现设备或部件发生异常时，应转入精密诊断。此时应该对异常状态进行多方面的分析。这种分析包括收集设备运行的历史资料、对简易诊断的结果进行审核，同时进一步合理地选择测量仪器对设备的各种参数进行监测，对监测得到的特征信号在时域、频域、幅值域以至倒频域等各个方面进行全面分析，以便从特征信号中提取各种症兆，对设备做出综合判断。通常所称的“故障诊断”不是简易诊断，而是指比较复杂的精密诊断。

根据诊断环境，诊断有离线人工分析、诊断和在线计算机辅助监视诊断两种。在线监测是在线诊断的基础，信号分析及数据处理方法是在线诊断中的重要手段。

目前所采用的监测诊断技术可以概括为三类。

（1）以检测仪表为主体的监测装置

目前许多引进的大型设备都配备了这种装置，如本特利(Bently)公司的Bently序列和飞利浦(Philips)公司的Philips序列都有引进，其中以Bently序列较多，如Bently7200、9000、3300型等。它的主要构成部件是传感器和指示仪表箱，有用于测温度的，但大多数是用于测量振动的。其主要缺点是：

1）检测信号是随机的，幅值并不能全面地表达动态过程的特性；

2）机组在强烈振动之前，故障症兆并不很明显，有时振幅变化并不很大，但机组确有故障，如半倍频是故障的重要信息之一，但检测仪表并未显示出来。而一旦振幅突然增大，则为时已晚，即不能防止突发性故障；

3）读数式检测仪表本身并无分析功能，依赖于人的经验判断。

（2）检测仪表配备硬件分析装置

这种系统是第一种装置的改进与补充，所用的装置主要是频谱分析仪，有些名为故障诊断仪，而实际主要功能还是频谱分析，也有的分析功能是用计算机软件去实现。如早几年本特利公司的ADRE3和恩特克（Entek）公司的PM等系统就具有频谱分析、谱阵图、波特图（Bode Plot）、轴心轨迹图等功能，故有帮助人们提高诊断的准确性的优点，但也仍存在以下缺点：

1）分析装置只是一种工具，不能自动判断，诊断决策仍需依赖于领域专家；

2）不能连续地自动分析，容易丢失故障信息，不能预防突发性故障；

3）大型机械设备的结构复杂，故障与症兆之间并无一一对应的因果关系，难免有误诊的。

（3）计算机辅助监测诊断系统

这种系统主要结构是由传感器、接口装置及计算机组成。其中接口装置具有电平转换、采样、存储等功能。它可以实时监测和自动诊断，对防止突发性故障有利，是工况监测与故障诊断技术的主要发展领域。但目前的水平主要是计算机辅助监测与诊断系统，还不能真正达到自动诊断的水平。国内外都有这种系统的开发与应用，但仍无商品，除了技术成熟性不足之外，主要原因还是由于大型机械设备故障诊断的针对性很强，且领域专家知识仍然是故障诊断不可缺少的一部分，而商品型诊断系统必须充分考虑通用性。

6.2 专家系统在诊断中的应用

6.2.1 专家系统简介

设备故障诊断工作属于工程实际中的新兴学科，它具有如下特点。

1）知识可以从类似机械和工作实际、诊断实例中获得，即知识来源比较规范。

2）诊断对象多为复杂的、大型的动态系统，这种系统的大部分故障是随机的，普通人很难判断，这时就需要通过讨论或请专家来进行诊断。但对于一些新型机器，可能无处获得诊断知识；或者对于非定型生产的机器，由于其工作特性和常用机器相比差异很大，知识获取也十分困难。

3）实际应用后，经济效益和社会效益十分明显。

而专家系统恰恰适用于复杂的、知识来源规范的大型动态系统，它可以汇集众多专家的知识，进行分析、比较、推理，最终得出正确的结论。现场技术人员可以充分利用各种信息和症兆，在计算机系统的帮助下有效地解决工程实际问题，这也是专家系统在设备故障诊断过程中得到成功应用的原因。

设备故障诊断专家系统是将人类在设备故障诊断方面的多位专家具有的知识、经验、推理、技能综合后编制成的大型计算机程序，它可以利于计算机系统帮助人们分析解决只能用语言描述、思维推理的复杂问题，扩展了计算机系统原有的工作范围，使计算机系统有了思维能力，能够与决策者进行“对话”，并应用推理方式提供决策建议。

设备故障诊断专家系统的另一个特点是通过对话窗口能使那些不具备编程能力的工程技术人员建立功能强大的程序系统。这样，实际工作中生产一线的工程技术人员就不会再受计算机语言和编程技巧的限制，可以将他们的故障诊断经验、知识和遇到的故障实例输入专家系统，使专家系统不断学习、提高，丰富其知识库，提高故障诊断的准确率。与此同时，专家系统提供的处理问题的对策也将更具有针对性、更为有效。

典型专家系统的基本结构如图5.2-69所示。它包括知识库、推理机、数据库和用户界面四个部分。

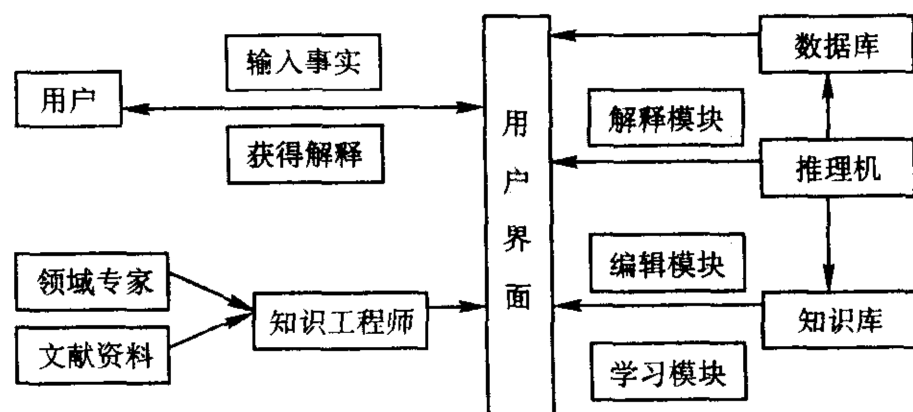


图5.2-69 专家系统结构示意图

1）知识库 知识库是用来存放专家知识、经验、书本知识和常识的存储器。知识工程师通过与领域专家交流和阅读文献资料，获得领域知识和经验，并把它们以一定的形式表示出来，形成知识的外部表示方式，然后将知识的外部表示转换成计算机能识别的内部表示形式，存放于知识库中，知识的外部表示一方面要适应特定领域、特定问题的特点，一方面要易于转换成能节省储存空间和运行时间的内部表示形式。

常用的知识表示形式有四种：产生式、框架式、语义网络、对象式。其中产生式表示方法比较简便、易于推理，在专家系统中应用得最广泛，特别适于构建领域知识相对简单的专家系统。框架式知识表示方式特点是领域知识可以用框架来表示，父框与子框间的属性可以自动继承或重新定义，

特别适合构建领域知识相对复杂的专家系统，这种方式在专家系统中也应用得很广泛。语义网络表示方法是对框架式的推广，领域中的各个概念通过表示它们之间的关系的纽带联系起来。对象式表示方法是对框架式的改进，这种方式实现了数据的封装和继承，符合当今面向对象的程序设计方法的需要。

专家系统的知识库应该能够修改、更新和扩充，这是通过机器学习过程来实现的，学习方式有机械记忆学习、提问学习、实例学习等形式，无论采用那种学习方式，专家系统都应进行知识的冗余性检查和一致性检查，以避免新知识的加入所带来的副作用。

2）推理机　推理机是专家系统的核心之一。它按照一定推理策略，根据用户输入的数据，运用知识库中的知识进行推理，得出结论，并能解释其推理过程，推理机的形式取决于知识的表示形式。常用的推理方式有以下三种。

① 正向推理　正向推理又称数据驱动推理，它从已知的事实出发，推出未知的结论。以产生式专家系统为例，它从用户输入的事实出发，寻找前提条件与之匹配的规则，如果找到这样的规则，就把这些规则的结论加入已知事实中，再重复上述过程。

② 逆向推理　逆向推理又称目标驱动推理，它从假设的结论出发，寻找可靠的前提以证实这个结论。以产生式专家系统为例，它从某种假设出发，寻找结论与之匹配的规则，如果找到这样的规则，就以这些规则的前提为新的假设，再重复上述过程。

③ 混合推理　混合推理是对上述两种推理过程的综合应用。单纯的正向推理，目的性不强、搜索效率低，而单纯的反向推理，初始假设盲目性大，因此，常使用混合推理方式。

例如，在旋转机械故障诊断中，振动信号是主要的诊断依据，大量的诊断知识是以振动理论为基础的。在振动信号中，频谱又是诊断的首要依据。因此，通常旋转机械的故障诊断是以频谱作为正向推理的症兆，得出故障假设集，再以假设集去指导逆向推理。

专家系统在推理时，既可以进行精确地推理，又可以在事实不确切、不完整，或者事实与结论之间的因果关系是不确定的情况下，进行不确定性推理。

专家系统的解释机制有两种，一种是在推理过程中，系统根据用户提供的原始事实无法完成推理时，就会要求用户提供一些额外的事实，当用户询问为什么需要这些事实时，系统能做出解释。另一种是系统得出结论后，能解释其整个推理过程。通过“记忆”其推理过程中所用的规则，就可实现这两种解释机制。

3）数据库　专家系统的数据库又称为黑板结构，用于存放用户输入的原始事实、推理的中间结果和最终的结论。

4）用户界面　专家系统的用户界面应该是一个“友好界面”，便于用户操作。这一方面是指通过在线帮助，即使初次使用该系统的用户仅略懂计算机操作常识，也能够正确地进行操作。另一方面是指当用户的输入发生错误时，系统具有容错能力，并提示用户重新输入。

6.2.2　专家系统应用

专家系统是人工智能的一个分支，目前在许多领域都得到了广泛的应用。其中在失效诊断领域中也进行了相当的研究，表5.2-20列出了国内外用于失效诊断的专家系统。

其中，DELA/CAST-1系统由美国通用电气公司在1984年研制成功，用于电机车的诊断和维护。该系统知识库中有350条规则，并对规则进行了分组。该系统采用混合推理方式，能解决80%的实际问题。汽车发动机诊断系统由我国研制，该系统将知识分为静态知识和动态知识，前者用90个框架表示，后者用200条产生式规则表示。该系统采用正向推理方式，对四类十几种故障诊断的正确率高达90%。

表5.2-20　用于失效诊断的专家系统

名　称	用　途
DELA/CAST-1	电机车诊断和维护
汽车发动机诊断系统	诊断汽车发动机的故障
PDS	汽轮发电机的监控和诊断
FIES	故障探测和隔离分析
NTC	故障检查
ACE	电话网维护
REACTOR	核反应堆事故诊断
ENGING COOLING ADVISOR	汽车发动机冷却系统噪声诊断
XMAN	飞机喷气式发动机故障诊断
直升机发动机转子诊断系统	直升机发动机转子故障诊断
DRILLING ADVISOR	油井钻探问题诊断
DART/DASD	计算机故障诊断
CRIB	计算机硬件故障诊断
列车机车诊断系统	列车机车故障诊断
大型直流电机诊断系统	大型直流电机故障诊断
梁内裂纹诊断系统	梁内裂纹诊断

6.3　失效案例库及其应用

失效分析工作意义在于能够找到事故的原因，并制定相应的预防措施，以防止将来再次发生。然而失效分析又是一项非常复杂的工程，它涉及到机械、物理、冶金、化学、力学、摩擦学等许多不同领域。对于一名从事失效分析工作的专家来说，他不仅仅要掌握上述领域内的知识，更重要的是要将这些内容有机的联系起来，这需要专家有非常丰富的实践经验。但是这样的专家太少了，对于广大工程技术人员来说，根本无法快速成为一名专家，况且即便是专家，也会有疏忽的时候。另外，专家可能会遇见一些难题，一时无法解决，这时就需要查找一些相关的资料来参考，而相似的案例是最好的参考资料，通过案例的比较，专家就能很快找到线索。

另外一方面，在设备的设计、制造、使用和管理过程中，设计者、制造者、使用者和管理者都需要丰富的经验和专家的指导才能够保证设备安全、稳定的运行。但是实际的情况并非这样。根据调查，有许多的事故有着惊人的相似之处，如设计时采用的材料不当，导致设备很快失效，甚至发生灾难。而并非所有设计人员都能及时得到事故分析的报告，他们仍然采用这种材料来设计这些设备，这样就可能导致同样的事故再次发生，相反，如果设计人员或设备的操作者能及时的得到案例信息，就可以及时的采取措施来预防事故的发生。

失效事故的发生常常造成巨大的财产损失和人员伤亡。失效的时间相关性、随机性和突发性特点，促使我们要深入总结和研究，充分利用这些数据，从而可以正确诊断失效模式，吸取以往的经验和教训，准确预测和有效预防或减少类似事故的再次发生。

大量的失效案例分散在国内外多种出版物及资料上，查找比较困难，如果采用人工方式对腐蚀案例数据进行整理归类将是一项很庞大的工程。当前计算机技术，特别是数据库技术飞速发展，完全可以将计算机领域的研究成果运用到失效案例的管理中。

数据库技术是在文件系统的基础上发展产生的。数据库中数据是按三级模式组织，用户使用的数据是由外部存储器

中真实存在的数据经过二级映射而得到。另外数据库中的数据文件之间的联系是由数据库管理系统自身实现的，而与应用程序无关。

一个完整的数据库系统由数据库、数据库管理系统、数据库管理员和应用程序四部分组成，如图 5.2-70 所示。其中数据库是相关数据的集合，以一定的组织形式存储在永久存储介质上；数据库管理系统是管理数据库的软件，实现在数据库上的各种操作，它是数据库系统的核心；数据库管理员是一个或一组人员，是数据库的高级用户，他完成数据库的规划、设计、维护和管理等工作；应用程序是各种用户使用数据库的界面，它在数据库管理系统支持下为用户提供各类服务。数据库管理系统有两个最基本的特点，这两个特点使它区别于其他类型的程序设计系统：管理长期应用的数据的能力；高效地存取大量数据的能力。

失效案例库实际就是一个数据库系统，它的功能很多，图 5.2-71 为一个腐蚀失效案例库的主要功能，其中查询、添加、修改等功能是案例库的主要功能，同时，案例库中还应具有一些辅助功能。

查询是案例库的主要功能之一，可以设计简单查询、高级查询等方法，图 5.2-72 是一个简单查询的流程图。

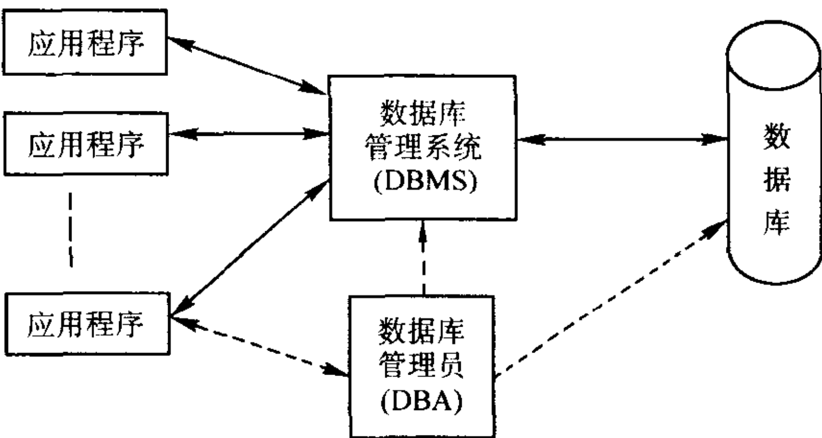


图 5.2-70 数据库系统的组成

数据库技术的发展不仅表现在所支持的数据模型的发展，即支持的数据模型越来越复杂，包含越来越多的语义，同时还出现了许多新的发展方向，支持特殊领域应用要求的数据模型的数据库技术，其中多媒体数据库和 Web 上的数据库非常适合失效分析领域的应用。不仅可以存储文字数据，还可以存储图像、声音等数据。通过互联网可以实现数据资源共享。这些新技术的发展，为案例库的应用提供了更加广阔的前景，为失效诊断和预防提供了新的手段。

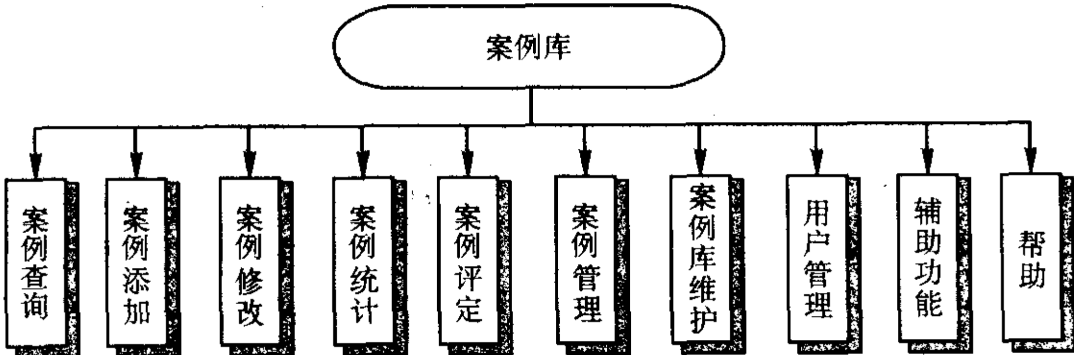


图 5.2-71 案例库的主要功能

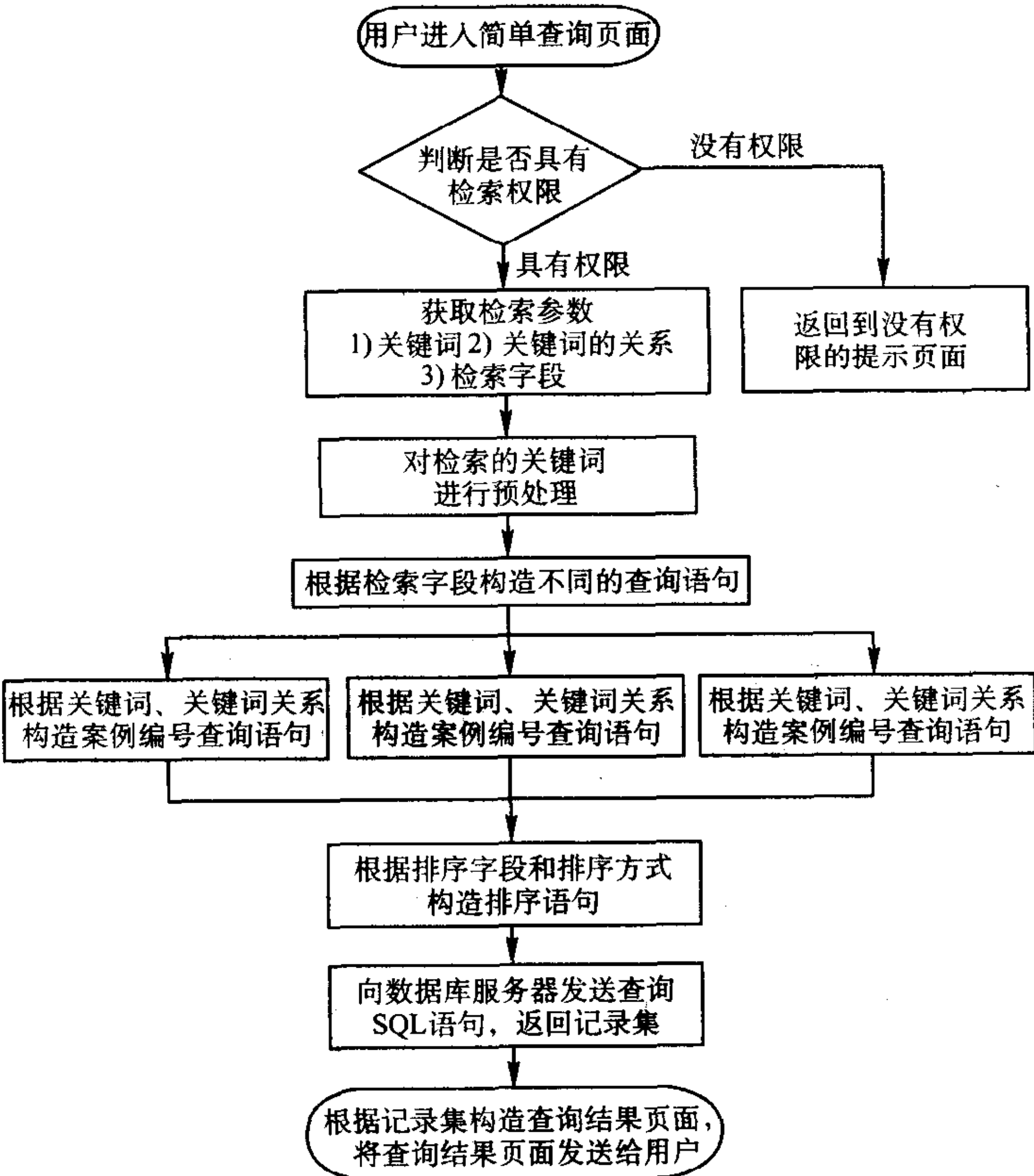


图 5.2-72 简单查询流程图

编写：付国如（北京航空航天大学）

第3章　失效预测技术和方法

1　以失效模式为基础的失效预测技术和方法

1.1　金属材料失效抗力预测

1.1.1　金属材料的强化手段

不同的机械构件具有不同的失效方式，即使是相同的构件，如果选用的材料不同或状态不同，构件断裂的原因、过程、断口的特征上也会出现差异。因此，机械构件的失效除取决于构件所受外力及环境、温度等因素之外，还取决于构件所选用的金属材料本身的性能是否满足使用要求。

金属基体不同、强化方式不同、热处理状态不同，使得材料的性能不同，应用的范围不同，构件失效的方式也不同。

机械产品是由众多的机件或构件组成的，每个构件都有各自的性能要求。构件的服役条件包括构件所承受的应力大小、状态及其随时间的变化、构件的工作环境（温度、介质等)、应力分布等。

根据构件的服役条件，设计者对每一个构件都提出了特定的要求。这些要求概括起来有如下几点：构件应具备规定的使用寿命，安全可靠，成本低，效能高，结构紧凑，质量小（对大多数构件）。为满足这些要求，必然对材料提出了一系列相应的要求，如材料的物理性能：密度、线胀系数、导热系数、电导率、磁导率、比热容等。材料的化学性能：抗氧化性能及抗各种介质的腐蚀性能。材料的力学性能：抗拉强度（σ_b），屈服点（σ_s），伸长率（δ），断面收缩率（ψ），冲击韧度（a_K），断裂韧度（K_{IC}）以及弹性模量（E）等。此外，材料还应具有较好的工艺性能：铸造性能、压力加工性能、焊接性能、热处理性能及各种冷加工性能等。

当然，对于某一种构件，在选材时，根据构件的要求，对材料性能的某些方面应提出较严格的要求，而对另一些方面的要求则可以相对放松些。人们不可能，也没有必要对材料提出全面苛刻的要求。

为了满足机械产品设计者对材料所提出的要求，材料工作者设计和生产了几百种，甚至上千种牌号的材料，如金属材料，复合材料，高分子材料及陶瓷材料等。作为工程结构材料，目前使用最广、最多的是金属材料。

依据金属材料基体的不同，可将金属材料分为六大类，铁及其合金、铝及其合金、镁及其合金、钛及其合金、铜及其合金、耐热合金。

纯金属在工业上的应用是极有限的，这是由于它们很难满足强度上的要求。为了满足对工程结构材料的强度及综合性能的要求，必须进行金属合金化以及采取其他强化方法。这些方法可归纳成如下几种。

1）固溶强化　在基体金属中加入适量的合金元素，形成置换式或间隙式固溶体而强化。加入合金元素之后，如形成置换式固溶体，由于基体（溶剂）和合金元素（溶质）的电子结构不同，切变弹性模量也不一致，使其发生固溶强化。形成间隙式固溶体时，由于基体晶格畸变而强化。

2）冷作硬化　金属材料在再结晶温度以下进行变形时，所引起的强度、硬度的增加，延伸率、断面收缩率下降的现象叫冷作硬化或冷变形强化。

冷作硬化的原因是，在外力作用下，金属材料的晶粒破碎，晶格畸变，晶体内位错密度增加，位错互相缠结，有时形成胞状结构，使位错的运动愈加困难。由此引起材料的变形抗力和强度提高，而材料的塑性降低。

3）细化晶粒　细化晶粒是提高金属材料室温性能的有效方法之一。随着晶粒的细化，材料的综合性能逐步提高。晶粒越细小，材料中单位体积内的晶界面积就越大。晶界上的原子排列是最紊乱的，而且存在大量的缺陷，这些都会阻碍位错的运动。这就是细化晶粒得以提高材料强度的原因。

细化晶粒的另一好处是可以有效地降低材料的韧－脆转变温度。

4）第二相强化　第二相强化也称为弥散强化。当合金元素的含量较高并超过一定限度之后，淬火可获得过饱和固溶体，将它再在较低的温度加热，过饱和固溶体便发生分解而析出第二相，由此合金获得弥散强化。也可以采用将过量的合金元素加入到基体金属中去的方法，其中的一部分溶入固溶体，而另一部分则形成第二相。

当金属中存在着弥散分布的第二相质点时，这些第二相质点与基体之间存在着一定的键合力，使得第二相质点均匀地钉扎在基体内部，阻碍位错运动。第二相质点越细小、分散、致密，对位错的钉扎作用就越大。

改变金属材料中第二相的大小和分布的最有效的办法就是热处理。所以，热处理是提高金属材料强度的有效措施。

5）相变强化　相变强化的效果很大程度上取决于相变温度，低温下固溶体过饱和度增大，强化效果好。同时，相变温度愈低所产生的新相的晶粒愈细，第二相质点也愈细，这些都增加了相变强化的效果。

在合金设计和使用中，往往将上述几种方法同时采用，以达到最佳强化效果。

1.1.2　钢的基体组织对其失效抗力指标的影响

（1）冶金因素对失效抗力指标的影响

任何一种断裂失效过程都可以分为两个阶段，即裂纹的萌生阶段和扩展阶段。在外界条件确定的情况下，就材料本身而言，裂纹总是沿着薄弱环节萌生并沿着某一薄弱地区扩展。晶体内部的各种界面往往成为这些薄弱环节。这些界面包括：晶界面（大角度晶界、亚晶界、孪晶晶界)，相界面（硬性第二相界面、软性第二相界面)、解理面。

设法改善上述界面的物理、化学、力学状态可以适当地起到抑制裂纹萌生，提高材料的抗断裂失效的能力。

裂纹总是沿着能量消耗最小的方向即阻力最小的途径扩展。就裂纹扩展所受到的阻力和消耗的功而言，沿晶断裂小于穿晶断裂，穿晶断裂中的解理断裂所消耗的功远小于韧窝断裂。因此，应采用合金强化和软化手段，设法减少沿晶断裂和解理断裂。

下面对影响钢的失效抗力指标的某些冶金因素加以论述。

1）回火温度对钢的力学性能的影响　淬火中所得到的不稳定状态的马氏体和残余奥氏体在回火过程中会分解。其中，低碳马氏体中的碳向位错线附近偏聚，使碳原子处于偏聚状态。高碳马氏体中碳原子在孪晶界界面上聚集成富碳区，并逐步形成 ε 碳化物。进一步加热则形成稳定的碳化物，即渗碳体。在回火过程中铁素体也进行再结晶。这时的组织叫做回火索氏体。淬火钢中的残余奥氏体在回火过程中也会分解为碳化物和铁素体。在回火的最后阶段，铁素体的晶粒不断长大，渗碳体的颗粒位于铁素体的晶界，渗碳体颗粒的平均尺寸的增大是靠小颗粒的溶解和大颗粒的长大进行的，这也促进了铁素体晶粒的长大。

这些转变必然使材料的强度和硬度降低，韧性提高。随着回火温度的增高钢的硬度逐渐下降，但在 180℃以前，硬度降低很少。如果在同一温度回火，含碳量较高的钢比含碳量较低的钢硬度高。

强度韧性的变化规律与硬度变化规律相似，而冲击韧度则反之，随着回火温度的提高，其室温的冲击韧度逐渐提高。但在 230～370℃之间回火时，冲击韧度反而下降。这被称为低温回火脆性，而且是不可逆的，又被称为不可逆回火脆性或第一类回火脆性。大量的构件是在 370～675℃之间进行回火的，因为其韧性作为最重要的指标之一。在这一范围内，如果在较低（370～540℃）温度回火，可使构件获得良好韧性的同时具有一定的强度和硬度。如果在较高（540～675℃）温度回火，则可使构件具有最高韧性，但强度有所降低。

某些合金结构钢在 540～650℃之间回火后，如果缓慢冷却会使韧性降低，但可以通过再次回火并快速冷却而进行弥补。这种现象被称为高温回火脆性，又称为可逆回火脆性或第二类回火脆性。

2）合金元素对钢淬火－回火状态力学性能的影响　钢中的合金元素可以分为二大类：一类是碳化物形成元素（如锰、铬、钼、钨、钒、钛、铌、锆等）；另一类是非碳化物形成元素（如镍、硅、钴、铜等）。

淬火加热时，除锰之外，所有的合金元素都在不同程度上细化奥氏体晶粒，其中尤以碳化物形成元素的作用最大。因为这些元素在钢中形成高稳定性的弥散状态合金碳化物，在淬火加热中它们又很难熔入奥氏体，因而起到一种钉扎作用，机械地阻止了奥氏体晶粒的长大。这正是人们所希望的。

淬火冷却时，除钴以外，熔入奥氏体的合金元素都增加过冷奥氏体的稳定性，推迟奥氏体的等温转变过程。使 C 曲线右移，从而使钢的淬透性提高。

回火过程中，所有合金元素均在不同程度上增加钢在回火时的组织稳定性，这是由于合金元素降低了碳在钢中的扩散速度，因而将马氏体分解和渗碳体析出过程推移到更高的温度。因为合金元素能提高钢的抗回火软化能力，所以在同一温度下回火，合金钢比碳钢具有更高的硬度。

总之，合金元素阻止奥氏体晶粒长大，增加钢的淬透性，使钢的抗回火软化能力提高，当合金元素形成合金碳化物时，钢的强度升高，当合金元素熔入基体而使固溶体强化、韧性略有降低。但当含量不超过一定量时，由于合金元素有细化晶粒的作用，这使钢的韧性不会显著降低，有时反而会在强度提高的同时，韧性也有适当的提高。

（2）奥氏体化温度对钢的力学性能的影响

通常淬火加热温度的确定是以把钢加热到奥氏体相变温度之上 20～40℃为准，这样保证了淬火加热时钢的组织全部转变成奥氏体，同时又防止了晶粒过度长大。从而保证了淬火后获得细小的马氏体组织，而回火后可获得满意的综合力学性能。

当淬火加热温度高于上述温度时，奥氏体晶粒长大。奥氏体晶粒的大小对钢件冷却后的组织及性能有重大影响。奥氏体晶粒大小不同，回火后所得到的马氏体组织不同，力学性能也就不同，特别对回火后的韧性有很大的影响。粗晶粒淬火后，将得到粗大的马氏体，使钢的韧性显著降低。因此，粗大的马氏体是钢发生过热的一种表现。

（3）非金属夹杂物对钢的力学性能的影响

1）非金属夹杂物的分类和鉴别　钢中的非金属夹杂物多是在冶炼、铸造过程中形成的。主要是脱氧产物，如氧化铝、氧化硅。有时还形成复杂的硅酸盐和氮化物（AlN，TiN …）和硫化物。钢中的非金属夹杂物的另一种均是即浇铸时炉衬、锭模的耐火材料剥离而掺杂于钢液中，钢液凝固时来不及浮出而部分地存在于钢中形成了夹杂物。这些非金属夹杂物的成分、性质、数量、分布、尺寸、形态等，全面的影响着材料的冷、热加工性能和材料的力学性能。因此，钢中夹杂物的成分、数量、分布、尺寸等在相应的国标及冶标中都有明确的规定。

钢中的非金属夹杂物主要有氧化物夹杂、硫化物夹杂、硅酸盐夹杂及氮化物夹杂。

2）非金属夹杂物对钢的力学性能的影响　非金属夹杂破坏了金属基体的连续性，引起微区的应力集中，降低了钢的力学性能，尤其是冲击韧度、疲劳极限和断裂韧度。同时增加变形后钢材性能的方向性差异。

夹杂物对钢的力学性能的影响不仅取决于夹杂物的数量及化学性质，更与夹杂物的某些物理特性密切相关，如夹杂的可变形性，线胀系数及弹性模量。这些物理性能与钢相应的物理性能之间的差异促使钢中局部应力升高，严重者产生内部微裂纹，轻者会在外力与局部应力集中叠加后使钢材产生早期开裂失效。造成局部应力集中的原因有：

① 热工艺过程中冷却时热收缩不匹配所造成的嵌镶应力。

嵌镶应力的一般关系式为

$$\sigma' = \phi(\alpha_2 - \alpha_1)\Delta T \tag{5.3-1}$$

式中，ΔT 为温度变化范围；α_2、α_1 为夹杂物和基体的线胀系数；ϕ 为夹杂尺寸、形状、分布及基体与夹杂的弹性模量的系数。

② 热变形过程（轧制、锻造）中，如果夹杂物具有较高的塑变能力，则随基体一同变形，不会产生附加应力；但如果夹杂物的塑性很低或基本不变形，则钢在变形过程中就会在夹杂与基体之间产生微裂纹。

③ 在外力作用下，由于夹杂与基体之间的弹性模量不匹配而造成局部应力集中。

上述三种局部应力集中叠加的结果会在不同程度上影响材料的力学性能，特别是在交变应力情况下，夹杂周围的应力集中会不断增加，导致微裂纹的萌生。

3）非金属夹杂对钢的应力腐蚀开裂的影响　非金属夹杂对钢的应力腐蚀开裂也有明显影响。实验证明，在某些应力腐蚀条件下，钢产生沿晶裂纹，裂纹的萌生点就在硅酸盐、硫化物等非金属夹杂物附近。

由于冷变形而使大多数非金属夹杂呈条带状沿变形方向分布，这时，施加应力的方向性对应力腐蚀开裂稳定性起着重要的作用。电化学的非均匀性和夹杂周围的应力集中是产生这种差异的直接原因。当夹杂物细小、分散，而且分布与轧制方向无关时，其应力腐蚀中的各向异性效应就缓和或消失。

（4）钢中微量元素对钢的力学性能的影响

任何一种工业用钢都是由特定的若干组元所组成的，每个组元的含量百分比是一定的，即材料的成分是一定的。此外，每个钢种都含有多种不希望存在的元素，这是由于钢在冶炼时受到所用原料、冶炼方法和工艺操作等的影响，难免在钢中残留下少量的元素，如铝、硫、磷、氧、氮、铅、锡、硒、铜、铬、锗、镍等，这些元素不能被视为合金元素，而是属于杂质或残余元素。杂质元素并不都是有害元素，其中大部分是有益的。如碳钢中无意中加入的少量锰、铬、镍、铝等，仍可提高碳钢的强度和韧性。

钢中杂质元素在钢中存在的方式是多种多样的，可以形成固溶体，化合物及游离态。这些杂质存在的方式不同对钢的影响不同。其中以杂质元素在晶间偏聚的危害性最大。杂质元素晶间偏聚能降低晶界的聚合能，当杂质元素在晶界处的含量达到一定的浓度时，会使钢发生脆性断裂，断裂途径

由解理断裂改为沿晶断裂。如回火脆性断裂、低温冷脆断裂、金属脆性断裂、晶间疲劳断裂，晶间应力腐蚀、氢脆等。杂质元素在晶界的临界浓度取决于材料的屈服强度、晶粒尺寸及显微组织。

1）杂质元素与回火脆性　对于低合金钢，最常见的有害杂质元素是周期表中Ⅳ～Ⅵ主族元素（表5.3-1）。这些元素大部分为钢中常见的痕量杂质元素。钢的回火脆性与这些元素的分布有着密切的关系。

表5.3-1　常见脆化元素

Ⅳ族	Ⅴ族	Ⅵ族
Si	P	S
Ge	As	Se
Sn	Sb	Te

关于回火脆性的机理，人们的认识还不十分透彻。普遍认为与碳化物在马氏体界面和原奥氏体晶界的薄片析出有关。痕量杂质元素的存在是产生第一类回火脆性的必要条件之一，第一类回火脆是由碳化物在晶界析出和杂质元素沿原奥氏体晶界偏聚而引起的。这种偏聚是在奥氏体化过程中产生的平衡偏析和回火过程中由于碳化物的排斥而产生的偏析所造成的。杂质元素对第二类回火脆性的影响可归纳成以下几点。

① 如果把杂质完全除掉，钢在回火过程中就不会产生回火脆性，其脆性断裂方式则是解理断裂。如果杂质在基体中，则观察到的脆性断裂方式仍然是解理断裂。但是，一旦杂质开始向晶界偏析，便产生二类回火脆性，断裂方式为沿晶断裂。

② 杂质元素中Sb对二类回火脆性的影响最大，其次是Sn，P。

③ 可以把合金元素对二类回火脆性的影响分为三类：一类与杂质元素同时扩散产生晶间偏析，促进杂质的偏析（如Ni)。另一类自身并不形成偏析，但可促进杂质元素偏析（如Cr)，还有一类元素能把杂质元素钉扎在基体中，从而阻止了杂质元素的偏析（如Ti、Mo)。

2）金属镉引起的钢的脆性断裂　镉并不是结构钢中的合金元素或杂质元素，在一般情况下结构钢中很难发现有镉的存在。镉的来源通常是构件表面镀镉或有外来铺附在构件表面上。过去人们认为钢件表面的固态金属不会引起钢材的脆性断裂（固体金属脆）。后来，高强度钢固体金属脆现象的经常出现，引起了人们的注意，镉脆就是其中之一。

固态镉存在于钢制构件表面时，在外加拉应力的长期作用下，构件表面逐渐产生裂纹。如果温度升高，裂纹的扩展速度增加。裂纹的扩展可以是穿晶解理，也可能是沿晶开裂，但更多的是混合型。当温度升高时穿晶开裂成分减少而沿晶开裂成分增加，裂纹多有分叉现象。

产生镉脆的原因在于固态镉与钢表面紧密接触时，镉原子向钢中扩散，其扩散过程由下式决定：

$$D_s = X^2/2t \tag{5.3-2}$$

式中，t为时间；X为扩散距离；D_s为镉表面的扩散系数。

3）杂质元素对钢的应力腐蚀开裂的影响　可以肯定，大量多数微量元素对钢的应力腐蚀速率均有影响，而这种影响是由多方面因素决定的，改变其中任何一个因素往往引起其他因素作用的大小、程度、甚至性质上的改变。如外界条件不同，钢种不同，钢中所含微量元素的配比不同等，都会影响微量元素的作用。因此，同一元素不同的试验者，在不同的条件下试验，结果可能不一。

4）钢中的白点　钢中白点是钢产生氢脆的一种形式，它是由氢分子气泡造成的，是一种不可逆的氢脆。

钢中的氢来自于冶炼。氢在液态铁中的溶解度很高，钢锭凝固时因溶解度降低而析出，氢来不及逸出钢锭而留在内部。压力加工前加热时，氢又溶入钢中，加工后冷却过程中氢的浓度再次降低。氢原子从固溶体中析出并存留在某些显微空隙处并结合成分子，由此产生很大的压力。这种氢分子析出造成的内应力与相变所产生的组织应力相叠加，并达到较高值时，会造成脆性开裂。这种开裂的断口上往往存在椭圆形的白点斑点。如果钢中存在白点，会引起钢材性能的全面下降。

1.1.3　有色金属基体组织对其失效抗力指标的影响

（1）铝合金的基体组织对其失效抗力指标的影响

铝及其合金在有色金属中是应用最为广泛的一类金属结构材料。它具有高的导热性、导电性、抗蚀性、易于铸造、切削和加工成型。铝合金还有高的比强度和比刚度，是航空工业中最重要的结构材料，在国防、民用建筑、造船、电器以及日常生活用品上都获得广泛的应用。

铝及其合金的力学性能和断裂特性与其冶金因素，合金元素浓度，杂质成分、分布、形状、强化相的组成、大小、分布，晶界形态，热处理状态等都有密切的关系。

1）纯铝的特性　纯铝的屈服强度很低，不适合做承力结构件。但纯铝塑性好，易成型，且导电、导热性好，常用作导电导热材料。纯铝不能通过热处理强化，冷变形是提高纯铝强度的惟一手段。纯铝具有优异的抗腐蚀性能，这是由于铝在空气中易于在表面形成致密的氧化膜，防止了铝的进一步氧化。

纯铝的力学性能和抗腐蚀性能受杂质含量的影响较大。铝中的主要杂质元素是铁和硅，其次有铜、锌、锰、镍和钛等。铁与铝形成针状$FeAl_3$，硅与铝的化合物呈条状或块状，性质硬而脆，使铝的塑性显著降低。铝中铁与硅还可以形成Fe_2SiAl_8和$FeSiAl_{15}$以及亚稳定相$FeAl_6$等。这些杂质相对塑性极为不利，随着铁、硅杂质含量的增加，其强度和硬度也随之提高，但塑性确大幅度下降。

2）冶金因素对铝合金力学性能的影响

① 主要合金元素及其沉淀相对铝合金力学性能的影响

用于铝合金的主要合金化元素为铜、镁、锌、锰、硅及其他微量元素如锆、钛、镍等。由于合金中添加元素的种类和浓度的不同，合金中沉淀硬化相也不相同，这在很大程度上决定了合金的性能。为了获得高强度，必须配合有效的沉淀硬化处理，在沉淀过程中形成均匀、弥散的共格或半格过渡相，使其在基体中造成强烈的应变场，提高对位错的阻力，反之强化效果不高。

二元系铝合金的强度过低，满足不了承力结构件对材料力学性能的要求，为了获得高强度，合金中必须形成具有较大强化效果的强化相。为此，在合金中加入多元合金元素。如硬铝合金，添加元素以Cu、Mg为主。

在硬铝合金中，铜是最主要的合金元素，随着铜含量的增加，在任意镁含量的情况下，强度都随之增加。增加合金中的含镁量也有类似的效果，但必须和铜含量保持适当的比例。在其他系铝合金中，含铜、镁元素的作用基本上和硬铝相同，但随着其含量的变化，并增添不同的元素，其性能则相应略有些变化。

② 杂质相对铝合金力学性能的影响　铝合金中主要杂质元素是铁和硅，铁和硅的含量超过一定限度时就与铝中合金元素形成粗大难溶硬相质点，如在Al－Cu－Mg系合金中，当铁和硅含量大致相等时，有利于形成α（AlFeSi）和Al_6（MnFe）相的粗大质点（20～30 μm），这些相与含铁的其他金属间化合物一样，实际上不溶于固态铝中，本身的断裂强度低，脆性大，在变形及断裂过程中起到裂纹源的作

用。从而降低了塑性，特别是冲击韧性，并使断裂韧性显著下降。降低铁和硅的含量，进一步减少不可溶硬质点的数量可提高断裂韧性。

铝合金中由于含有杂质铁、硅形成粗大难溶硬质点是造成合金塑性，特别是断裂韧性降低的原因。为了提高材料的断裂韧性，对合金中的杂属含量做了严格地限制。

③ 微量元素对铝合金力学性能的影响　为了改善合金的力学性能，常在主要合金元素中添加微量元素，如锰、铬、锆、钛等，这些元素的金属间化合物以细小弥散质点存在时，强烈的影响着合金的再结晶过程，提高再结晶温度和阻止晶粒长大。这样，在热加工和热处理后容易保持非再结晶或部分再结晶的纤维组织，强度也有所提高。

3）热处理和组织状态对铝合金力学性能的影响

① 退火组织对铝合金力学性能的影响　铝合金在低温下均匀退火可得到未再结晶部分再结晶的纤维状组织，并具有均匀细小的第二相质点。高温均匀化退火后为完全再结晶的粗大等轴晶粒。具有未再结晶组织的材料在相同应力状态下具有较高的强度和持久强度，但延伸率和断裂韧性下降。加入抗再结晶元素锰、锆能有效推迟再结晶过程。合金中加入锰时，会使合金中形成含锰的弥散质点，使合金持久强度提高，断裂韧性下降。用锆取代锰会给断裂韧性带来良好的影响。锰、锆同时存在，容易出现锆浓度偏析，阻止晶界流变，对改善合金的韧性有利。

② 时效沉淀相对铝合金力学性能的影响　铝合金借助时效硬化以提高合金的强度。硬铝沉淀相主要为 s′相，超硬铝为 η′相，其尺寸和形态主要受热处理控制。硬铝在自然时效状态下，沉淀相 s′的质点较小，与基体局部共格，位错容易切过，晶内滑移的分布较均匀，具有良好的冲击韧性和断裂韧性，疲劳裂纹的扩展速率也较低。

人工时效，特别是高温长时间时效和过时效后，s′过渡相尺寸增大，强度极限和屈服极限增加，但塑性、韧性和抗蚀性明显下降。但超硬铝一般采用人工时效，而不采用自然时效。这是由于自然时效的时间过于漫长且降低了抗应力腐蚀能力。

③ 晶粒尺寸对铝合金力学性能的影响　晶粒尺寸直接影响塑性变形的均匀性，拉伸时，细晶和粗晶试样在局部塑性变形方面显示出很大差异，而且随着变形温度的提高，这种差异增加。其他铝合金，如硬铝、防锈铝等结果也大致相同。粗晶组织、由于变形波动范围大，变形不均匀必导致局部应力集中，当应力达到一定值时而萌生裂纹。

④ 晶界无析出物区对铝合金力学性能的影响　铝合金晶界无析出物区宽度对力学性能的影响至今尚未取得一致的看法。

一种看法是宽的无析出物区在受力状态下，塑性变形容易集中在晶界两侧的无析出物区，从而形成裂纹和引起沿晶断裂。不过在无析出物区加宽的同时，晶界沉淀相尺寸和数量也加大，甚至形成连续的网膜。因此认为无析出物区加宽而造成塑性降低。另一种观点认为无析出物区加宽能增加塑性，无析出物以将优先发生塑性屈服而引起应力松弛，使合金塑性得以改善。

综合分析，由于无析出物区的变化总是伴随着基体组织和晶间沉淀相的变化。它们的影响相互交织在一起，在不同情况下，无析出物区对性能影响有所不同。短时时效，晶内变形均匀，断裂韧性较高。随时效时间的增加，晶内形成分布不均匀的滑移带，韧性下降，大量滑移带与晶界的交互作用，形成裂纹源。进一步时效，出现明显的无析出物区。晶间断裂的宏观应变过程主要取决于无析出物区与基体间屈服强度之差。两者之间的差别愈大，晶间断裂的倾向就愈高。加入微量元素铬、锆有利于减少这种差别，可减少无析出物区的不利作用。

(2) 钛合金的基体组织对其失效抗力指标的影响

钛及其合金具有密度小（4.5 g/cm³），其密度介于铝（2.7 g/cm³）和铁（7.8 g/cm³）之间，而比强度高于铝和铁，高温和低温性能好，耐腐蚀、无磁性，目前成为航空、航天、化工、造船以及一般民用工业中不可缺少的重要材料。

1）纯钛的特性和断裂失效　钛具有同素异构转变，转变温度为882.5℃，在882.5℃以下为密排六方的 α-Ti，高于882.5℃为体心立方 β-Ti。密排六方 α-Ti 的塑性变形是通过滑移和孪生进行的，滑移系统比其他六方金属多。钛也易于孪生变形，所以纯钛具有优良的塑性变形能力。

高纯钛和工业纯钛经缓慢冷却退火可得到多面体 α 晶粒组织。快冷时发生马氏体相变，其形态变化不大，只是晶界不完整呈锯齿状。工业纯钛在铸造和淬火后得到过饱和 α′相，α′相称为铁马氏体，呈针状。纯钛加热到相变点以上，然后缓慢冷却到室温时得条状 α 组织。

纯钛的力学性能与其显微组织关系不大，而与其纯度关系密切。室温高纯铁具有优良的塑性变形能力，伸长率可达72%，而工业纯钛的塑性随着杂质含量有很大的变化。

2）钛合金的特性和断裂失效

① α 型钛合金　α 型具有优良的低温塑性，好的焊接性以及较高的高温蠕变强度，在低温，超低温下都具有较好的韧性。

② α+β 型钛合金　α+β 型钛合金由其所含成分和浓度的差异，加上不同的热处理制度，其性能变化范围大，适应性好，可以制成综合性能良好的中强钛合金、高强钛合金、耐热钛合金以及高塑性低强度钛合金。因此，工业上的应用最为广泛。与 α 型钛合金相比，α+β 型钛合金性能对组织敏感性强、加热工艺要求严格，但随着合金化程度的提高，焊接性能降低。

③ 介稳定 β 型钛合金　在工业上应用的 β 型钛合金通常是介稳定 β 型钛合金。介稳定 β 型钛合金在固溶处理状态下具有良好的塑性，冷加工成形性能也好。介稳定 β 型钛合金在使用中断裂失效的部位多在焊缝处，这与 β 型钛合金可焊性差有关。采用焊后两次时效，接头性能可获得改善。

3）微量元素对钛合金力学性能的影响　钛合金中常见的有害杂质元素有 H、C、N、O、Fe 等，其中危害最大的是气体元素 H、N、O。此外，钛合金与金属镉接触时，也产生类似钢的金属脆。

① 氢对钛合金力学性能的影响

a）应变时效氢脆　氢在钛合金与局部应力场交互作用，在此处形成局部氢浓度偏聚。金属内部某些局部应力愈大，即畸变或应变愈大的区域，氢的浓度也就愈大。如微裂纹的尖端是氢最富集的地方。在这些地方氢在位错线上沉淀下来，此时位错群好似收集氢的一口深井，同时也是一个陷井，使氢在此处逐步达到较高的浓度，位错与氢混合形成气团。氢对位错起到了钉扎的作用，使位错不能再自由的运动，于是产生了局部硬化。在外力作用下，局部硬化区不易通过位错的移动产生塑性变形，应力得不到松弛，由此产生脆性断裂。此种应变时效氢脆也称为可逆氢脆，当应力消除时，上坡扩散停止，在热扩散的作用下，部分已经偏聚了的氢可能扩散到其他氢浓度相对较低的地方，从而消除了产生氢脆的危险。

b）钛合金中的氢化物氢脆　如果合金中含有不太高的氢原子，室温下合金的平衡组织为氢固溶体及氧化物两相。但是合金在淬火后形成了过饱和的氢固溶体。由于加载荷之

前试样并不存在氢化物。因此，在高速变形时不会发生。如果这种淬火状态的铁合金，经低应力慢速变形后或在静载长期作用下，过饱和固溶体中的氢原子析出，并形成 TiH，造成氢脆。此种氢脆是不可逆的。

② 氧、氮对钛合金力学性能的影响　在高温下，氧、氮在 β-Ti 中有较高的溶解度，并与基体形成间隙式固溶体。在冷却过程中这些间隙杂质元素的含量对 β 固溶体分解动力学有着强烈的影响，使时效后合金的强度大为增加，而塑性、冲击韧度、断裂韧度显著下降，并出现低温脆断。

③ 镉对钛合金力学性能的影响　钛合金产生镉脆的条件有三，其一是镉必须与钛合金紧密接触（如两者之间存在压应力）两者之间不应有其他物质隔离；其二是需要持久的拉应力存在，产生镉脆的第一步就是在拉应力作用下镉向钛合金中扩散；其三是只有变形钛合金才产生镉脆。这种现象的原因至今还不清楚。

(3) 镁合金的基体组织对其失效抗力指标的影响

镁合金的主要优点是密度小（1.78 g/cm³），易于切削加工，而且具有比较高的比强度、比刚度，较好的减振性和热加工成形性。

纯镁的强度和延性都很低，屈服强度仅有 60～150 MPa，而且耐腐蚀性能极差。纯镁在空气中形成氧化膜，但这种氧化膜不致密，腐蚀介质可以通过氧化膜中的空隙渗入表面，氧化膜很脆、易脱落。因此氧化膜的保护性极差，这是造成镁耐蚀性差的基本原因。纯镁除某些特殊用途之外，不能用来作为工程结构材料。

镁的合金化主要是固溶强化和通过时效处理所造成的沉淀硬化来提高合金的强度。在这一点上与铝合金一致，通过合金化可以使镁合金的屈服强度增加好几倍。主要合金元素是铝、锌、锆、稀土。在一定的含量范围内，这些元素可以与镁形成固溶体。并有强烈的固溶强化效应。这些元素的固溶度随温度变化明显。时效过程中析出，并形成有明显强化效果的第二相。

镁合金中的冶金缺陷和热加工缺陷严重影响着其力学性能，同时也降低其抗应力腐蚀能力。镁的化学活性高、热加工变形性相对较低，使得镁合金加工过程中易于受到污染和产生各种缺陷。

(4) 铜合金的基体组织对其失效抗力指标的影响

铜及其合金具有优异的导电性和导热性，良好的力学性能和抗腐蚀性能。由于塑性高，很容易进行各种加工。因此铜及其合金是电器、仪表、化工、机械等工业部门中的重要金属材料。

纯铜也称紫铜，其导电、导热性能仅次于银，优于其他任何工业金属。纯铜的强度低，耐蚀性也差，作为工程结构材料，在使用上受到限制。

铜锌合金称为黄铜，铜中几乎可溶 40% 锌，形成 α 固溶体，α 黄铜的强度、延性都很好，适合于剧烈的冷加工成形。

α 黄铜易发生两种腐蚀：脱锌和季裂。

α 黄铜脱锌是一种气蚀现象，黄铜与海水、含氧及二氧化碳较高的淡水接触时，会产生合金中锌的脱溶现象，使零件表面上残留一层多孔的纯铜，呈海绵状。如果这种选择性的腐蚀不能被及时阻止，则将深入合金横断面的深处，最后产生断裂。

黄铜的季裂是零件在冷加工后，消除应力退火不完善，在潮湿大气，特别是含氨盐的大气及汞盐溶液中产生的应力腐蚀开裂。消除季裂的方法是零件冷加工后及时退火(260～300℃)，以消除残余拉应力或采用零件表面镀锌、锡保护。

铜能与多种合金元素形成固溶体，而且溶解度都很大，固溶强化效果明显，并且有优良的耐腐蚀性能。如磷青铜，其强度高、韧性好、抗蚀性良好，摩擦因数小，并且不产生季裂。良好的综合性能使磷青铜广泛地被用来制造波纹管、膜片及弹簧等。

铍青铜可进行时效强化，其在固溶状态具有良好的可加工性、成形性，时效后又有良好的力学性能，物理性能和耐蚀性。因此，铍青铜广泛地用于要求耐腐蚀、高强度和相当高的电导率的零件，如接触电桥、特殊性能的螺栓，还用来生产耐磨性能良好的零件，如撞针、无火花工具。

随着杂质元素的增加，铜合金的耐磨蚀性能全面下降，杂质元素显著地降低材料的塑性、韧性。

铜中常见的杂质元素有铝、镍、锡、铋、硫、铅、磷、氧、氢等。这些元素大部分（如 Al、Ni、Sn、S、P）在铜中的溶解度都很大，它们对铜的力学性能的影响很小。但其中的铋和铅在铜中的溶解度极小，几乎不能溶入固态铜内，而是在低温形成共晶，这些共晶体多以极薄的薄膜分布在基体的晶界上。当进行热加工时，共晶熔化或强度极低，造成合金的热脆性。因此，需热变形的铜合金对铋和铅的含量应严格控制。

氧、硫与铜形成 Cu_2S 及 Cu_2O，它们是脆性化合物，这些脆性化合物使材料冷变形发生困难。

氧含量较高的铜如果在还原性气氛中退火，易发生氢脆，氧化铜被氢还原后生成的水蒸气、二氧化碳等在铜内形成高压气室，促成构件的早期失效。

1.1.4　金属材料各种失效抗力指标间的换算关系

金属材料的各种失效抗力指标的大小在实际上存在一定的对应关系，研究者进行了许多探讨。其中主要是根据金属的硬度评价其他的失效抗力指标。这主要是由于金属的硬度试验简便、迅速，人们很早就探讨了如何通过所测定的硬度值来评定金属的其他力学性能指标。这无论对于实际试验或对于材料的科学研究，都具有十分重要的意义。

至今还没有从理论上确定金属材料的硬度与其他的力学性能之间的内在联系，只是根据大量的试验确定了硬度与某些力学性能指标之间的关系。

试验证明，金属材料的压入硬度与抗拉强度之间成正比关系，即

$$\sigma_b = k\text{HB} \tag{5.3-3}$$

式中 k 为比例系数，不同的金属材料其 k 值不同；同一类金属材料经不同热处理后，硬度和强度发生变化，其 k 值基本保持不变。但若经过冷变形提高硬度，则 k 值不再是常数。

关于不同金属材料的 k 值有许多经验数据。对于钢铁材料可粗略认为 $k \approx 3.3$。

我国计量科学院等单位通过大量试验和分析研究已经指定出黑色金属硬度及强度换算表，如表 5.3-2 所示，以及铝合金和铜合金的换算表，具有一定的实用价值。但若要求得到较精确的强度数据，还需通过拉伸试验测定。

有人设想找出疲劳极限与硬度之间近似的定量关系，试图通过测定金属材料的硬度 HB 估算 σ_{-1}，但至今尚未取得理想的结果。

疲劳极限与抗拉强度值之间的关系式为

$$\sigma_{-1} = m\sigma_b \tag{5.3-4}$$

对于不同的金属材料、不同的试验条件，其 m 值也不同。一般来说 $m \approx 0.4 \sim 0.6$，平均为 0.5，即 σ_{-1} 大致相当于 σ_b 的一半。对于钢铁材料，已知 σ_b 约为 HB 的 3.3 倍，因此，σ_{-1} 就大约为 HB 的 1.6 倍左右，即 $\sigma_{-1} \approx 1.6\text{HB}$。表 5.3-3 列出了某些退火金属材料的 HB、σ_b 与 σ_{-1} 的实验数据。由表 5.3-3 可见，黑色金属基本满足上述关系。

表 5.3-2 钢的硬度与强度换算表

洛氏硬度		表面洛氏硬度	布氏硬度		维氏硬度 HV	抗拉强度/MPa					抗拉强度近似值 /MPa （不分钢种）
HRC	HRA	HR15N	压痕直径 d10/3 000/mm	HB		碳钢	铬镍钢	铬镍钼钢	铬锰硅钢	超高强度钢	
70	86.6				1 037						
60	86.1				997						
68	85.5				959						
67	85.0				923						
66	84.4				889						
65	83.9	92.2			856						
64	83.3	91.9			825						
63	82.8	91.7			795						
62	82.2	91.4			766						
61	81.7	91.0			739						
60	81.2	90.6			713					2 639.1	2 556.7
59	80.6	90.2			688					2 508.6	2 447.8
58	80.1	89.8			664					2 390.9	2 344.9
57	79.5	89.4			642					2 281.1	2 248.7
56	79.0	88.9			620					2 181.1	2 158.5
55	78.5	88.4			599		2 057.5		2 045.7	2 089.9	2 074.2
54	77.9	87.9			579		1 985.9		1 971.2	2 005.5	1 994.7
53	77.4	87.4			561		1 917.3	1 946.7	1 900.6	1 929.0	1 919.2
52	76.9	86.8			543		1 850.6	1 881.0	1 833.9	1 857.4	1 848.6
51	76.3	86.3	2.73	510	525		1 785.9	1 818.2	1 769.2	1 791.7	1 781.9
50	75.8	85.7	2.77	488	509	1 710.3	1 724.1	1 758.4	1 708.4	1 730.9	1 719.2
49	75.3	85.2	2.81	474	493	1 653.4	1 665.2	1 699.6	1 650.5	1 674.1	1 659.3
48	74.7	84.6	2.85	461	478	1 699.5	1 608.3	1 643.7	1 595.6	1 620.1	1 603.4
47	74.2	84.0	2.886	449	463	1 550.4	1 553.4	1 588.7	1 542.6	1 569.1	1 550.5
46	73.7	83.5	2.927	436	449	1 503.4	1 501.5	1 536.8	1 492.6	1 520.1	1 499.5
45	73.2	82.9	2.967	424	436	1 459.3	1 451.4	1 486.7	1 445.6	1 473.0	1 451.4
44	72.6	82.3	3.006	413	423	1 417.1	1 403.4	1 438.7	1 399.5	1 426.9	1 406.3
43	72.1	81.7	3.049	401	411	1 347.9	1 358.3	1 392.6	1 357.3	1 381.8	1 362.2
42	71.6	81.1	3.087	391	399	1 340.6	1 314.1	1 348.5	1 316.1	1 355.7	1 321.0
41	71.1	80.5	3.130	380	388	1 305.3	1 272.9	1 305.3	1 276.9	1 289.6	1 281.8
40	70.5	79.9	3.171	370	377	1 271.0	1 232.7	1 265.1	1 239.6	1 242.5	1 243.5
39	70.0	79.3	3.214	360	367	1 238.6	1 195.5	1 225.9	1 204.3	1 194.5	1 208.2
38	69.5	78.7	3.258	350	357	1 207.2	1 159.2	1 188.6	1 171.0	1 144.5	1 173.9
37	69.0	78.1	3.299	341	347	1 176.8	1 125.8	1 153.3	1 138.6	1 092.5	1 140.6
36	68.4	77.5	3.343	332	338	147.4	1 093.5	1 119.0	1 108.2	1 038.6	1 109.2
35	67.9	77.0	3.388	323	329	1 119.0	1 063.1	1 086.6	1 079.8	980.7	1 078.8
34	67.4	76.4	3.434	314	320	1 091.5	1 033.7	1 056.2	1 052.3	—	1 049.3
33	66.9	75.8	3.477	306	312	1 065.0	1 007.2	1 026.8	1 052.8	—	1 021.9
32	66.4	75.2	3.522	298	304	1 039.5	981.7	998.4	1 000.3	—	955.4

续表 5.3-2

洛氏硬度		表面洛氏硬度	布氏硬度		维氏硬度HV	抗拉强度/MPa					抗拉强度近似值/MPa（不分钢种）
HRC	HRA	HR15N	压痕直径 d10/3 000/mm	HB		碳钢	铬镍钢	铬镍钼钢	铬锰硅钢	超高强度钢	
31	65.8	74.7	3.563	291	296	1 014.0	957.2	971.9	976.8	—	969.9
30	65.3	74.1	3.611	283	289	989.5	934.6	947.4	954.2	—	945.4
29	64.8	73.5	3.655	276	281	965.0	912.1	922.8	932.6	—	921.0
28	64.3	73.0	3.701	269	274	942.5	894.4	900.3	912.1	—	899.7
27	63.8	72.4	3.741	263	268	918.9	875.8	879.7	892.4	—	877.7
26	63.3	71.9	3.783	257	261	896.4	859.1	859.1	874.8	—	857.1
25	62.8	71.4	3.826	251	255	874.8	843.4	—	857.1	—	837.5
24	62.2	70.8	3.871	245	249	853.2	828.7	—	839.5	—	818.9
23	61.7	70.3	3.909	240	243	832.6	815.0	—	823.8	—	800.3
22	61.2	69.8	3.957	234	237	813.0	803.2	—	809.1	—	783.6
21	60.7	69.3	3.998	229	231	793.4	791.4	—	794.4	—	766.9
20	60.2	68.8	4.032	225	226	774.8	781.6	—	780.6	—	752.2
19	59.7	68.3	4.075	220	221	756.1	772.8	—	766.9	—	737.5
18	59.2	67.8	4.111	216	216	738.5	764.0	—	754.2	—	722.8
17	58.6	67.3	4.157	211	211	721.8	757.1	—	742.4	—	710.0
16	58.1	66.8	4.19	208		706.1	750.2	—	730.6	—	697.3
15	57.6	66.4	4.21	205		690.4	744.4	—	719.8	—	685.5
	57.1	65.9	4.25	201							674.7
	56.6	65.5	4.28	198							663.9
	56.1	65.0	4.32	195							654.1

表 5.3-3　退火金属的 HB、σ_b 与 σ_{-1} 的关系

金属及金属名称		HB	σ_b/MPa	$k=\frac{\sigma_b}{HB}$	σ_{-1}/MPa	$\alpha=\frac{\sigma_{-1}}{HB}$
有色金属	铜	47	220.30	4.68	68.40	1.45
	铝合金	138	455.70	3.30	162.68	1.18
	硬铝	116	454.23	3.91	144.45	1.24
黑色金属	工业纯铁（0.02%C）	87	300.76	3.45	159.54	1.83
	20 钢	141	478.53	3.39	212.66	1.50
	45 钢	182	637.98	3.50	278.02	1.52
	T8 钢	211	753.42	3.57	264.30	1.25
	T12 钢	224	792.91	3.53	338.78	1.51
	1Cr18Ni9	175	902.28	5.15	364.56	2.08
	2Cr13	194	660.81	3.40	318.99	1.64

此外，也有人利用硬度试验间接测定屈服强度、评定钢的冷脆倾向，以及借助特殊硬度试样近似地建立真实应力－应变曲线等。这些将在金属材料检验及力学性能研究中得到应用。

1.2　断裂失效预测

机械构件的主要失效形式是断裂、腐蚀及磨损。美国在1983年估计这三类失效方式造成的当年国民经济的损失都超过1 000亿美元，断裂是首当其冲的，年损失近1 400亿美元。

在机电装备的各类失效中以断裂失效最主要，危害最大。因此，国内外对断裂失效进行了大量的分析研究，迄今为止，断裂失效的分析与预防已发展为一门独立的边缘学科。目前对断裂行为的研究有两种不同的方法。一种是断裂力学的方法，它是根据弹性力学及弹塑性理论，并考虑材料内部存在有缺陷而建立起来的一种研究断裂行为的方法。另

一种是金属物理的方法，从材料的显微组织、微观缺陷、甚至分子和原子尺度上研究断裂行为的方法。而断裂失效分析则是从断口的宏观、微观特征入手，研究断裂过程和形貌特征与材料的性能显微组织、零件的受力状态及环境条件之间的关系，从而揭示断裂失效的原因。它在断裂力学方法和金属物理方法之间架起联系的桥梁。

根据零件断裂前所产生的宏观塑性变形量的大小分为：塑性断裂，断裂前发生较明显的塑性变形；脆性断裂，断裂前几乎不产生明显的塑性变形。通常材料的塑性变形小于2%～5%（横截面积）的断裂均可称为脆性断裂；塑性－脆性混合型断裂，又称为准脆性断裂。当变形在5%～10%范围内出现的断裂，均可称为塑性－脆性混合型断裂。塑性断裂对装备与环境造成的危害远较脆性断裂小，因为它在断裂之前出现明显的塑性变形，易引起人们的注意。与此相反，脆性断裂往往引起危险的突发事故。在断裂失效事故中，主要危险是脆性断裂事故。

上述的断裂方式分类尚嫌粗放，特别是对脆性断裂，因为这样分类时，脆性断裂方式中除低应力脆断、应力腐蚀断裂等之外，还包括了在断裂事故中占很重要地位的疲劳断裂事故。疲劳断裂的断口部位在宏观上也没有明显塑性变形，断口表面也比较平齐，但是因为这类断裂事故发生较多，断裂过程又有特殊性，故在工程界常把这种断裂方式不列入脆性断裂范围。以下讲的脆性断裂，也不包括疲劳断裂内容，只讲比较狭义的脆性断裂。

脆性断裂（狭义）包括以下两类。

1）低应力脆断　是指通常在弹性应力范畴，在许用应力条件下一次加载引起的脆性断裂事故。这类事故的发生常常有外在原因或有内部原因，外在原因包括受载时的加载速率，环境温度对材料性能的影响，零部件或构件形状设计中引起的应力集中等。材料的内部原因包括材料内部存在的宏观缺陷（裂纹、空洞、大块夹杂、缩松等），材料本身的质量问题，如钢的低温脆性、蓝脆、回火脆、不锈钢中的475℃脆，σ相脆性等。

2）环境介质条件下的脆断　实际上也是一种低应力脆断事故，在上述情况下增加环境介质因素，是指构件在受载情况下同时接触到会使材料性能或表面状态恶化的环境条件，例如潮湿空气、水介质、熔盐、硫化氢气氛、融熔金属、辐照环境等。这种环境介质条件下的断裂事故可分为：应力腐蚀开裂、氢致开裂、融熔金属脆、碱脆、辐照脆等几种。这类断裂事故的发展过程包含了时间因素，因之与疲劳断裂在这一点上有相似之处，都有一个裂纹形成和发展的时间过程，都不是受载后立即发生断裂，属滞后破坏断裂范畴，这与上述受载后立即断裂的低应力断裂事故是有差别的。

低应力脆断事故发生带有突发性，事故前预兆监测通常很困难，发生预兆到事故发生时间很短，因之这类事故常常造成灾难性后果。环境介质脆断事故，大都属于滞后断裂问题，事故的发展有一个过程，通过监测或定期检查，观察裂纹的亚临界扩展，在达到临界裂纹长度以前及时更换零部件或采取适当措施，对防止恶性事故发生还有一定的可能性，但由于这类亚临界裂纹扩展的距离通常很短，故预防监测难度很大。

1.2.1　金属材料的脆性断裂失效

如前所述，金属材料的断裂是金属材料失效的一种主要形式，引起的破坏最大，且比较难加以控制和预测预防。金属材料的断裂可分为以下几种类型。

（1）按照构件断裂前所产生的宏观塑性变形量的大小分

1）塑性断裂，断裂前发生较明显的塑性变形。

2）脆性断裂，断裂前几乎不发生明显的塑性变形。通常材料的塑性变形小于2%～5%的断裂均可称为脆性断裂。

3）塑性－脆性混合型断裂，又可称为准脆性断裂。当变形在5%～10%范围内出现的断裂，均可称为塑性－脆性混合型断裂。断裂后的断口上可观察到解理、准解理与韧窝等多种形貌特征。

（2）按照断裂路径的走向分

1）穿晶断裂。裂纹穿过晶粒内部进行扩展。穿晶断裂可以是塑性的，也可以是韧性的。这两者的主要区别是前者的断口形貌以韧窝为主而后者以解理花样为主。

2）沿晶断裂。裂纹沿着晶粒的边界进行扩展，可分为沿晶脆断和沿晶韧断（在晶界面上有浅而小的韧窝）。

（3）按照断裂方式分

1）正断断裂。受正应力引起的断裂，其断口表面与最大正应力方向垂直。断口宏观形貌较平整，微观形貌有韧窝、解理花样等。

2）切断断裂。在切应力的作用下引起的断裂。断面与最大正应力方向成45°角，断口宏观形貌较平滑，微观形貌为抛物面状的韧窝花样。

3）混合断裂。正断与切断两者相混合的断裂方式，断口呈杯锥状，混合断裂是最常见的断裂类型。

（4）按照断裂机理分

可分为解理、准解理、韧窝、滑移分离、沿晶及疲劳等多种断裂，在本篇第2章中有详尽的介绍。

这些分类方法在失效分析工作中是同时使用的，并且彼此之间没有固定的对应关系。例如一个构件的脆性断裂的断口可能是穿晶断裂，也可能是沿晶断裂；可能发生了正断断裂，也可能发生切断断裂；可能呈现解理形貌，甚至也可能呈现韧窝形貌等。这是由于各种分类方法的出发点不同造成的。

塑性断裂对装备与环境造成的危害远小于脆性断裂，因为它在变形前出现明显的塑性变形，易引起人们的注意。与此相反，脆性断裂往往引起危险的突发事故。

下面主要介绍各种脆性断裂失效的预测方法和技术。

1.2.2　环境导致的脆断失效预测

通常把材料在负载条件下，由于环境的原因而发生的低应力延迟的破断失效，通称为环境破断失效或环境诱导开裂（Environment Assisted Cracking，EAC）。它的力学行为特征，一是延迟（或滞后）断裂，经过一段时间的孕育后发生失效，一是低应力脆断，名义应力往往远低于材料的屈服强度，破断前没有明显的塑变。断口微观形貌特征，则因环境条件及应力大小的不同而有很大差别，可以是沿晶断口，也可以是解理、准解理断口，有时也会出现浅的韧窝形貌。在有腐蚀的环境条件下，断口常黏有腐蚀产物。

环境脆断失效是应力与环境因素对特定材料综合作用的结果，如图5.3-1所示。因此，环境破断失效的机理研究，往往需要综合应用几门学科的理论。

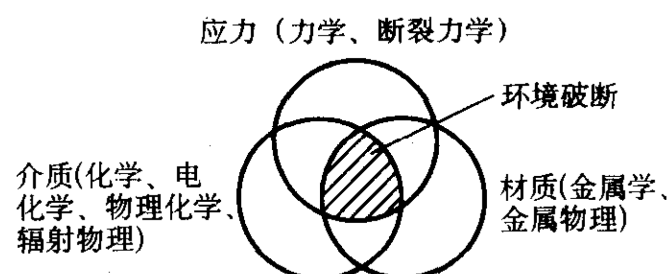


图5.3-1　环境破断发生的条件与涉及的学科

环境破断失效通常包括应力腐蚀断裂（应力腐蚀有两种含义。一种为广义的应力腐蚀，它包括以阳极溶解为控制过程的应力腐蚀和以阴极吸氢为控制过程的氢脆；一种是指狭义的应力腐蚀，即是指以阳极溶解为控制过程的应力腐蚀）

(Stress Corrosion Cracking, SCC)、氢脆(Hydrogen Embrittlement, HE)、腐蚀疲劳(Corrosion Fatigue, CF)及液态金属脆化(Liquid Metal Embrittlement, LME)、辐照脆化(Irradiation Embrittlement)等。本章将分别讨论这些环境破断失效方式，但不述及温度条件下的蠕变失效。

(1) 应力腐蚀断裂

1) 应力腐蚀断裂现象　金属材料在特定介质环境中，在拉应力作用下，经过一定时间后发生开裂及断裂的现象称为应力腐蚀断裂。由于近代工业技术的发展，材料工作的环境条件越来越苛刻，零件及材料的应力腐蚀问题也日益突出，应力腐蚀与低应力脆断、疲劳断裂并列为当今工程断裂事故的三种主要形式。

应力腐蚀断裂发生的条件和特征可以归纳如下。

① 特定的腐蚀介质与材料的组合　一定的材料只有在与一定的介质环境组合时才会发生应力腐蚀。表5.3-4列出了一些常用材料发生应力腐蚀的环境介质。

表5.3-4　发生应力腐蚀的材料－介质组合

材料	介　质
铜合金	NH_3蒸气或NH_3溶液；水；水蒸气；$AgNO_3$；湿H_2S；水银；$FeCl_3$；含氮有机物；柠檬酸；石碳酸
Au－Cu－Ag	$FeCl_3$水溶液
铝合金	($NaCl+H_2O_2$)水溶液；NaCl水溶液；海水；水银；($NaCl+NH_4Cl$)水溶液；($CaCl_2+NH_4Cl$)水溶液，工业大气
碳钢和低合金钢	NaOH水溶液；NH_3水溶液；硝酸盐水溶液；H_2S水溶液；含HCN水溶液；NH_4Cl水溶液；($NaOH+NaSiO_3$)水溶液；海水；CH_3COOH水溶液；$CaCl_2$水溶液；$FeCl_3$水溶液；$(NH_4)_2CO_3$水溶液；碳酸盐和重碳酸盐水溶液；混合酸($H_2SO_4+HNO_3$)水溶液；熔融Zn·Li或Na·Pb合金；潮湿的含CO·CO_2空气；海洋大气或工业大气
镁合金	Na_2SO_4；氟化物；水；(氯化物$+K_2CrO_4$)水溶液；热带工业大气或海洋大气
镍基合金	熔融苛性碱；浓NaOH(260～427℃)；HF酸；浓缩锅炉水(260～427℃)；硅氧氟酸；含氧及痕量铅的高温水；熔融铅；浓Na_2S水溶液；含SO_2的水蒸气
Ni－Cr奥氏体不锈钢	严重污染的工业大气；湿润空气(湿度＞90%)；水蒸气(260℃时)；海洋大气；($NaCl+H_2O_2$)水溶液；H_2S水溶液；(NaOH＋硫化物)水溶液；热浓碱；($CuSO_4+H_2SO_4$)水溶液；(H_2SO_4＋氯化物)水溶液；浓缩锅炉水；H_2SO_4(260℃时)；过氯酸钠；20%～50% $CaCl_2$水溶液；酸式亚硫酸盐；硫胺饱和溶液；($NaHCO_3+NH_3+NaCl$)水溶液；粗苏打和硫化纸浆；明矾水溶液；二氯乙烷；邻二氯苯；体液(汗液，血清)；甲基三聚氰胺；联苯、二苯醚；氯乙醇加水；聚连多硫酸H_2SmO_4($m=2～5$)
马氏体不锈钢	工业大气；海水；氯化物；酸性硫化物
铁素体不锈钢	海洋大气；工业大气；海水；高温水；水蒸气；氯化物溶液；($NaCl+H_2O_2$)水溶液；NaOH水溶液；NH_3水溶液；H_2S水溶液；($H_2SO_4+HNO_3$)水溶液；硝酸盐；硫酸；硝酸；高温碱

续表5.3-4

材料	介　质
钛及钛合金	H_2；甲醇蒸气；CCl_4；溴、氟蒸气(－196℃)；发烟硝酸；汞；Cd(＞327℃时)；Ag(＞466℃)；AgCl(371～482℃)；Ag－5Al－2.5Mn(343℃)；海水；氯化物盐(288～427℃)；甲醇、乙醇、乙烯二醇
锆及锆合金	甲醇、乙醇；氟利昂；CCl_4；卤族元素及其蒸气；Hg；Cs；氯化物水溶液；热盐；NaCl(300～350℃)。LiCl＋KCl(300～350℃时)，KNO_3+NaNO_3+KCl(或KI、KBr)(200℃)

有时在介质浓度很低的条件下也会发生应力腐蚀断裂。通常，介质浓度和环境温度越高，越易发生应力腐蚀。

② 拉应力的存在　拉应力是发生应力腐蚀的必要条件之一。拉应力可以是工作载荷引起的，也可以是装配应力及材料的残余应力。通常应力越大，发生应力腐蚀开裂的时间越短。

③ 材料纯度和组织状态的影响　通常认为纯金属不会发生应力腐蚀，但存在极少杂质时，也有可能发生应力腐蚀。如纯铜含0.001%杂质，纯铁含0.004%杂质时，观察到出现应力腐蚀现象。材料的组织状态对应力腐蚀的敏感性影响很大。通常，稳定的组织对应力腐蚀敏感性较小，例如在H_2S气氛中工作的碳钢，一般规定硬度不得超过220HB，当硬度达到＞250HB时，即明显地发生应力腐蚀现象，硬度越高(即组织越不稳定)则应力腐蚀敏感性就越大。

④ 应力腐蚀的开裂时间　应力腐蚀断裂是一种延迟断裂破坏。其裂纹的扩展速度比无应力情况的纯腐蚀要快的多。裂纹在应力和介质的综合作用下先发生亚稳态扩展，当扩展到临界裂纹长度时，即发生失稳扩展而快速断裂。通常，应力腐蚀裂纹的亚稳态扩展速度为0.001～1 cm/h。结构钢、不锈钢、铝合金一般为0.1 cm/h以下，而钛合金、高强钢则可达1 cm/h。除材质因素以外应力腐蚀裂纹扩展的速度还取决于介质类型、浓度、温度和应力大小。

2) 应力腐蚀断裂的基本规律

① 应力腐蚀应力强度因子门槛值(K_{ISCC})和应力强度因子门槛值(K_{IC})　早期的应力腐蚀试验多是采用光滑试样在介质和恒载荷的作用下，测定其发生断裂的时间。研究发现，当试样承受的应力小于某一值时，不发生开裂或发生开裂的时间很长，此应力值即为发生应力腐蚀的应力门槛值或应力腐蚀的临界应力。应力门槛值与温度、介质浓度及材质有关。

断裂力学应用到应力腐蚀断裂的试验研究中后，则用裂纹试样进行应力腐蚀试验，对大多数的金属环境组合均有如下现象，即应力强度因子低于某一门槛值K_{ISCC}后，裂纹不再扩展，即不发生应力腐蚀开裂。实际实验测定时，一般以100 h(或1 000 h)不发生断裂的应力强度因子K_I值作为应力腐蚀门槛值K_{ISCC}。通常认为，当$K_{ISCC}/K_{IC}\geqslant 0.6$时，材料对应力腐蚀不敏感。在使用$K_{ISCC}$时，必须注意其时间条件。

② 关于应力腐蚀断裂的机理　应力腐蚀开裂是局部腐蚀和应力共同作用下而发生的复杂的物理－化学过程。关于它的机理有各种理论，迄今比较公认的有阳极溶解作为控制过程的电化学机理和阴极吸氢为控制过程的氢脆机理。

③ 材料组织结构对K_{ISCC}的影响　应力腐蚀断裂应力强度因子门槛值K_{ISCC}是一个组织结构敏感的参量。同种材料，当处理到不同强度级别时(即不同组织状态时)，强度越高，则K_{ISCC}越低，其K_{ISCC}/K_{IC}的比值也越低，即应力腐蚀敏感

性增高。相同强度级别而不同材料的 K_{ISCC} 值也不同。通常，低碳钢板条马氏体组织的抗应力腐蚀能力比中高碳钢的针状马氏体组织要高。

(2) 氢脆

金属的氢脆现象早已为人们所关注。氢进入金属后，一般都使材料的性能变坏。金属中的氢可以是在材料生产工艺过程中吸收的，例如冶炼及焊接时液态金属吸收的氢保留在材料中，以及酸洗、电镀中材料吸收的氢，也可能是材料在氢环境下服役而吸收的氢，例如高压氢容器以及潮湿空气、水介质、碳氢化合物、酸、硫化氢等工况环境，因而氢损伤能在各种环境下发生。近代由于能源工业、化学工业的高速发展，使得吸氢环境增多，使金属的氢脆成为日益突出的问题之一。如高强钢在水介质中的应力腐蚀开裂已经公认是氢脆机理，因此金属的氢脆问题日益引起人们的关注。在材料科学和工程中，氢脆是一个十分活跃的领域。

氢损伤或氢脆的表现形式很多，一般可以分为不可逆氢脆和可逆氢脆两类。前者是指氢造成的永久性损伤，如低强度钢在石油化工设备中或酸洗处理后因吸氢而出现的氢致鼓袍，炼钢时未采取脱氢措施或后处理而在大型锻、轧件中出现白点、发纹等，可逆氢脆是指排除氢后力学性能可以恢复的氢脆现象，钢及一些合金中的氢致延迟破坏（或称静疲劳破坏）属于这一类。我们主要讨论后一种氢脆问题。

金属特别是高强钢的氢脆一般表现为低应力下的延迟破坏，当应力低于门槛应力值 σ_{th} 时，不再发生氢断裂。温度是氢脆的重要影响因素，高强钢氢脆通常发生在 $-100 \sim 100℃$ 的温度范围内，由于氢在更高的温度会逸出及材料的升温软化，则不再显示氢脆；而在更低的温度下，由于氢的扩散速率慢，不足以形成局部偏聚，也不会显示氢脆。在室温附近，氢脆的敏感性最大。应变速率也影响氢脆的敏感性。高应变速率（冲击、爆炸成形、高速拉伸）下，溶解氢来不及通过扩散形成局部偏聚，材料就断裂了，故反映不出塑性的下降。在低应变速度下，氢脆敏感性明显增加，表现出塑性的下降及低应力延迟破坏。材料中氢的浓度越高，越容易生局部偏聚，则氢脆现象越容易发生。

不同的材料氢脆敏感性不同，同种材料而组织结构不同，其氢脆敏感性也各不相同。钢的氢脆主要表现为塑性指标（δ、ψ）的大为降低，但其强度基本无明显变化，成分及组织结构对氢脆的影响规律与应力腐蚀相同。过去认为只在体心立方金属中发生氢脆，但近年来的大量研究发现，铝、铁及奥氏体钢中同样存在氢脆问题，非晶态金属对氢脆还特别敏感。

关于氢脆的机理，目前有多种学说。氢压理论，认为金属的氢在缺陷处以分子态聚集，形成高压气体而造成鼓泡、白点、发纹等缺陷，由此造成材料脆化；氢化物理论认为氢与金属形成化合物而造成脆化，如 Ti、Zr、Nb 等合金在氢环境中形成氢化合物而脆化，一般认为钢中不会形成氢化物。减聚力理论认为，固溶氢使金属原子间的结合力减弱而使金属变脆。以上各种意见可以说明不同情况下的氢损伤或氢致开裂的原因。

对固溶氢导致氢脆的现象，多半是因为使金属中位错增殖，并促使位错的运动，氢和位错的交互作用使得裂纹前沿塑性区及三向应力集中区发生氢富集，富集的氢引起原子结合力下降，在一定应力状态下，因氢浓度达到临界值而致裂纹扩展。裂纹扩展致使裂纹尖端氢浓度降低，使裂纹扩展停止。在新形成的高应力应变区再度吸引氢。当富集到临界氢浓度时裂纹再扩展，裂尖氢浓度又降低，出现止裂，这过程不断反复，在主裂纹扩展到临界尺寸 α_0 时失稳。

钢中的白点（即髮裂）是钢被中氢浓度超过 2×10^{-6} 后，钢在冷却凝固时，奥氏体状态还可固溶较多的氢，转变为铁素体时氢便要析出，于是在相变组织应力的协同作用下析出的原子氢，在钢中缺陷处结合成分子氢增大压力，出现白点（即髮裂）。

氢鼓泡是钢材在高温氢环境中工作，环境中氢被钢材吸收，与钢中的碳形成甲烷，在缺陷处因气体体积膨胀产生压力，导致工作表面层发生鼓泡。钢中如有碳化物组元这类合金元素，可抑制甲烷形成。

还有一类氢损伤是在高温裂解炉管中发生的，因生成甲烷引起脱碳，使炉管强度降低，失效方式为蠕变断裂。改善措施是增加合金中的碳化物含量，形成合金碳化物防止甲烷形成。

以上所说的白点（髮裂）、氢鼓泡、氢脱碳等造成的氢损伤都是氢造成的不可逆损伤，而延迟氢致开裂则是可逆的。

(3) 腐蚀疲劳

金属材料在腐蚀介质和交变应力联合作用下发生的低应力开裂即为腐蚀疲劳失效。广义地说，除了在真空和纯氮或其他惰性介质环境下的疲劳破坏，绝大多数的疲劳破坏都可以认为是腐蚀疲劳。已为实验证实，即使在一般的空气环境中，材料的疲劳寿命也比惰性环境下要低得多。

腐蚀疲劳与应力腐蚀不同，其区别不仅在于加载方式的不同（交变应力和静应力），而且腐蚀疲劳不是像应力腐蚀只限于某些材料、介质的组合，而是几乎在所有的腐蚀环境中都可发生。当然，也存在个别例外情况，如镍基合金在高温氧化环境中会改善疲劳性能。通常，如果腐蚀产物（如氧化物）能遮蔽裂纹或使裂纹尖端钝化时，疲劳性能会得到某些改善。

腐蚀疲劳与常规机械疲劳不同，腐蚀疲劳一般不存在疲劳极限，对机械疲劳影响不大的应力循环频率、应力波形，对腐蚀疲劳性能却有着显著影响，频率低（<10 Hz），应力为正弦波，正三角波形的通常加速裂纹扩展。

一般腐蚀介质的浓度越大，环境温度越高及最大应力值越高，腐蚀疲劳裂纹的扩展也越快。腐蚀疲劳的裂纹通常由表面点蚀坑引发，因之常常是多源的，裂纹发展一般是穿晶裂纹，因是多源的，故裂纹通常是成群发展。这是由于一条裂纹扩展到一定深度后，电化学腐蚀通道受阻，而在另外的地方又生成新的裂纹。腐蚀疲劳断口的显微形貌兼有疲劳断口应力腐蚀断口的特征，当腐蚀比较严重时，腐蚀产物的存在往往会掩盖或破坏交变应力造成的疲劳条带等微观特征。

金属腐蚀疲劳机制通常认为与应力腐蚀阳极溶解的钝化膜的破坏－修复机制相似。在腐蚀疲劳的交变应力作用下，滑移过程引起的表面挤入挤出作用是可以累积的，因此钝化膜的破坏比应力腐蚀更易发生。腐蚀疲劳比应力腐蚀的材料－介质组合条件也更加广泛。由于钝化膜破坏后发生点蚀，不断提供新的疲劳源，因而使腐蚀疲劳裂纹成群萌生和扩展。因腐蚀疲劳的 S－N 曲线看不到水平段，即不存在疲劳极限，所以腐蚀疲劳中许用应力随周次增加而下降。

腐蚀疲劳的裂纹扩展速率。有人用线性迭加法来处理，即腐蚀疲劳的裂纹扩展速率为应力腐蚀的裂纹扩展速率与机械疲劳的裂纹扩展速率之和。但进来的研究已经发现，单纯的迭加不尽合理，可能是一种竞争机制，有主有从，也可能有交互作用。

(4) 液态金属脆化

液态金属脆化是指固态金属与液态金属接触时造成固态金属塑性或强度的下降。它包括直接与液态金属接触的脆化及与低熔点金属在低于其熔点时接触的脆化。前者如不锈钢与液态钠的脆化，后者如高强钢与镉的接触脆化。

液态金属脆化问题是核技术，航天航空等高技术领域发展逐渐突出的材料或零部件失效问题。液态金属冷却的反应

堆由液态金属的蒸气压低，可以高温低压运行，从而提高热效率。液态金属冷却可以用于快中子增殖堆，从而使堆芯小型化。在航空技术中使用的铁合金及高强钢零部件，为了提高其抗氢脆能力及电化学稳定性，常使用镀镉层，因而引起镉脆。

不同金属与合金对不同的液态金属的敏感性不同，表5.3-5列出使一些材料致脆的脆化剂。

表 5.3-5　各种金属发生液态脆化的脆化剂

碳钢和低合金钢	黄铜、铝青铜、铜、锌、铅－锡焊料、铟、镉、锂、钾、钠、汞齐合金
不锈钢	同上，但敏感性低的多
铝及铝合金	汞、镓、铟、钖、碱金属（锂除外）
铜及铜合金	汞、锡、铅及其合金，铋，铋－铅合金，锂
镁合金	钠、锌
钛及钛合金	汞、镉

通常认为，液态金属脆化的必要条件是所涉及的金属间既不能生成高熔点化合物，也不具有大的溶解度量；同时液态金属对液态金属的表面张力小，易于浸润。脆化的机理目前还没有详尽的阐述，通常认为是液体金属降低了金属的表面能，并且以液体或蒸气的形态扩散到裂纹尖端区域而使裂纹扩展。液态金属脆化时，固体金属表面原子的离解能比水溶液应力腐蚀时的离解能低得多，前者是固态金属的溶解及质量转移过程，后者是电化学作用的电子转移过程。浸泡在液体金属中的材料，常因液态金属腐蚀造成材料损耗而失效。

纯金属发生液态金属腐蚀，是纯金属的溶解及液态金属原子扩散到固态金属晶格中去，有时还引起相变及形成化合物，合金在液态金属中的腐蚀则常是由于选择性溶解而使某些合金成分从表面下浸出，造成晶间腐蚀或表面下的孔隙，或者引起合金的相变，或者是液态金属与某些合金组元形成化合物。由于溶解度随温度的升高而增大，所以温度是影响液态金属腐蚀的主要因素，当系统中存在温度梯度时，往往造成热区的腐蚀加剧，而在冷区发生材料的沉积。

通常液态金属是流动的，但实际工程中也发生过材料在低于低熔点金属熔点的温度下与低熔点金属接触的低应力脆断。温度范围通常为（1～3/4）T_m（低熔点金属的熔点）。发生这种脆断的裂纹，通常为多叉的树枝状裂纹或与主裂纹相通的网状裂纹，断口多为沿晶形貌，但也有穿晶解理断裂。断口表面被低熔点金属污染，这是判断低熔点金属接触脆断的主要依据。

由于裂纹扩展是因为低熔点金属原子通过表面扩散到裂纹尖端致脆引起的，故裂纹扩展速度由低熔点金属原子扩散系数控制。低熔点金属原子表面扩散系数有如下关系：

$$D_s \approx x^2/2t \qquad (5.3\text{-}5)$$

式中，D_s 为低熔点金属原子的表面扩散系数，为温度的指数函数，cm^2/s；x 为扩散距离，cm；t 为扩散时间，s。

可用式（5.3-5）来估算低熔点金属接触脆断的寿命。钢的强度水平对液态金属脆化敏感性有很大影响，一般在低于某一强度极限后，对液态金属脆化就不敏感。

（5）辐照脆化

在核技术的设备中，一些零部件处于高能电子、中子等粒子流的辐射下，结构材料因高能粒子辐射而发生的脆化即辐照脆化。

通常，固体材料的辐照损伤主要表现在：几何尺寸的变化，体积胀大，密度减小，强度、硬度上升而塑性下降，以及电阻上升、磁导率变化等物理性能的变化。金属材料的辐照损伤主要是由中子辐照造成的。通常，受强辐照的材料，组织结构中存在大量空穴。

材料因辐照损伤导致韧－脆转变温度升高，这对压力容器会造成危险。但在温度较高的情况下，相同中子辐照量时的损伤要少些，这可能和温度提高，促进回复有关。

1.2.3　材料冶金质量引起的脆性断裂失效预测

材料的冶金质量通常是指对材料的低倍缺陷和高倍缺陷的控制，也包括化学成分、工艺过程及材料组织结构的控制等。因材料的冶金质量影响引起的脆断可分成二大类，一类是破坏了材料连续性的缺陷，例如在材料内部存在孔洞、裂纹与基体界面结合很差的夹杂物等，可看作脆性断裂失效的已有裂纹来处理，即用断裂力学概念作为脆断判据去分析。

另一类是材料的成分或生产工艺不当，造成不正常的组织结构，以致材料的断裂韧性 K_{IC} 或缺口韧性 α_K（或 CVN）下降，或脆性转变温度（T_{V28J} 或 FATT）提高了，达不到预期要求的冶金质量。属于这类情况的有害杂质的含量及其分布、合金组元的配比、第二相的组成及其分布、基本组织情况等等。

（1）材料不连续性缺陷引起失效

1）孔洞　孔洞一般指材料内部具有三维形貌的空间，破坏了材料内部的连续性。影响脆性断裂的孔洞一定会在失效件的断口表面上显示。为控制工件的质量，要用无损探伤手段确定其分布位置，三维尺寸，然后根据工件承受载荷时的应力应变分布去判断孔洞存在的危害性，确定工件是否报废或尚能降级使用。

材料中孔洞的形成可能有以下几种情况。

① 缩孔　液态金属凝固时，要发生很大的体积收缩，若凝固过程中不断有液态金属补充，则不会在材料内部形成缩孔。铸件缩孔一般只会在最后凝固的部位形成，例如钢锭顶部或铸件冒口处。这些有集中缩孔的部位都应在钢锭或工件加工成形过程中切除掉。但如果在液态金属凝固过程中控制不当，或铸件工艺设计不合理，液态金属不能随时补充凝固收缩所需的量，则会在工件内部留下缩孔。由于金属凝固时晶体是按树枝状方式生长的，所以缩孔一般都是存在于相邻的树枝间隙间，此时分散的小缩孔排列成树枝状，集中一点的大缩孔则在孔壁留有枝晶形貌，故缩孔的内表面是不平整的。铸件缩孔的部位常位于工件厚截面心部（特别是在变截面处）。

在大型铸锻件中，容易存在缩孔，缩松（密集的小缩孔）等冶金缺陷，可用超声波探伤测定其部位和尺寸。

铸件中具有三维空间形貌的缩孔经锻轧变形后，随着该部位的锻压比增大会改变形貌，在长度方向伸长，厚度方向压缩，可以成为一条二维的内裂缝，这种裂缝断口的相邻表面有顺长度方向沟槽纹路。

焊件中焊缝金属也有凝固过程，但一般不会有缩孔，因凝固时不断有新熔化的焊条金属补充，但在焊缝终端或焊接过程中换焊条时有可能出现缩孔或缩松。

② 气孔　气孔是因为液态金属中气体溶解度比固态金属中大得多，液态金属中溶入气体量多（如在大气中熔化温度过高，则溶入气体多），在凝固时，便有气体形成气泡排出，此时如气体来不及从未凝固的液态金属中排出，便会以气孔的形式留在固态金属内部。炼钢时温度未控制好脱氧不完善常会存在气孔。

气孔是液态金属中的气泡在凝固时形成的，铸件中气孔大都是球形或椭球形，表面光滑。判断是气孔还是缩孔可从孔洞外形和内壁表面光滑与否来确定。作为已有裂纹考虑时，气孔与缩孔的危害是同等的，按临界尺寸判断其影响。计算证明一次加载下不发生脆断的孔洞，也可因为应力强度因子达到疲劳或应力腐蚀门槛值，发生亚临界扩展，扩展到

临界值后骤然断裂。对强韧性要求不高的钢材，用有皮下气泡的沸腾钢或半镇静钢钢锭锻轧，由于碳含量低，在轧成带材或板材时变形量大，气孔可以在高温热轧过程中焊合，可以提高钢材的成材率。但对合金钢，有色金属或合金，因为在锻轧变形中气孔很难焊合而以裂纹的形式出现，所以这种钢锭中就不允许出现气孔，钢锭浇铸时采用发热剂或铸件烧注时采用发热冒口，是排除内部气孔，使缩孔集中到冒口中去的有效措施。钢中气孔主要是由氧引起，钢锭因脱氧不完善，塔形试样也可因氢含量高（$>2\times10^{-6}$）而上的气孔出现白点。铝合金铸件中气孔则主要是氢造成的。焊接时因冷凝速度快，焊缝也易出气孔，通常用X射线探伤来确定。

③ 冷拔空洞　在钢材冷拔时，如冷拔孔型选择不当，一次变形量太大，就可能出现类似单轴拉伸时出现缩颈的现象，即在通过模孔部位出现三向拉应力，使钢材内部出现微孔洞（未达到全部断裂阶段），这种微孔洞在后续变形中伸长。这类缺陷在生产中很少发生，20世纪70年代曾在制作航空螺栓的一批冷拔钢材中，出现过这类缺陷，幸好在螺栓验收过程中发现，未造成严重事故。

2）裂纹　各种加工工序中都有可能在坯料或工件表面或内部形成裂纹。裂纹相当于压扁了的孔洞，它破坏了整体金属的连续性，在传递应力时会在裂纹两端造成应力集中。在断裂失效分析中，可根据裂纹形状，尺寸及在工件中的部位及外载等形成的应力状态进行断裂力学计算，如其尺寸已超过疲劳门槛值或应力腐蚀门槛值，则该裂纹就可发生亚临界扩展，当扩展到临界值 a_{cr}时即骤然失稳扩展，发生断裂。在脆性断裂中，达到 a_{cr}的裂纹就是断裂的主要原因。

在成形工序中各类裂纹的形成原因及其特征分述如下。

① 体积收缩裂纹　液态金属凝固时要有较大的体积收缩，凝固后的金属在冷却过程中也会因冷却先后，后来冷却部分受已冷却部分牵制使进一步收缩受到拘束而产生内应力，这种情况在铸件和焊缝中都会出现。当这一内应力过大时，就会出现裂纹。如果这裂纹是在高温时形成的，则称为热裂纹，由于材料在高温时晶界强度小于晶内强度（即在晶界-晶内等强温度以上），故这种裂纹具有沿晶断裂特征。在铸件中这种裂纹常出现在最后凝固部分，因之沿裂纹还可能伴有晶间显微缩孔，和有害的低熔点杂质偏聚。铸件中热裂纹与表面贯通，裂纹表面有氧化色，这是判断热裂纹的主要依据。

收缩裂纹还有一类是冷裂纹，发生在凝固温度以下，也是由于冷却过程导致收缩不同引起。冷裂纹一般是穿晶的，裂纹平面与导热方向垂直，反映冷却先后次序。冷裂纹即使穿透表面，裂纹面上氧化也不严重，不像穿透表面的热裂纹，一般都严重氧化。

焊缝中热裂纹发生在焊道金属中间，因这部分是熔化后最后凝固冷得最慢部分。但冷裂纹也可发生在热影响区，是在焊后相隔一定时间才出现的延迟开裂裂纹，主要是焊接过程中进入了氢的关系，是一种氢致开裂裂纹，可以是沿晶，也可以是穿晶解理（或准解理），裂纹表面几乎没有氧化色。

工件在加热或冷却过程中，由于存在温度梯度，收缩或膨胀不均匀会造成内应力，如同时又有相变发生，相变体积变化也会产生内应力，两部分迭加起来的总应力，如超过了材料的断裂强度，就会出现裂纹。热处理淬火时，对于变截面工件，由于各部位冷速相差太大，虽然膨胀系数造成的热应力和抵消部分马氏体相变产生的组织应力，但也容易在变截面处造成淬火裂纹。淬火裂纹有的是在淬火过程中产生的，而有的是在淬火后未及回火时产生的，对于这种延迟断裂，有人认为除相变造成的组织应力外，还有氢元素扩散偏聚造成氢浓度富集而导致的。

② 锻轧件过烧裂纹　过烧是锻造时工件加热温度过高，工件表面上的易熔部位发生熔化而流失。例如有的钢中存在的少量共晶组织处，或者含有低熔点杂质的晶界处，这些都是结晶时低熔点部位。这种裂纹都是一般因加热是在氧化性气氛进行的，所以裂纹中填满氧化物，在金相试样中可看到黑色网络，破坏了金属的内在联系。

高速钢类钢坯锻打时，会出现因锻造变形升温造成过烧熔化开裂的问题，这类钢铸态是有熔点低的共晶组织，锻造加热温度已接近共晶温度，在自由锻方坯时，坯料上下表面与锤、砧有摩擦力出现锥形不变形区，四周自由表面往中心是易变形区，变形热会导致锻坯升温，如锻打速度快，变形大，就会使中心升温，出现十字形亮线，若温度过高，便顺方坯对角线开裂。在高温合金锻造时，如合金中低熔点有害杂质量高，使晶界熔点降低，锻打时因加热温度接近熔点，也会出现类似裂纹。方坯中间形成的十字形锻造过热裂纹在镦粗后也可保留十字形。

③ 锻轧中过度变形开裂及折叠裂纹　不论在热状态或冷状态使材料发生变形时，如形变量超过材料在该温度时的断裂应变值，就可出现裂纹。锻、轧件表面如有不平整，则在下一道次中会压平而形成折叠裂纹，轧钢件上的折叠裂纹是由表面弯向中心，裂纹面纵向伸展，通常贯穿全长。锻件上折叠不规则。所有折叠裂纹判断时的主要特征是裂纹由表面向里，裂纹面有氧化皮，钢中裂纹两侧有脱碳层，脱碳层一直包裹住裂纹尖端。这是因为钢坯加热时表面都会有脱碳层，锻轧时表面不平整处微凸起部位一定有脱碳层，故折叠成的摺缝一定在脱碳层中间。

④ 纵向发裂　发裂有的是铸坯中原来的不连续空洞或裂纹，在轧制成圆棒过程中顺纵向延伸（孔洞闭合但未焊住）形成的，也有是延性夹杂物顺轧向变形，或脆性夹杂轧碎后顺轧向排列，在酸浸底蚀时形成的。还有一种可能是钢中白点的纵截面。

钢中白点是由于钢液中溶氢过高。在凝固后固态金属中熔氢过饱和，在冷却发生相变时由于铁素体溶氢能力比奥氏体差的多，过量溶解的氢要以分子态的氢气泡析出，这种气泡有巨大的内压，在相变应力协同作用下能使钢材造成内裂纹。在横剖面上，经酸浸后可看到辐射状的裂纹，在塔形试样上就表现为发裂。在淬透性高，风冷能硬化的结构钢中，氢含量超过2×10^{-6}，就可能出现白点。

⑤ 回火裂纹　高强钢的焊接性能较差，有些构件在焊完后的去应力退火过程中，还会出现开裂现象，这种开裂通常与焊件拘束应力过大有关。

高硬度的冷轧辊由于低温回火不能去除残留的内应力，在回火后的储存过程中，或者在后继的表面磨削过程中，由于内应力分布状态改变（也有人认为是溶解氢富集）而发生脆性断裂。有的即使当时未察觉出现开裂，在工作时因轧制压力便发生断裂。

⑥ 磨削裂纹　经淬火硬化的工件，在用钝砂轮磨削，而且进刀量较大时，就会在被磨削工件表面形成网状微裂纹，破坏了表面层金属的连续性，成为脆断或疲劳的裂纹源。这种缺陷形成是与砂轮钝化又未及时修正或所选砂轮过硬，进刀量太大有关。如是表面渗碳工件，还应注意是否工件表层有网状碳化物，如有网状碳化物，这也是形成磨削裂纹的主要原因。

（2）冶金因素对材料韧性的影响

材料的韧性从最根本上说是与键结合状态，晶型和有序化程度有关。具共价键的材料有金刚石、硅、碳化物、氮化物和硅化物等，键合方式是相邻原子间电子共用，这类材料表现最脆。离子键结合对价电子位置的约束相对较少，只是电正性元素的原子（金属）把价电子交给了电负性元素（如周期表右边的非金属元素）的原子，基本上靠静电力量键

合。如氯化钠晶体是每个钠原子将一个价电子转移给一个氯原子形成的离子晶体。这类晶体也还是很脆的，但是单价的离子键合晶体的变形能力要比多价的好些。金属键合的晶体是金属原子把价电子都交出来，这些价电子为整个晶体，而不是几个金属离子共有，也就是说这类晶体是正离子的有序排列形成的晶格被自由电子气连结在一起。这样的自由电子云的运动所受约束最少，故这类材料的变形能力最大。

低对称性的晶体点阵排列滑移困难，因之紧密排列、高对称性的晶体（如面心立方或体心立方晶型）更易产生滑移，韧性要好些。同时固溶体中有序化程度增加，也会使滑移变得困难，长程有序的影响比短程有序更为严重，也就是脆化倾向增加。

以下分述各类冶金因素对断裂韧性及脆性转变温度的影响。

1）晶体结构与变形方式　脆性转变温度标志了断裂方式由剪切向解理转变的温度效应，这种现象并非在所有铁基合金中都能看到，而仅仅在部分体心立方晶型的铁基合金中才有，对其他面心立方晶型的工程合金，脆性转变的温度效应是不明显的。有些合金含量偏高的钢，如马氏体时效钢，虽然仍是体心立方晶型，但含碳量极低，放在低温下仍是韧性断裂（剪切型）。例如 18Ni－5Mo－9Co 马氏体时效钢，低到－196℃还是剪切型断口，除非是在这么低的温度时冲击加载，才会在这种和类似的其他马氏体时效钢（和一些严密六角晶型的合金）中出现解理和准解理断口。

对面心立方的铁基合金，滑移系比较多，容易发生交叉滑移，不易出现解理断裂，但铁基体心立方晶型的合金在温度降低时交叉滑移的能力会大幅度下降，变形方式要变成以孪晶为主，这样就易发生解理断裂，而且使脆性转变温度（T_T）以下的韧性降低，但在体心立方铁基合金中如堆垛层错能低，则容易发展交滑移，铁基合金中加入置换型组元会使固溶体的堆垛层错能降低，也许上述的 18Ni－5Mo－9Co 马氏体时效钢在－196℃还是剪切型韧断，除了碳低以外，还与此有关。

面心立方晶型的合金材料，如奥氏体钢，铜，铝等即使在低温也易发生交滑移，因之这些材料都有良好的塑性，但是面心立方晶型的金属，堆垛层错会抑制交滑移发生（这作用恰恰和在体心立方晶型金属中相反），因之如合金中存在大量堆垛层错，就会降低材料的宏观塑性。如 19Cr－10Ni 型奥氏体不锈钢，在温度降低时韧性也要缓慢下降，低温下的冲击断口看不到由剪切到解理的转变，也就是说没有明显脆性转变温度。但 C－M－N 型奥氏体不锈钢，含氮量到 0.15%，降低了堆垛层错能，就可看到在某个温度区间有韧性突降现象。

2）炼钢脱氧制度的影响　碳钢由于脱氧程度不同可分为沸腾钢，半镇静钢，镇静钢三种，脱氧程度按此顺序增加，钢液凝固后的密实度和均匀度则随氧含量增加而下降，冲击韧性也是这样变化。脆性转变温度则因氧含量增高而上升。

在相同脱氧制度下，终轧温度对脆性转变温度也有影响，轧制温度增加，终轧温度高，则晶粒粗，故脆性转变温度提高。

3）晶粒度的影响　通常在材料的强度级别提高时，韧性都会降低，只有在晶粒尺寸减小时，才可使强度升高，脆性转变温度降低。

在北方严寒地区冬季发生的许多脆断事故，基本上都与晶粒粗大有关。现许多部件采用亚临界处理（A_{c3}与A_{c1}温度之间加热），可显著降低脆性转变温度，这与改善晶粒度有关。在不改变材料化学成分的条件下，不论什么显微组织形态，细化晶粒都是可以有效改善韧性的。但是要注意一点，前面讲到的晶粒度对改善韧性的影响都指的是脆性转变温度，也就是说晶粒大小只对涉及解理断裂方式才有重大影响，对韧窝断裂方式，几乎没有影响，所以要改善高台阶能，细化晶粒是不会有明显效果的。

4）化学成分的影响　合金的化学成分是通过不同工艺，形成合金的各种组织结构来反映各种性能的，包括工艺性能和使用性能。钢中合金组元的主要影响如表 5.3-6 所示。

表 5.3-6　钢中合金元素的主要作用

元素	合金元素的作用
碳	钢中主要强化组元。间隙式固溶强化或形成碳化物造成弥散强化。但大颗粒碳化物也可成为裂纹源，若成网状分布时导致脆断。碳量增加时提高脆性转变温度，增 0.01%碳，约使 T_T 增加 4℃
镍	主要韧化组元。降低相变温度，增加淬透性，稳定奥氏体
铬	碳化物形成组元。在结构钢中增加淬透性，保证不锈钢抗腐蚀性能，固溶强化
钼	强碳化物形成组元。增加结构钢淬透性，抑制回火脆性，有固溶强化作用
硅	脱氧剂。固溶强化组元，同时提高脆性转变温度，在结构钢淬火回火时提高碳化物的转化温度
锰	脱氧剂。固定硫形成 MnS，可防止 FeS 形成热脆，增加结构钢淬透性，Mn－C 比增大可降低脆性转变温度
钴	在马氏体时效钢中钴有细化时效沉淀相的强化作用
钛	结构钢中加钛形成 TiN 或 TiCN，抑制奥氏体晶粒长大。在马氏体时效钢中作为沉淀相组元 Ni_3（TiAl）起强化作用，在硼钢中起定氮保硼作用
铝	强脱氧剂。也可形成 AlN，钉扎奥氏体晶界抑制晶粒长大（直到 980℃），由于能减少固溶体中 N 量，故可降低脆性转变温度，在马氏体时效钢中作沉淀相组元组成 Ni_3（TiAl）起强化作用

合金化学成分中这些有益组元含量在合金设计及合金牌号标准中都已作了规定，对机械构件的钢材成分进行分析时，首先应该要看是否含量超标，不超标时也要考虑合金配比是否合适，成分在牌号规范内，但配比不合适，工艺性能或使用性能上达不到要求的事例是很多的。在许多合金结构钢中，成分配比上的被动常造成钢种淬透性不稳定，这就使同一零件热处理后残留应力及性能发生变化引起失效事故。这些问题在生产厂大都有内控标准，出现问题的原因或是失控，或是内控标准不合理，或是合金成分上的超标。

合金的化学组成中除希望的合金组元（即从使用性能或工艺性能出发所配入的成分）外，也还会有原材料或生产过程中不可避免或很难避免的，不希望的，对性能有害的组元掺入。这些有害杂质一类是生产过程中带进来的气体，如氧、氮、氢。还有一类主要是原材料或辅助料中带进来的有害杂质，这些主要是周期表右侧的一些非金属元素。

氧含量对钢的韧性或脆性转变温度造成一定的影响。脱氧制度越不完善，脆性转变温度便越高。氮在铁素体中的作

用类似于碳，也起提高脆性转变温度的作用。氮是导致低碳钢发生蓝脆的主要原因，这通常是因为脱氧不充分，含氮量较高，热轧钢材加热到 200 ~ 400℃范围（表面氧化色为蓝色）时，氮形成柯垂耳气团钉扎，位错在运动时摆脱不了气团的钉扎，因之屈服点升高，韧性下降。

在钢中利用脱氧剂固定氧（用铝时也包括氮），形成的氧化物如来不及由钢液中浮出，便会成为氧化物或硅酸盐，铝酸盐类夹杂，这类夹杂物有的在高温时可发生塑变（如硅酸盐类），在轧材中会顺轧向延伸，也有的是脆的，轧制过程中会碎裂且呈虚线状顺轧向排列。这些夹杂物（包括 MnS 等）与基体界面结合很差，类似于空洞。

金属中存在氢时会导致氢脆，过去认为这种现象主要存在于体心立方晶型的钢中，后来发现钛合金也有，近年来的工作证明在面心立方金属，如稳定的奥氏体不锈钢及镍基合金，铝合金中也有，只是在体心立方的金属中更敏感。

在钢中氢的溶解度应该是很低的，氢量超过溶解度就可能出现白点。但由于钢中有许多缺陷，内界面都可吸氢，故公认低于 2×10^{-6} 不出现白点。如果有 2×10^{-6} 的氢是均匀分布在钢材内部，一般认为不会导致脆性。但是在工件受载时，如内部存在应力梯度，则固溶氢会向高应力区及高塑变区富集，使断裂塑变下降造成脆断。

5）组织的影响　材料内部组织结构及其形态分布，对韧性或塑性指标有明显影响，以下由宏观和微观两方面来讨论。

① 宏观组织的影响　热轧碳钢板在不同方向取样制成的构件的冲击性能是不同的。这是由于原来钢坯中的合金元素（包括溶解的杂质组元）的偏析和夹杂物在轧制变形后形成的流线产生的影响。在使用锻轧的方法制造构件时，应特别注意流线的分布与构件受力的关系。

② 微观组织的影响　由实验可得，马氏体状态的碳钢冷脆转变温度最低，贝氏体组织的次之，珠光体（含铁素体）的脆性转变温度相对要高一些。这些是在奥氏体晶粒度相同的情况下对比的。晶粒度的影响比原始组织是珠光体还是马氏体的影响要大。

1.2.4　材料的冷脆性及脆性转变预测

由于脆性断裂事故常常发生在气温较低的环境中，所以人们也研究了温度对材料性能的影响。在某一较低的温度下，正常屈服条件下的试件出现脆断；而在存在约束应力的情况下，在某一较高的温度下就可以出现脆断。在两个温度之间，对光滑无缺口试样，无脆断问题，对缺口试样，则存在缺口脆性，引起脆断事故。在静拉伸试验中，可以看到随应变速率（加载速率）增加，材料的屈服强度又有相应增高的现象，这和温度降低的影响类似。普遍通用的表达式是

$$\sigma = f(\varepsilon, \dot{\varepsilon}, T) \tag{5.3-6}$$

式中，σ 为应力；ε 为应变；$\dot{\varepsilon}$ 为应变速率；T 为试验温度。

（1）T_{TV27J} 判据

1978 年联邦德国材料试验协会提出由夏比冲击试验提出如下判据：

$$T_{工作} \geqslant T_{构件} = T_{TV27J} + \Delta T \tag{5.3-7}$$

式中，$T_{工作}$ 为最低工作温度，℃；$T_{构件}$ 为构件的韧 - 脆转变温度，℃；T_{TV27J} 为夏比 V 形试样系列温度冲击试验中由 27 J 能力确定的脆性转变温度，℃；ΔT 为冲击试样的脆性转变温度与构件的脆性转变温度的差别。

（2）T_{TE50} 判据

日本焊接协会低温用钢材料标准 WES - 135 提出了构件防止低温脆断的标准，采用的是 CVN 系列试验中高阶能一半时标志的脆性转变温度 T_{TE50}，用经验统计总结了工作应力与 T_{TE50} 之间的关系。提出：

G 类：在一般结构中采用，保证 $a = 10$ mm 裂纹不发生失稳扩展条件：

$$0.317\sigma = \begin{cases} 360f(t)\exp\left[3\,100\left(\dfrac{1}{T_\sigma+273}-\dfrac{1}{T+273}\right)\right] \\ (用于 -18℃ \leqslant T_{TE50}) \\ (260+275T_C)f(t)\exp\left[3\,100+ \right. \\ \left. 22.5T_\sigma\left(\dfrac{1}{T_\sigma+273}-\dfrac{1}{T+273}\right)\right] \\ (用于 -58℃ \leqslant T_{TE50} \leqslant -18℃) \end{cases} \tag{5.3-8}$$

A 类：对防止脆断有特殊要求的结构，保证 $a \leqslant 100$ mm 时不发生失稳扩展事故。

$$1.02\sigma = 150f(t)\exp(2\,680+120T_C)\left(\frac{1}{T_\sigma+273}-\frac{1}{T+273}\right) \tag{5.3-9}$$

式中

$$f(t) = \begin{cases} 1-0.05(t-30) & 板厚\ t \leqslant 35\ \text{mm} \\ 0.75 & 板厚\ t > 35\ \text{mm} \end{cases}$$

$$T_C = \begin{cases} 18 + T_{TE50} & -58℃ \leqslant T_{TE50} \\ 25.5 + 1.13T_{TE50} & -196℃ \leqslant T_{TE50} < -58℃ \\ T_{TE50} & T_{TE50} < -196℃ \end{cases}$$

σ 为工作应力，MPa；t 为板厚，mm；T 为最低工作温度；T_{TE50} 为系列温度冲击试验高阶能 50% 所确定的脆性转变温度，℃。

（3）NDT 判据

美国海军实验室曾对低强度铁素体钢 NDT 的应用，提出以下建议：

1）$T_{使用} \geqslant$ NDT，只允许所加应力低于 35 ~ 56 MPa；

2）$T_{使用} \geqslant$ NDT + 17℃，允许所加应力低于 $\sigma_s/2$；

3）$T_{使用} \geqslant$ NDT + 33℃，允许所加应力低于 σ_s；

4）$T_{使用} \geqslant$ NDT + 67℃，所加应力达到强度极限才会发生韧性断裂。

NDT 与冲击试验测定的韧 - 脆转变温度 FTE、FTP 对应关系大致如下：

$FTE \approx NDT + 33℃$，$FTP \approx NDT + 67℃$

1.3　疲劳断裂失效预测

人类对机械构件疲劳失效的认识和研究，经历了漫长的道路。早在 19 世纪中叶，随着铁路工程的迅速发展，火车轮轴在运行中，在其轴颈圆角根部处发生突然断裂，而在断裂之前轮轴没有显示出任何的宏观可见的塑性变形。从断裂轴上取样作静强度检验，结果表明轮轴材料完全符合设计规定的要求。从此，人们开始对这种新的失效形式进行分析研究，并把这类承受交变载荷构件所发生的断裂称为疲劳断裂。

为了适应铁路工程的迅猛发展，从设计上避免由轮轴疲劳失效而导致灾难性事故的发生，研究者对钢材的疲劳性能作了大量系统的研究，从中获得了评定材料疲劳性能的曲线，并且至今这种曲线仍是疲劳设计中应用的最基本的数据。主要结论有两条：第一，钢材的疲劳断裂寿命取决于交变应力范围而不是交变应力的最大值；第二，存在一个最低应力范围极限，在低于此极限时，钢材则不再发生疲劳断裂。以后，开始对各种金属材料的疲劳 $S-N$ 曲线进行大量的试验测定，并且又进一步考虑温度、加载频率、平均应力、应力集中、环境等各种因素对 $S-N$ 曲线的影响，把由试验获得的疲劳性能数据汇编成册，供设计者使用。

由此可见，机械构件在服役中出现的失效以及对这种失效所进行的分析研究，推动了由静强度到疲劳强度这一设计思想上的进步。

随着运载工具和动力机械的迅速发展，许多构件承受的交变载荷由简单的对称循环发展到非对称循环。为满足这类构件的疲劳设计需要，Goodman（1862～1935年）试验建立了考虑平均应力的疲劳等寿命图，它是疲劳设计中应用的另一种基本数据。与此同时，材料工作者对材料的疲劳断裂机制也展开了广泛而深入地研究，20世纪初，用光学显微镜来观察材料在疲劳过程机械产生疲劳失效分析，研究预防疲劳失效的各种措施和手段，对于正确地进行机械构件的设计、提高机械产品质量，具有十分重要的意义。

随着现代工业的发展，疲劳设计思想也在不断地发展和演变：由传统的无限寿命设计发展到安全寿命设计，又进入到现代的破损安全设计（或损伤容限设计）。20世纪60年代发展并建立起来的断裂力学，很快地也被用于研究材料的疲劳裂纹扩展特性。Paris的研究指出，裂纹尖端的应力场强度是控制疲劳裂纹扩展速率（da/dN）的主要参量，由此建立了裂纹扩展速率与应力强度因子范围（ΔK）之间的一般关系，即 $da/dN-\Delta K$ 曲线。试验测定获得的各种材料的 $da/dN-\Delta K$ 曲线，是损伤容限设计的基本数据。

由上述可见，人类从发现疲劳破坏这一现象以后，逐渐地认识和部分地掌握了疲劳断裂的一些规律，并且遵循这些规律在机械设计、选材以及制造工艺等各个环节中去寻求改善机械构件疲劳断裂抗力的各种途径，在整个历程中的每一个进步，都与失效分析工作密切相关。

1.3.1　疲劳断裂现象及特征

（1）静载断裂与疲劳断裂

根据以上所述，在静载条件下，材料在发生断裂之前，通常需要经历弹性变形、塑性变形和塑性失稳等几个阶段。断裂后的塑性变形总量取决于材料的塑性以及环境条件。上述就是塑性材料在静载条件下发生断裂过程的主要特征。但是，如果材料承受的不是静载荷，而是交变（循环）载荷，加在材料上的交变应力的最大值即使是低于材料的弹性极限（σ_e），经过若干次应力循环之后，材料也会发生断裂，在交变载荷作用下发生的断裂，称为疲劳断裂。通常，用 $S-N$ 曲线（应力－断裂循环数）来表征材料的疲劳断裂行为。材料承受的交变应力越高，则发生断裂的循环数就越少；反之，其断裂循环数就越多。

以往有不少的试验研究，企图在材料的静强度与疲劳强度之间建立起某种经验关系，这样便可通过材料的静强度去推断或定量地给出疲劳强度。这些工作所获得的结果为机械疲劳设计提供了在工程上可使用的若干图表，但未曾获得任何有普遍意义的精确关系。这方面的许多研究结果指出，静载断裂和疲劳断裂是两种互相区别的断裂过程，它们分别受控于各自的断裂机制，所以材料的静强度和疲劳强度都是在一定条件下材料所表现出的力学特性。因此，了解和掌握疲劳断裂的各种特性，是进行疲劳失效分析工作的最基本的条件。

（2）疲劳断裂现象及一般特征

疲劳断裂的各种特征是与静载断裂和其他方式断裂的特征相比较而存在的，可以归纳成如下几种。

1）脆性断裂方式　构件或试样在整个疲劳过程中，不发生肉眼可见的宏观塑性变形。对于应力疲劳（即高周循环疲劳），甚至发生疲劳断裂之后，整个构件或试样也不呈现出宏观塑性变形。所以，可以将一般的疲劳断裂归入“脆性断裂”。在多数情况下构件的疲劳断裂是突然发生的，这种断裂方式给构件失效前的预报和预防工作带来极大的困难。

2）断口形貌　疲劳断裂具有区别于其他任何性质断裂的断口形貌，呈现出独特的断口花样。一个典型的疲劳断口往往由裂纹源区、裂纹扩展区和瞬时断裂区三个部分组成。疲劳断口上的这种独特形貌是区别于其他任何性质的断裂的极为重要的凭据。

3）表征疲劳性能的 $S-N$ 曲线　疲劳断裂寿命（即断裂循环数）随交变应力值的降低而增高。对于低碳钢，$S-N$ 曲线在低应力区逐渐趋于水平，而低于该水平线的交变应力，不会再使低碳钢引起疲劳断裂，通常将此水平线上的交变应力称为疲劳强度极限，或简称为疲劳极限。除了低碳钢之外，合金钢、铝合金、铁合金、镍基合金等金属材料，它们的 $S-N$ 曲线不存在水平线段。对于这类材料，通常以 $N=10^7$ 或 $N=10^8$ 次循环不发生断裂所对应的交变应力值定义为条件疲劳强度极限，或简称为条件疲劳极限。

4）疲劳强度的分散性　在同一个交变应力下，试样（或构件）的疲劳断裂寿命呈现出一定的分散性。材料的表面完整性（其中包括表面粗糙度，残余应力，表层的组织结构等）、材料的组织结构均匀性、试样尺寸波动、缺口效应以及环境等，都是引起断裂寿命出现分散性的因素。所以，材料的 $S-N$ 曲线是一个具有一定分散度的曲线（如图5.3-2所示）。每一应力水平下试样的断裂寿命分布服从统计规律。所以，可以将 $S-N$ 曲线分散带按不同存活率划分为若干条相互平行的 $S-N$ 曲线，其中越靠近纵坐标轴的 $S-N$ 曲线，它的存活率越高。设计者根据构件承载的重要性，选择一定存活率的 $S-N$ 曲线，由此进行疲劳强度设计和寿命估算。

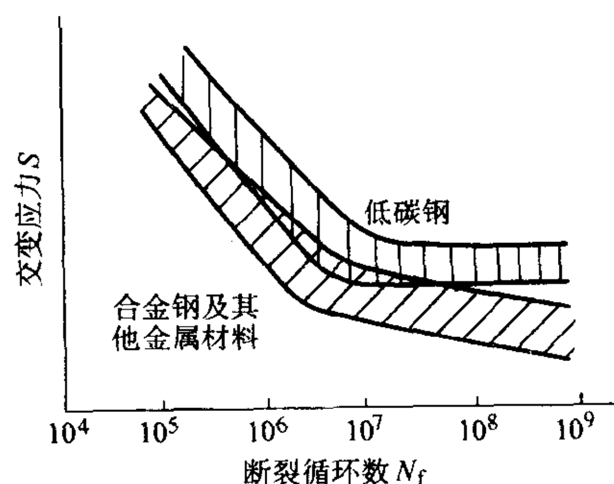


图5.3-2　各种金属材料的疲劳 $S-N$ 曲线示意图

塑性材料在静载条件下，缺口试样的抗拉强度通常高于光滑试样。但是在交变载荷条件下，缺口试样的疲劳强度通常低于光滑试样的疲劳强度。材料的塑性越高，缺口试样的疲劳强度越接近于光滑试样，而材料的塑性越低，则两种试样疲劳强度的差距就越大，这是因为材料的塑性越低，其疲劳缺口敏感性越高的缘故。

5）表面粗糙度的影响　表面粗糙度对材料的静强度几乎没有多大的影响，但对疲劳强度却有非常明显的影响。在一般情况下，任何材料的疲劳强度都随其表面粗糙度的降低而增高。

6）环境因素的影响　腐蚀性环境对材料的静强度虽有一定的影响，但其影响程度远不如它对疲劳强度的影响。一般情况下，对材料敏感的腐蚀环境能够显著地降低材料的疲劳强度。对于一般中等强度的合金结构钢，腐蚀环境可使其疲劳强度极限下降，甚至能使钢的疲劳强度极限消失。

7）表面滑移带形式　与单调加载（静载）的情况相似，在交变载荷下材料表面也出现滑移带，但其分布形式及密集程度各不相同。交变载荷下形成的滑移带比单调加载下形成的更为密集，滑移带之间的台阶高度，前者远比后者为高。此外，均质材料的光滑表面，在交变载荷作用下会出现“挤出”或“挤入”现象，而在单调加载条件下则不会出现这种现象。

8）疲劳裂纹扩展门槛值　材料的疲劳 $S-N$ 曲线上存在着一个疲劳强度极限值，而疲劳裂纹的扩展也存在着一个

极限值，通常称为疲劳裂纹扩展门槛值（ΔK_{th}）。断裂力学的研究指出，对于一个长疲劳裂纹（长度约大于0.5 mm），当裂纹尖端的应力场强度因子范围（ΔK_{I}）超过材料本身的ΔK_{th}时，疲劳裂纹便开始扩展，而当$\Delta K_{I} < \Delta K_{th}$时，裂纹不再扩展。

1.3.2 疲劳断裂规律

(1) 疲劳断裂的宏观规律

新材料和新工艺的研究，需要测定材料的疲劳性能及研究各种因素对疲劳性能的影响，构件的疲劳设计需要以材料的各种疲劳性能为依据，有时还需根据全尺寸构件的疲劳试验来确定构件的疲劳寿命，疲劳失效分析，不仅需要了解材料的疲劳性能，有时还需要做构件的故障再现性模拟疲劳试验。因此，了解材料的疲劳断裂的宏观规律及影响这些宏观规律的各种因素，不仅对新材料和新工艺的研究及构件的疲劳设计是必要的，而且对进行构件的疲劳失效分析也是十分重要的。

1) 高周疲劳

① 疲劳载荷形式　构件在服役中所承受的交变载荷形式繁多，基本上可分为以下几种形式：拉伸（或拉压）脉动载荷，弯曲载荷（旋转弯曲，反复弯曲），扭转载荷，双轴载荷（拉-拉，拉-压，拉-弯，弯-扭，拉-扭等）。

构件在服役中所承受的加载方式基本上可分为以下三种：恒幅加载、程序加载和随机加载。

② 疲劳 $S-N$ 曲线　材料的 $S-N$ 曲线是疲劳设计、材料和工艺研究以及疲劳失效分析的基本依据，$S-N$ 曲线是材料处于弹性应变（或与弹性应变相比出现的塑性应变可被忽略）的范围内的应力与循环次数间的关系曲线。在此条件下，材料的疲劳断裂寿命通常处在10^3～10^8次循环之间。

③ 疲劳等寿命曲线　有些构件在服役中承受非对称循环载荷，为了满足这类构件的疲劳设计和疲劳失效分析的要求，需要提供在不同的平均应力条件下的 $S-N$ 曲线。为此，首先需要得出在各种平均应力条件下的 $S-N$ 曲线，然后再由这些 $S-N$ 曲线计算获得各种寿命条件下的$\sigma_a-\sigma_m$曲线。正平均应力降低材料的疲劳强度（或疲劳寿命），而负平均应力增高材料的疲劳强度（或疲劳寿命），疲劳断裂的寿命越高，平均应力的影响就越大。

④ 疲劳应力集中系数及缺口敏感性因子　一个含有缺口应力集中的构件或试样，当对它施加的静载使得光滑表面处的截面处于弹性应力状态时，则缺口处的平行于外力的最大应力σ_{ymax}与净截面应力σ_n之间有以下关系：

$$\sigma_{ymax}=K_t\sigma_n \tag{5.3-10}$$

式中，K_t为弹性应力集中系数。

$$K_t=\frac{实际的最大应力}{名义应力}$$

当构件承受交变载荷时，在缺口或内外圆角等出也产生疲劳应力集中，这里相应地有一个疲劳应力集中系数K_f，被定义为：

$$K_f=\frac{有效疲劳应力}{名义疲劳应力}$$

然而，与静载下的弹性应力集中系数不同，疲劳应力集中系数不仅是几何形状的函数，还是材料和加载方式的函数，为了考虑材料特性的影响，引进一个“缺口敏感性指数”q，用它来评价弹性应力范围内缺口对材料疲劳强度的影响。在高周疲劳条件下，q被定义为：

$$q=\frac{K_f-1}{K_t-1} \tag{5.3-11}$$

⑤ 累积损伤原理　一些构件在服役中不是简单地承受单一恒幅交变载荷，而是承受一系列由各种幅值组成的交变载荷，即一种谱载荷。为了对这种构件的疲劳寿命进行估算，首先需要解决在谱载荷作用下构件的疲劳损伤形式。Miner对此提出了一个累积损伤假说，并给出了简单的数学表达式，称为线性累积损伤规则。

⑥ 影响疲劳 $S-N$ 曲线的因素　$S-N$ 曲线是衡量材料疲劳性能的最基本的数据，但是它的形状和位置受着许多因素的影响。在需要借助于疲劳性能（强度和寿命）进行构件的疲劳失效分析时，就必须了解各种因素对材料的 $S-N$ 曲线的影响及其影响程度，以便做出正确的判断和处置。

影响材料 $S-N$ 曲线的因素繁多，可大体上分为三大类：物理因素，几何因素和外界因素。下面分别讨论各种影响因素。

a) 材料种类　不同类型的金属材料具有不同的组织结构和静强度，所以它们表现出不同的疲劳强度。在一般情况下，对于完全不同类型的材料，其静强度越高，它所表现的疲劳强度亦越高。一般的变化规律是，疲劳强度随其静强度的增高而增高。材料的疲劳强度随其显微组织内的非金属夹杂物及弥散强化相的数量和分布状态而变化。

b) 材料的晶粒及亚晶粒尺寸　一个承受应力$\Delta\sigma$范围的光滑试样，在表面层切应力范围$\delta_\tau=\tau_{max}-\tau_{min}$的作用下，疲劳裂纹由表面萌生。对于光滑表面，疲劳裂纹通常产生于驻留滑移带上，而驻留滑移带的长度通常即为晶粒尺寸。

此外，晶粒取向也会影响疲劳性能。晶粒取向与试样长轴方向一致的疲劳强度最高，而与试样长轴方向相垂直的疲劳强度最低。

c) 材料的静强度　许多试验结果表明，材料的疲劳强度有随着其抗拉强度提高而增高的变化趋势。当材料的σ_b低于某一定值时，σ_w与σ_b之间存在线性关系，但当σ_b高于某一定值时，则非线性关系。

材料的疲劳裂纹萌生及其扩展取决于晶体滑移的难易程度，而材料的屈服强度是衡量晶体滑移难易的力学性能指标之一，所以有理由认为材料的屈服强度与σ_w之间存在着比与σ_b更为密切的关系。

d) 材料的表面完整性　这里所谓的表面完整性，主要包括表面粗糙度、表面残余力、表面层显微组织以及表面类裂纹（机械划伤，电弧烧伤、锻、铸、焊裂纹，热处理裂纹）等。因为绝大多数情况下，疲劳裂纹产生于构件的表面，所以表面完整性可在宽广的范围内影响其疲劳强度。

Ⅰ) 表面粗糙度　一般地说，材料的疲劳强度随其表面粗糙度的降低而增高。在相同粗糙度的条件下，材料的强度越高，表面粗糙度对疲劳强度影响越大，表面粗糙度越低，疲劳强度越高。

Ⅱ) 表面残余应力　表面完整性的诸因素中，表面残余应力是影响疲劳性能显著的因素之一。在所有情况下，残余拉应力降低而残余压应力提高材料的疲劳强度。因此，为了改善构件的疲劳性能，应该采用各种工艺技术使构件表面层引入残余压应力，或采用热处理、机械振动等其他工艺消除由于机械加工或装配带来的有害的残余拉应力。相反，由于镀镍、镀铬等，对构件表面形成的残余拉应力，会使构件的疲劳性能发生大幅度的下降。

Ⅲ) 表层显微组织　材料表层的晶粒、亚晶粒尺寸、位错密度、相结构以及元素贫化等，都会强烈地影响其疲劳强度。因为疲劳源优先产生于表面，所以只要改变材料表层的组织结构就能显著地影响其疲劳性能。表面脱碳，不仅会降低钢的疲劳强度，而且会增大疲劳强度的分散性。

Ⅳ) 表面类裂纹　由于各种原因使机械构件表面存在着的类裂纹会使疲劳强度下降，如果对划伤表面采用喷丸强化引入压应力，则其疲劳强度可获得明显的提高。非金属夹杂物或其他脆性相如果处于材料表面，它们与基体相交的界面往往起到类裂纹的作用。

e) 缺口效应 对于绝大多数材料，增大缺口根部局部地区的应力，均造成疲劳强度的下降。疲劳强度随着应力集中系数的增高而下降。

f) 尺寸效应 试样或构件的尺寸大小影响疲劳强度的高低。通常，弯曲扭转疲劳受尺寸效应的影响比单轴拉伸者要大些，而其中光滑试样又比缺口试样的影响为大。

尺寸对疲劳性能的影响主要有两方面的原因：尺寸越大，外表面的面积越大，表面上形成疲劳源的几率也就越高，所以疲劳性能下降；另一方面，在相同的循环应力水平下，尺寸越大，表层所承受的平均应力越高，并且试样表层处在高应力水平的体积越大，出现疲劳源的几率也越大，所以疲劳性能降低。

对于缺口试样，由于缺口根部的应力集中而使表层的应力分布趋于缓和，不同尺寸的表层平均应力值不像光滑试样那样差别显著，所以尺寸效应对缺口试样的疲劳性能的影响较小。

g) 腐蚀效应 工业中使用的大多数金属材料，其疲劳性能会不同程度地受到腐蚀环境的影响。一般情况下，加载频率越低，腐蚀环境对疲劳性能的影响越大。

腐蚀介质对光滑和缺口表面的疲劳性能具有不同的影响。盐水降低钢的两种试样的疲劳强度，但并不影响钛合金的疲劳性能。对于铝合金，盐水只降低光滑试样的疲劳强度，但基本上不影响缺口试样的疲劳强度。

在一般情况下，腐蚀环境对 $S-N$ 曲线上的长寿命区的影响比短寿命区的影响要大。通常，那些在低应力长寿命条件下疲劳破坏的构件，腐蚀环境往往表现出更为强烈的影响。

h) 温度效应 绝大多数的金属材料的疲劳强度在室温以下随着温度的降低而逐渐增高。但是在室温以上时，材料的疲劳强度随着温度的增高其总的变化趋势是逐渐下降的，不过这种变化不是单调的，各种材料的变化规律不尽相同。

i) 微振效应 微振疲劳是由微振效应引起的一种疲劳破坏现象。所谓微振效应（作用）可定义如下：两个紧密接触的固体匹配面之间在往复振动滑动情况下发生的机械和化学的综合效应（作用），而作用在两个匹配面上的正应力须足够大，往复振动滑动的振幅应能使表面产生微振磨粒。影响微振效应的因素众多，其中最主要的有匹配面之间的相对位移量，匹配面之间的压力值及其分布，匹配面双方的材料及其表面状态，振动频率，温度，环境，匹配面之间的润滑等。

微振疲劳裂纹的扩展起始于微振效应引起的损伤区，这种表面损伤有以下几种形式：磨粒起坑；粗糙表面接触形成的微裂纹；摩擦诱生应力导致的微裂纹；次表面循环切应力导致的表面分层等。各种机制给匹配面表面带来的损伤形貌（即微裂纹形貌）各不相同。磨粒犁沟机制认为，由微振产生的磨粒与粗糙表面的微振运动，使表面上产生与运动方向相平行的微细犁沟。粗糙表面接触微裂纹萌生机制认为，粗糙的匹配表面上的一些突起点（表面波峰）之间在相对运动时产生接触力，这些突起点不可能经过几次循环就被切断，这样经过若干次微振循环后，便在粗糙表面的波谷应力集中处出现与运动方向垂直的微细裂纹。

匹配面上产生微振微裂纹之后，在构件承受的外加循环载荷作用下，其中的一个或几个裂纹开裂扩展，直至导致最终的疲劳断裂。

2) 低周疲劳 如上所述，疲劳断裂寿命随循环应力的增高而降低，当循环应力增高并超过材料的屈服强度时，$S-N$ 曲线上的这一线段几乎与横轴平行。这表明当试样发生全面屈服之后，断裂寿命的变化对循环应力的高低已不敏感，这种现象通常在寿命低于 10^3 次的范围内，所以把低于低寿命范围内的疲劳称为低周疲劳（或应变疲劳），试验研究发现，在低寿命范围内，应变范围（$\Delta\varepsilon$）这个参量对循环周次很敏感，所以在低周范围内，不再采用应力范围而是采用应变范围－寿命（$\Delta\varepsilon-N$）曲线来描述材料的低周疲劳性能。

对于低周疲劳的研究范围和深度，目前远不如对高周疲劳的研究，这里仅就目前能收集到的资料，讨论一些影响低周疲劳性能的因素。

① 加载频率 与高周疲劳相比，加载频率对低周疲劳性能有着更为显著的影响。有的试验数据表明，低周疲劳强度随着加载频率的降低而逐渐下降。

② 温度效应 在一般情况下，低周疲劳抗力随温度的升高而下降。但是对于一些高温热强合金，低周疲劳强度与温度之间不一定呈单调变化关系。因此对于一些高温热强合金，其高温疲劳强度会出现峰值，峰值的高低及其所处的温度区间，取决于材料的热强性。

③ 驻留时间 所谓驻留时间就是交变载荷在最高应变（或应力）、高低应变（或应力）下保持恒定的时间。驻留时间对材料的低周疲劳性能有很大的影响。无论是提升驻留时间还是压缩驻留时间，均降低材料的疲劳断裂寿命。

④ 缺口效应 与高周疲劳的情况相类似，应力（应变）集中能减低应变疲劳强度。但是，与高周疲劳不同，缺口应力集中系数的变化对应变疲劳强度的影响无显著差别，这是缺口对应变疲劳强度影响的一个特点。

⑤ 平均应变 平均应变对铝合金和钢的低周疲劳强度的影响是，拉伸和压缩平均应变均降低低周疲劳强度，并且只要应变比相同，相同的拉伸平均应变和压缩平均应变对低周疲劳性能具有相同的影响。

(2) 循环硬化和循环软化

材料的组织结构在低周应变循环作用下发生变化，这种变化又反过来影响材料的低周疲劳性质。通常用恒应变范围 $\Delta\varepsilon$ 的试验方法研究材料的循环特性。在恒应变条件下，循环的初期阶段只有在不断地增高应力范围时 $\Delta\sigma$ 才能维持恒定的 $\Delta\varepsilon$，这种现象叫做循环硬化，反之，如在循环初期阶段，只有在不断降低 $\Delta\sigma$ 时才能维持恒定的 $\Delta\varepsilon$，这种现象叫做循环软化。根据疲劳过程中循环应力（或应变）的变化，可将金属材料分为三大类：循环硬化材料，循环软化材料以及循环稳定材料。

大部分金属材料的循环硬化（软化）并不是贯穿于整个疲劳断裂的全过程，而只发生于循环的初期阶段，经过一段循环之后逐渐趋于稳定，即达到饱和状态，此时滞后回线也处于一个稳定的环。通常，达到稳定滞后环的循环数大约是疲劳断裂总寿命的10%到30%。

循环应力与应变之间有如下关系：

$$\sigma_a = k'\left(\frac{\Delta\varepsilon_p}{2}\right)^{n'} \tag{5.3-12}$$

式中，σ_a 为应力振幅，$\sigma_a=\Delta\sigma/2$；$\Delta\varepsilon_p$ 为塑性应变范围；n' 为循环应变硬化指数，其值为0.025～0.25之间；k' 为与材料相关的常数。

表5.3-7列出了若干常用结构材料的单调和循环加载条件下的 $\sigma_{0.2}$、n、n' 等数据及其循环特性。

(3) 疲劳断裂的微观规律

如上所述，试样和构件的疲劳裂纹大都萌生于外表面，但其萌生方式（机制）却是各式各样的。对于实际构件，各种加工工艺给表面带来的各种形式的类裂纹，往往就是导致构件疲劳失效的先天性疲劳裂纹源。本节所要阐述的是一个经过精细抛光的光滑表面上萌生疲劳裂纹（即疲劳源）的一

表 5.3-7 几种材料的循环应变特性

材料	状态	$\sigma_{0.2m}/\sigma_{0.2c}$	n/n'	循环特性
OFHC 钢	充分退火	21/140	0.40/0.15	硬化
	不充分退火	259/203	0.13/0.16	稳定
	冷作硬化	350/238	0.10/0.12	软化
2024 铝合金	自然时效	308/455	0.20/0.11	硬化
7075 铝合金	人工时效	476/525	0.11/0.11	硬化
Man - Ten 钢	验收状态	385/350	0.15/0.16	软化和硬化
Ti - 8Al - 1Mo - 1V	淬火并回火，HV = 350	1 190/770	0.066/0.14	软化
Waspaloy	双重退火	1 015/805	0.078/0.14	软化和硬化
SAE1045 钢	退火	553/714	0.11/0.17	硬化
	淬火并回火，HV = 595	1 890/1 750	0.071/0.14	稳定
	淬火并回火，HV = 500	1 715/1 295	0.047/0.12	软化
	淬火并回火，HV = 450	1 540/980	0.041/0.15	软化
	淬火并回火，HV = 390	1 295/770	0.041/0.17	软化
SAE4142 钢	淬火，HV = 670	1 645/ -	0.14/ -	硬化
	淬火并回火，HV = 560	1 715/1 295	0.092/0.13	稳定
	淬火并回火，HV = 475	1 295/1 365	0.048/0.12	软化
	淬火并回火，HV = 450	1 610/1 085	0.040/0.17	软化
	淬火并回火，HV = 380	1 400/840	0.051/0.18	软化

些主要机制。在此情况下，疲劳裂纹可能在表面滑移带上、晶粒边界或孪晶界上、第二相与基体的界面等处萌生，而每种萌生机制又有其各自的特点。

1）滑移带　与轴向单调加载的情况一样，循环加载也使金属表面产生滑移带，不过它们的分布及形态彼此不尽相同而已。单调加载下出现的滑移带分布均匀、密集，而循环加载下出现的滑移带分布极不均匀且粗大，在形态上二者表现出明显的区别。

实际构件发生疲劳失效的原因众多，选材不当可能是其中的原因之一。屈服强度低的材料容易在疲劳过程中过早地出现滑移带，并在驻留带上或挤入挤出处萌生疲劳裂纹。因此，在高周疲劳条件下，材料的疲劳强度在很大程度上取决于其屈服强度。屈服强度越高，滑移越困难，所以疲劳裂纹也就更不容易萌生。

2）孪晶界与晶界　在有共格孪晶界存在时，驻留滑移带通常优先在此出现，并导致疲劳裂纹的萌生。在锆、锑、镉、铜、金、镍以及 α - 黄铜等金属中均先后观察到这种现象。有人认为，这种现象与滑移带和孪晶界相交处发生的“挤出 - 挤入”有关。试验表明，孪晶尖端的平均位移远大于滑移带内的位移。

多晶体的晶界也是疲劳裂纹容易萌生的地方。萌生于晶界上的疲劳裂纹是晶体发生大量滑移的结果。滑移面的领先位错在晶界上受阻，由此形成位错塞积。位错的塞积在晶界造成应力集中，当此应力峰值达到某临界值时便引起晶界开裂，材料的晶粒尺寸越大，晶粒内可能形成的位错塞积越长，晶界上的应变量越大，更易形成疲劳裂纹。晶粒越细则高周疲劳强度越高，因此，细化晶粒（或亚晶粒）不仅能够抑制疲劳裂纹在晶界上萌生，同时还能使滑移线分布均匀，延迟或抑制在滑移带上萌生疲劳裂纹。

应该指出，疲劳裂纹在晶界上的萌生与温度颇有密切关系。在 300℃ 下纯铝的晶界开裂现象颇为显著，25℃ 和 - 73℃ 下则大为减少，而在 - 180℃ 下就不再出现晶界开裂。一些镍基和铁基高温合金，高温下（> 600℃）的疲劳裂纹往往萌生于晶界，而在较低的温度或室温下，则萌生于晶内。由此可见，疲劳裂纹在晶界上的萌生，还受到材料性质、温度等条件的影响。

3）相界面　工业上使用的许多金属材料，都不同程度地存在各种非金属夹杂物。对于钢，这类夹杂物如硅酸盐、硫化物、氧化物、氮化物等等。此外，为了强化金属材料往往使材料中形成第二相强化相（如弥散状的或大颗粒状的强化相）。这些夹杂物和强化相与基体相交的界面也能成为疲劳裂纹优先萌生的部位。有的试验表明，夹杂物和强化相的形态、数量及其分布对材料的疲劳性能有相当显著的影响。硬度高于基体的夹杂物，一般来说对材料的疲劳性能是有害的，对于高强钢尤其如此。存在于构件表面的大块夹杂物，往往是疲劳裂纹引发的场所。

1.3.3 疲劳断裂失效门槛值预测

如前所述，对于存在某一裂纹尺寸的构件，只有疲劳裂纹尖端的应力强度因子达到或超过门槛值 ΔK_{th} 时，裂纹才会在交变应力作用下发生扩展。当裂纹尖端应力强度因子小于 ΔK_{th} 时，交变应力不足以使裂纹以每周一个原子间隔的速率扩展，此时裂纹仅仅是张开和闭合，而不出现宏观的裂纹扩展，其相应的 da/dN 等于或低于 10^{-11} m/周。

门槛值 ΔK_{th} 亦称为界限应力强度因子，它和 da/dN 一样，是反映带裂纹构件抗疲劳性能的一个重要指标。在物理意义上它和缺口试样的疲劳极限概念十分相似。它是工程设计、选材和安全评定不可缺少的重要指标。在实际结构中，对于一些重要的受力构件，需要根据材料的 ΔK_{th} 来确定其工作应力水平，即根据构件存在的初始裂纹尺寸或无损探伤所能发现的初始裂纹尺寸 a_0，再根据门槛值去求出该裂纹在疲劳载荷下不发生扩展的界限应力，按界限应力确定构件的允许工作载荷。或者反之，根据 ΔK_{th} 值确定不发生扩展的允许裂纹尺寸等。

门槛值和材料的微观结构密切相关，同时受应力比、加载的历史、环境和热处理工艺等的影响。

为了估算材料的门槛值，研究者提出了一些估算的经验公式，日本的北川英夫用 16 种金属材料进行试验，提出了用材料的 E 和 σ_s 计算门槛值的公式：

$$\Delta K_{th} = AE - B\sigma_s \tag{5.3-13}$$

式中，常数 $A = 2.0 \times 10^{-3}\ \text{mm}^{1/2}$；$B = 0.210\ \text{mm}^{1/2}$。

于崇华等根据疲劳裂纹尖端的应变分布和塑性区尺寸的分析提出如下公式：

$$\Delta K_{th} = 2\sigma_s (E\varepsilon_f / 2\sigma_s)^{(n+1)/2} \sqrt{2\pi\rho} \tag{5.3-14}$$

式中，ε_f 为静拉伸断裂的真实应变，即 $\varepsilon_f = \ln[1/(1-\psi)]$，

ψ 为断面收缩率；ρ 为裂纹尖端的曲率半径。

当裂纹处于初始扩展时，可近似假定 $n=1$，则 ΔK_{th} 公式为：

$$\Delta K_{th}=E\varepsilon_f\sqrt{2\pi\rho_{min}} \qquad (5.3\text{-}15)$$

式中，ρ_{min} 为裂纹扩展时其尖端的临界曲率半径，相当于一个原子间隔。微观结构对门槛值 ΔK_{th} 的影响已经有了比较明确的结论，人们指出门槛值 ΔK_{th} 随着晶粒尺寸的增大而增加，这是因为在较大晶粒中，在它开裂之前有可能积累更多的塑性变形。反之，光滑试样的疲劳极限则随着晶粒尺寸的的减少而增大。

约达（Yoder）等提出

$$\Delta K_{th}=5.5\sigma_s\sqrt{d} \qquad (5.3\text{-}16)$$

式中，d 为材料的晶粒尺寸，m。该式适用于预估低碳钢的 ΔK_{th} 值时有良好的精度。

在估算铁素体钢时也可用下式：

$$\Delta K_{th}=3.8+1.4\times10^8\sqrt{d} \qquad (5.3\text{-}17)$$

式中，d 为材料的晶粒尺寸，m；ΔK_{th} 的常用单位为 $MPa\cdot m^{1/2}$。

表5.3-8为用降载法所测定的14种常用材料的金属材料在 $R=0.1$，$\alpha/W=0.3$ 时的门槛值。

表5.3-8　材料的疲劳裂纹扩展门槛值

材料	$\Delta K_{th}/(MN/m^{3/2})$	应力比 R 的影响
40车轴钢	11.19	$\Delta K_{th}=12.47-12.41R, R\leqslant0.65$ $\Delta K_{th}=4.34, R>0.65$
50车轴钢	9.18	$\Delta K_{th}=9.83-8.91R, R\leqslant0.7$ $\Delta K_{th}=3.61, R>0.7$
合金灰铸铁	13.86	$\Delta K_{th}=14.96-11.18R$
AP1轨钢	8.37	$\Delta K_{th}=9.33-5.8ZR, R\leqslant0.58$ $\Delta K_{th}=6.17, R>0.58$
20CrMnTi	9.61～10.23	$\Delta K_{th}=11.33(1-0.7R-0.21R^2)$
35CrMo	12.18	$\Delta K_{th}=14.42(1-R)$
42CrMo	7.13	$\Delta K_{th}=9.17-6.98R$
50CrMoA	9.3～10.85	$\Delta K_{th}=11.78(1-R)^{0.3}$
16Mn	8.06～8.37	
球墨铸铁	14.57	
QT-80	11.32	
22CrMnMo	10.85～11.78	
ZG25	11.78	
15MnVN		厚板 $\Delta K_{th}=9.08(1-R)$ 薄板 $\Delta K_{th}=11.2(1-R), R\leqslant0.6$

大量实验表明，应力比或平均应力对各种材料的门槛值有十分显著的影响。普遍认为，ΔK_{th} 随 R 的升高而降低。

众多研究者提出一些 ΔK_{th} 和 R 的关系式，主要为：

$$\Delta K_{th}=\Delta K_{th0}(1-R) \qquad (5.3\text{-}18)$$

$$\Delta K_{th}=\Delta K_{th0}(1-R)^{\gamma} \qquad (5.3\text{-}19)$$

式中，ΔK_{th0} 为 $R=0$ 时的门槛值；γ 为材料常数。

Garwood提出一个非常简捷的估算钢材门槛值下限的公式，即

$$\Delta K_{th}=6.0-4.55R \qquad (5.3\text{-}20)$$

表面看式（5.3-18）、式（5.3-19）、式（5.3-20）格式截然不同，但它们能很好地拟合大多数实验数据，仅仅是 $R\to1$ 时差别较大，事实上 $R>0.6$ 以后的数据很少。

相对裂纹长度 α/W 对于 ΔK_{th} 也有一定的影响，对于大多数材料来讲，随 α/W 增加，门槛值 ΔK_{th} 也略有增加，当 α/W 在0.2～0.5的范围内时，ΔK_{th} 基本可看作常数。当 $\alpha/W<0.2$ 时，ΔK_{th} 有下降的现象。

试样类型不同，测得的 ΔK_{th} 也有所不同。根据对42CrMo钢的测试结果，四点弯曲试样测得 ΔK_{th} 最大，三点弯曲试样次之，再次为紧凑拉伸试样，中心裂纹试样的 ΔK_{th} 值最低。

热处理工艺对 ΔK_{th} 有密切的影响。村上里一对35B钢试验，证明 ΔK_{th} 与回火温度有关系，一般 ΔK_{th} 随回火温度的增加而减小。西安交通大学金属材料及强度研究所深入研究了热处理工艺对 ΔK_{th} 的影响，并测定了45Cr钢多种热处理工艺的门槛值，证明了 ΔK_{th} 并不是随回火温度单调地变化，而是在中温回火时有一个峰值。

喷丸强化能大大提高材料的裂纹扩展门槛值。

1.4　磨损失效预测

机械构件的磨损失效是逐步发展的过程。有的构件只能工作几个小时，有的则可以使用几年甚至几十年的时间才需要更换。构件的使用条件不同，技术要求不同，磨损特征也千差万别。构件磨损的共同特征是表面受到损伤，发生了表面材料迁移。表面损伤或材料迁移到什么程度判为构件失效，根据构件的工作条件和要求是大不相同的，粗略分为三类：①与土砂、矿石相互作用的构件，多数情况下允许很大的磨损量，其磨损量可达到构件原始质量的10%～100%；②设备中要求尺寸配合的构件磨损后，尽管磨损量很小，用质量几乎测定不出，但影响设备的精度和效率，或增加其震动，这种情况必须更换构件。还有一种情况，虽然没有配合精度要求，但发生一定量的磨损（比如几微米到几百微米的磨损量）会影响设备的运行而造成失效；③磨损过程中引起构件其他类型的损坏，如断裂，或磨损加速了构件的腐蚀而失效。

（1）磨损的系统特性

材料的耐磨性不是材料的固有性质，而是某个磨损系统的综合性质，或称为系统性质。由机械构件、对磨物及工件运转参量和环境条件组成的磨损系统决定该构件的使用寿命。如果系统中某个参量变化，就会影响到材料的耐磨性。不同材料制成的构件所受影响大小不同，这是由于磨损发生过程的不同，即磨损机理的差别所造成的，说明磨损机理控制了材料的耐磨性变化的规律。

（2）磨损过程的动态性质

开始产生磨损或孕育着磨损的发生。磨损方式不同，开始产生磨损的时间长短不一。抗磨料磨损构件，从与磨料接触时就开始了构件的磨损，而齿轮，轴承等一类构件要工作一段时间以后才发生磨损，一旦表面产生材料流失，磨损就加速直至失效。

此外，当磨料颗粒刺入材料表面的瞬间，材料的性能与磨料颗粒在构件表面滑行过程中材料的性能也有区别。磨损

热的影响是个极为重要的因素，由于磨屑很小，热容小，温升的影响明显，但是磨料与磨损表面相互作用的微区温升也不容忽略，也可达到1 000℃。许多文献介绍热剪切层的出现。这些都是磨损过程中动态特性的表现。磨损热的形成可以使原始较硬的材料软化，可以使硬而脆的材料发生塑性变形，因而改变了材料原始磨损性能。

(3) 磨损类型和磨损机制的转化特性

机械构件运行过程中所造成磨损的形式是可能转化的。发动机中的凸轮与挺杆；活塞环与活塞等摩擦副，当润滑条件不好时，可能发生黏着磨损，如果较好的润滑就可能产生应变疲劳和应力疲劳磨损，如果空气中含有大量的尘埃，也可能主要是磨料磨损造成构件失效。如果有黏着剥落下来的磨屑或是疲劳产生的磨屑，在机件连续运转过程中，这些磨屑也可能起磨料的作用产生磨料磨损。

磨损类型的确定只有对构件磨损表面，磨损亚表层和磨屑进行分析以后，再加上环境条件的综合分析，提出可靠的依据后才能判定。实用构件的磨损失效往往不是一种磨损类型，而是几种类型的综合作用。

各类磨损形式中往往有几种磨损机制同时在起作用。如磨料磨损中就有切削、犁沟变形和断裂三种方式。不同的工况条件或不同材料的组织结构，产生磨损的主要机制也是不同的，随着工况条件的变化，三种机制所占的比例也不相同。

1.4.1 磨损失效的分类

(1) 摩擦

1) 摩擦的分类 摩擦的分类有三种。

① 按摩擦副的运动形式分类

a) 滑动摩擦，指接触表面相对滑动的摩擦。

b) 滚动摩擦，指物体在力矩作用下，沿接触面滚动时的摩擦。

② 按摩擦副的运动状态分类

a) 静摩擦，当物体在外力作用下，对另一物体具有相对运动的趋势，但还处于静止的临界状态时的摩擦，叫静摩擦。

b) 动摩擦，物体受力的作用，超过静止的临界状态而沿另一物体表面发生相对运动的摩擦，叫动摩擦。

③ 按摩擦的性质分类

a) 干摩擦，指物体表面无任何润滑剂存在时的摩擦。

b) 边界摩擦，指两物体表面被一种具有分层结构和润滑性能的边界膜所分开的摩擦。

c) 流体摩擦，指两物体表面被润滑剂膜完全隔开时的摩擦。

d) 混合摩擦，指半干摩擦和半流体摩擦的混合状态，也就是虽有润滑但效果很差的一类摩擦。

2) 摩擦与表面特征和亚表层特性的关系

摩擦的存在与物体的表面特征和亚表层的特性有着密切关系，也与表面相互接触时的力学行为和运动过程中表面和亚表层的变化有着密切的关系。

机械构件多半经过机械加工，表面不可能绝对平整光滑，是由众多的微凸体和凹谷组成的。这种粗糙表面特性对摩擦，磨损和润滑都有重要的影响。

机械构件加工过程中，表面层的材料经受着相当大的塑性变形，必然引起加工硬化现象及内应力，甚至产生表面裂纹。同样，两物体对磨过程中表面轮廓发生变化，表面以下的相当厚度的一层材料也会发生特殊的物理化学变化。最表层是一层氧化膜，下面的一层称为贝氏层（Bielby层），它是一层由于切削而使分子层熔化或相变而产生的非晶态层或微晶粒层。也有人称这层为绝热剪切层。加工时产生的热量不能迅速向外传递，而使表层分子熔化或相变，形成一层具有特殊物理化学性能的硬层。贝氏层下是一定变形度的变形层，变形层的厚度和这层的变形度与加工方式和材料性能有关。变形度由贝氏层向构件内部逐渐减弱，直到构件的原始组织。这表层厚度有时可达0.1～0.3 mm。

(2) 磨损

1) 分类 磨损的分类方法很多，概括起来有三种分类方法。

① 按对磨物特性分类 美国金属学会出版的金属手册第七版、第八版、第九版都采用按对磨物特性分类。

② 按磨损发生过程分类 由于按对磨物特性分类只考虑了对磨物和工作条件，但没有明确指出磨损的实际过程，所以也有人从20世纪50年代开始转向根据磨损发生过程分类，Peterson提出的，即按磨损碎屑脱离磨损表面的特性分类，他认为磨屑脱离母体的实际过程最能体现磨损过程本质——即磨损机制。他把磨损分为十类，而且认为黏着、腐蚀、切削、塑性变形、疲劳五种为主要磨损类型，其他如表面开裂、表面反应、撕裂、熔化、电化学五种，是在特殊情况下出现的磨损类型。

③ 按磨损机制分类 Burwell和Strang提出按磨损机制分类为：黏着磨损；磨粒磨损；疲劳磨损；腐蚀磨损。

在机械构件的磨损过程中往往同时存在几种磨损机制，如磨料磨损过程就有切削、变形、断裂和疲劳四种机制同时存在，而且材料性质改变或工况条件改变都会影响磨损过程的本质。

目前文献资料常用的分类方法是：磨料磨损；黏着磨损；疲劳磨损；冲蚀磨损；微振磨损。

2) 摩擦与磨损 摩擦与磨损几乎是同时发生的，二者既有联系，又有区别。人们通常会认为摩擦因数大的摩擦副，磨损一定严重，事实并非如此。摩擦主要涉及材料的变形，而磨损则涉及断裂、转移，甚至熔化。一般情况下，在一个系统中如果摩擦力有变化，就说明马上要发生恶性的磨损现象了，因之摩擦力的变化，就是要发生恶性磨损的早期讯号。

另一方面，也存在只有摩擦而几乎没有磨损的情况。例如轴承等的设计准则便是由此出发的。由表5.3-9中所列的数据也可看出摩擦和磨损的关系不是一个简单的问题。由摩擦去判断磨损是不妥当的。在磨料磨损的情况下，摩擦主要是由于变形消耗能量，这部分能量所占比例很大，而磨损过程中形成磨屑还要附加一个断裂功，N.P.Suh等认为这比变形功小得多，他们在讨论磨粒尺寸对表层变形、犁沟和切削功所需能量的比例时，指出表层变形所占的能量消耗比例是很大的。

表5.3-9 几种摩擦副的摩擦因数和磨损系数

摩擦副（在空气中）	摩擦因数 μ	磨损系数 K
金对金	2.5	10^{-1}
黄铜对硬钢	0.3	10^{-3}
聚四氟乙烯对硬钢	0.15	2×10^{-5}
聚乙烯对硬钢	0.6	10^{-7}

1.4.2 各类磨损失效的特点和预测

(1) 磨料磨损

中国机械工程学会摩擦学学会，对摩擦学名词术语Abrasive Wear一词定义为磨料（磨粒）磨损。从事摩擦和润滑研究的学者认为，配合表面之间，在相对运动过程中，把

由不期而来的硬颗粒或由表面微凸体的作用造成表面损伤的磨损称为磨粒磨损，认为磨粒磨损是一种机制，主要特征是表面被犁削形成沟槽。另外从事研究土砂、矿石、岩石等物料对构件的磨损也称为磨料磨损，这些物料与构件表面相互作用的方式不只是滑动，还有冲击、滚动、冲刷等方式，认为磨料磨损是一类磨损方式，而不是一种磨损机制，这强调的是对磨物和磨料与构件表面的作用方式。往往把那些由于硬质点使表面形成沟槽的磨损也包括在这类磨损之中。

1）磨屑的形成过程

① 滑动磨料磨损中磨屑的形成　当磨粒承受一载荷时，可分解为法向分力和切向分力。法向分力使磨粒刺入材料表面，切向分力使磨粒沿平行于表面方向滑动。

如果磨粒棱角锐利，又具有合适的角度，可以形成切削状的显微切削磨屑。磨粒在金属上形成切槽，切槽的两边及底部都要产生一定的变形层。变形层的厚度比切削加工构件的表面变形层厚一些，因为磨粒的棱角比刀具刃口要钝得多。

如果磨粒的棱角不适合切削，只能在样品表面形成犁沟变形，即将金属推向磨粒的前方，或推向磨粒运动方向的一侧或两侧。如果多个磨粒连续与金属相互作用，不适合切削的磨粒多次使金属变形，形成经多次辗压变形的台阶。某些变形台阶以平行于磨痕的方向产生裂纹，裂纹扩展形成薄片状磨屑，也可能产生垂直于磨痕方向裂纹，裂纹扩展形成磨屑，这与材料的性能有密切关系。抗应变疲劳性能高的材料，可承受较多次数的辗压变形；而抗应变疲劳性能低的材料，少数几次变形后便因超过材料强度极限而开裂、脱落。所以材料的耐磨性不只是决定于材料的硬度，还决定于材料的应变疲劳性能。

综上所述，当磨料颗粒沿磨损表面滑动时，所形成的磨屑有三种主要方式：

a）显微切削；

b）多次变形碎薄片；

c）硬质点的断裂和脱落。

② 变形磨料磨损或高温中磨屑的形成　与滑动磨料磨损过程中磨屑形成机制完全不同的磨损方式是变形磨料磨损。例如颚式破碎机的齿板、圆锥式破碎机的破碎壁、球磨机的磨球、衬板就是这种磨损的典型机械构件。磨料与构件表面相应作用时，如果磨料硬度超过材料硬度，磨料的棱角将刺入材料表面使材料发生塑性流动而挤到凹坑的四周，材料变形到一定程度，磨料粒被压碎。多个颗粒与坑周围变形材料接触，又使之变形、塑性流动、加工硬化程度进一步增加，当超过材料的应变疲劳极限时形成磨屑脱落。

③ 腐蚀磨料磨损中磨屑的形成　这种磨损是不产生显微切削和表面变形的磨料磨损的。例如水中耕作构件的磨损就是腐蚀与磨粒的擦伤综合作用的结果。磨屑多是呈氧化物微细粉末状。特别是拖拉机履带一类构件，水砂中运动之后在空气中停留一段时间后又进入水砂中，磨损加快，甚至比埋入水砂中连续运行所造成的磨损更快。用高锰钢制成的拖拉机履带板，在北方旱田耕作，可使用2 000 h，但在南方水田中耕作，其寿命仅500 h。这是一种低应力细磨料对腐蚀产物的研磨，使新鲜表面裸露出来。腐蚀－研磨－再腐蚀的循环过程，腐蚀加速了磨损，磨损又促进了进一步的腐蚀。

这类构件分析研究很少，是磨损失效分析的薄弱部分。

2）材料性能对机械构件的的影响

① 金属材料硬度与耐磨性的关系　纯金属和退火钢的相对耐磨性与硬度是单一的线性关系，硬度越高，相对耐磨性也越高。而淬火回火钢则偏离了线性关系，随着回火温度的升高而硬度降低，导致相对耐磨性也降低，最后落在纯金属与退火钢的线性关系上。如果提高奥氏体化温度或等温处理，则材料的相对耐磨性与硬度的关系将发生变化。

在试验过程中发现材料经过冷作硬化使硬度有所提高，但并不增加它的相对耐磨性，这个结果与生产中易磨损构件表现出的结果一致，即靠常规手段的冷作硬化不能有效地提高材料的抗磨损性能。

② 磨损表面硬度与材料耐磨性　对构件磨损残体的磨损表面进行硬度测定表明：材料被磨损之后，表面的硬度有相当大的提高，提高的幅度与构件的服役条件有关，也与材料特性有关。

③ 磨料硬度和材料硬度比值与材料耐磨性的关系　当磨料的硬度比材料的硬度高很多时，材料的相对磨损并不随磨料硬度的增加，这时磨损只决定于材料的原始硬度。但是，当磨料硬度低于金属材料的原始硬度时，材料的相对磨损急剧减少。

④ 影响材料耐磨性的工况条件

a）磨料颗粒形状　磨料颗粒的形状、粒度和硬度对材料耐磨性有重要影响。

Ⅰ）磨料颗粒形状　磨料颗粒的形状对材料从构件表面的流失有重要影响。磨粒迎角又称侵入角或倾斜角，系指磨粒的前导面与构件表面法线之间的夹角，用 α 表示，α 可能是正、零或负值，也有人将迎角定义为磨粒前导面与构件表面之间的夹角。

临界迎角又称临界侵入角，即由切削磨损转变为犁沟变形磨损的最大角度。用 α_c 表示。α_c 决定于材料性质、构件表面状态（是否有润滑剂）和磨料性质。当磨粒前导面与表面法线之间的夹角 $\alpha > \alpha_c$ 时，产生变形磨损，此时摩擦力大、变形层厚、磨损率低。反之 $\alpha < \alpha_c$ 时，产生切削磨损，摩擦力小、变形层较薄，磨损率大。

材料进行磨损试验时，砂纸（布）上的磨粒与样品表面法线之间的夹角是统计分布的。部分磨粒对金属产生切削机理控制的磨损，部分磨粒使金属产生犁沟变形磨损，少量磨粒只在样品表面上摩擦不产生磨屑。所以磨粒是否锐利，决定了材料的磨损机理。

Ⅱ）磨粒硬度及压碎强度　磨料颗粒的硬度与材料的耐磨性有着非常重要的关系。天然或人造磨料颗粒可能在很大的范围内变化，它们的数值可在有关资料中查到。磨粒的另一性能指标是压碎强度。例如白刚玉和棕刚玉，它们的硬度十分接近，但白刚玉较脆，含杂质较多的棕刚玉韧性较好，两种磨料对金属造成的磨损也是不一样的，特别是接触应力较大的情况下，白刚玉颗粒易被压碎，棕刚玉则很少被压碎，对磨损影响明显不同。

Ⅲ）粒度　许多研究者用不同的磨料磨损试验机研究了磨粒尺寸对材料磨损率的影响。他们都发现，当材料和磨料品种不变，而改变磨粒颗粒尺寸时，存在一个磨料的临界尺寸，比临界尺寸小的磨料颗粒，磨损率随磨粒尺寸增加以较快的速率增加。磨料颗粒尺寸小于临界尺寸时，在较低的载荷下，材料磨损率与磨粒大小无关，

在高载荷下，磨损率随着磨粒尺寸增加，以一定的斜率增加。临界尺寸约在60～100 μm范围内。为什么会存在这种磨粒的临界尺寸，虽有些解释，但均不能令人满意。特别是磨粒尺寸小于临界尺寸时，为什么磨损率随粒度增加的原因应进一步研究。

b）磨料与材料的相互作用状态　磨料与材料相互作用状态对材料耐磨性的影响包括载荷大小、接触方式、相对滑动速度对材料耐磨性的影响。因为载荷增大、相对滑动速度加快，使材料表面微区的温度急剧提高，对材料的磨损特性有重要影响。

Ⅰ）载荷的影响　材料的磨损体积与载荷大小成正比，这符合多数的使用条件下的磨损规律。在磨料不破碎的情况

下，载荷增大使磨沟深度增加，尺寸变宽、磨屑变大、磨损体积增加。过大的载荷使磨粒破碎，或使磨损表面的温度升高使韧性增加，对磨损率也有影响。

Ⅱ）相对速度的影响　随着相对速度增加，材料的磨损率增大，这是一般的规律，材料的品种不同，磨损变化的幅度也不同。通过对材料磨屑和磨损表面的分析可知，随着速度增加磨屑的内部组织和磨损表面内部组织，发生很大的变化。如残余奥氏体量可逐渐变少，又由少变多，甚至超过样品的原始组织中的残余奥氏体量。

(2) 冲蚀磨损

冲蚀磨损是一种常见的磨损形式。灰渣泵的护套和护板的磨损、火箭助推器尾翼排气管的磨损、直升飞机转子和燃气轮机叶片的磨损、靠褐煤粉运动的气轮发电机和靠增压式沸腾锅炉排气运行的气轮发电机的磨损、悬浮粒子的管道输送管发电设备的锅炉管、高速磨煤机的打击板等构件的磨损都可归于冲蚀磨损。

遭受气体或液体中固体微粒打击、冲刷或浸蚀的构件表面材料的不断流失现象称为冲蚀磨损。有人根据粒子流冲击构件表面的角度大小，把这类磨损方式划分为冲刷磨损和撞击磨损。冲刷磨损指的是磨粒运动方向与材料表面之间夹角较小，甚至接近平行，高速运动的微粒沿构件表面滑动并划伤表面，以切削或犁削和脆断等机制造成材料磨损。

1) 冲蚀磨损机制　虽然对固体粒子磨蚀过程中材料的流失机制做了不少研究，但是许多的磨损理论是根据单颗粒冲击坑实验观察总结出来的。近年来的研究表明，材料的流失通常都是复杂的多次冲击作用的结果，而且必须考虑冲蚀磨损过程中金属表面变形和粒子嵌入对磨蚀表面层力学性能的影响。

A. Hammarsten 对冲蚀磨损产物——磨屑和磨损表面及亚表层进行分析，提出了磨损机制。

冲蚀磨损的磨屑有三种类型：

① 由于冲击粒子的切削作用，材料以切削屑的形式脱离表面；

② 由于冲击时发生脆性断裂，材料以薄片屑的形式从冲击形成的层状表面脱离；

③ 冲击时由于表面屑状物或其他凸起部分发生断裂，材料以簇团屑的形式脱离表面。

第一种类型的磨损前面已作过叙述。薄片屑脱离表明是屡次冲击综合作用的结果。一般在单颗粒冲击条件下，表面材料仅被推到坑的边缘，但并不发生材料的流失，随后连续不断的冲击揉搓表面层，而形成一种强烈变形的表面层状结构，最后表面加工硬化，造成脆性断裂，薄片屑脱离表面。表面的层状结构和碎砂粒嵌入都有助于断裂。粒子嵌入的作用之一就是在表面形成一些“铁砧”，在遭受粒子冲击时，这些铁砧就是应力集中区。

簇团屑脱离是由于磨蚀形成了极度不平的表面形貌的结果。靶面吸收的大部分冲击能都转化为热能，耗散在材料内部。如果表面凸起部分受到粒子的冲击，表面的几何形状就会降低冲击点的散热速度，因而在局部地区产生高温，使凸起部分软化，这样，凸起部分就可能由于断裂而脱离表面。

2) 影响冲蚀磨损率的因素

a) 磨粒形状的影响　多角状磨粒比球状圆滑磨粒在同样条件下有更严重的冲蚀磨损。如果多角状磨粒的冲击角是45°，则比球状圆滑磨粒的磨损量大四倍，甚至多角状磨粒在硬度较低时也比硬球状圆滑磨粒有更大的磨损能力。球状圆滑磨粒是以犁削变形方式为主，而多角形磨粒则以切削为主，而且磨粒转动方位不同（向前或向后）使切削的深浅有所区别，一般磨粒在冲击接触过程中附带转动的几率较低。

磨粒形状对发生最大磨损的冲击角有一定影响。试验表明，对棱角状 SiC 和 Al_2O_3 磨粒 $\alpha_{max} \approx 16°$，而钢球产生的最大磨损角为 28°。许多实验结果证明，一般韧性材料 α_{max} 在 16°～30°之间，原因之一是与磨粒形状有关。对脆性材料而言，则磨粒形状对 α_{max} 影响较小。

b) 磨粒大小的影响　磨粒大小对冲蚀磨损有重要的影响，这可从两个方面来分析。

① 磨粒尺寸效应。与磨料磨损类似，冲蚀磨损也存在着临界尺寸。对这种尺寸的解释认为磨粒破碎所产生的二次磨损有关。当磨粒的粒度小于一定尺寸时，由于空气动力学效应，使微粒绕靶偏转而不冲击，冲击功很小，不产生磨损，有没有磨粒破碎，因而不产生二次磨损。

② 脆性材料（如玻璃）的磨损率与粒子尺寸呈现指数关系：$E = K_1 d^m$。对于退火态的玻璃，$m = 2$。K_1 为常数。

③ 磨粒的硬度　一般认为，磨损率与冲击粒子的硬度有关。对许多具有不同级硬度的天然和工业材料进行了磨蚀性测量。实验所用粒子尺寸为 125～150 μm，速度为 130 m/s，磨蚀率与硬度的关系可以表达为 $\epsilon = K_3 H^{2.3}$（K_3 为某一常数值）。由于软粒子易被磨圆，尚未得到精确关系，但锐利的软磨粒比钝的硬磨粒往往有更大的冲蚀磨损率。

c) 冲击角的影响　靶表面和粒子运动轨迹之夹角称为冲击角。冲击角的影响与材料类型有关。延性材料在 20°～30°角冲击时，破坏最大。而脆性材料在垂直冲击时破坏最严重。

d) 颗粒速度的影响　在冲蚀磨损过程中最重要的参量是磨粒冲击表面的速度，这是由材料流失公式中动能项所决定的，根据许多研究报告结果得，$\epsilon = Kv^n$ 式中 K 和 n 都是常数。

e) 环境温度的影响　根据不同条件的研究得到两种结果：①随着环境温度上升，材料磨损变得严重，②随着环境温度上升，材料磨损稍有减弱。一般来说，第一种比较容易接受，因为材料在高温下屈服极限下降，磨损同样体积材料所消耗的能量减少。对于同样能量粒子来说，当然在高温下产生更为严重的磨损。但某些研究证明，随着环境温度上升材料的磨损减弱。如铍－铜合金、铝合金、低碳钢等。从室温升温至 600℃，材料磨损都有不同程度的减弱。Tilly 认为这和材料的性能在高温下发生变化有关，铝合金、碳钢在较高温度下具有高的塑性，韧性金属随着延伸率的提高抗冲蚀磨损性亦可提高。11％铬钢的延伸率对温度不敏感，所以磨蚀状况随温度的变化不大。

研究发现在更高温度下（约 980℃），甚至在过氧化情况下不锈钢的磨损同样也会加剧，比室温磨损量约大一个数量级，而且随着温度提高。温度对冲击磨损的影响，有人建议采用材料的半熔点温度（材料熔点温度的一半），作为判定磨损是否加剧的标准。小于半熔点温度，磨损率随温度的升高有所减缓，反之则磨损率增加。

3) 汽蚀

① 发生的原因　汽蚀是磨蚀的一种特殊形式，它的产生是由于材料表面附近的液体中有汽泡产生及破灭造成的，在高速流体中由于压力变化（由高压至低压）就容易产生汽蚀，如螺旋桨、离心泵、水轮机叶片等一类的设备中经常看到这种破坏方式，汽蚀的出现伴随着噪声、振动、工作表面破坏以及机械效率降低等。汽蚀的研究首先从机械构件开始的，但近期引起了水力学工作者的注意，在水力建筑中出现汽蚀，如水坝闸门、溢洞、隧洞、桥墩等脆性材料等均发现汽蚀。由此，促使水力学者加以大力研究。

汽蚀的发生决定于环境的压力、温度和速度。此外增加流体中的含气量，或是流体在那种易于产生旋涡运动的零件外形，都会促使汽蚀的出现。

② 汽蚀的疲劳机理及对抗汽蚀材料的要求　汽蚀损坏

除与零件的线型设计、表面质量（有无缺陷）和加工粗糙度有关外，还与选用材料有关。应通过对汽蚀破坏的力学特性和微观分析来研究破坏的原因。通过微观分析认为，汽蚀作用的主要特征是“机械力作用”，实际工况条件下汽蚀破坏特征可能是腐蚀和机械力联合作用的结果。在一定条件下认为是腐蚀疲劳损坏机制。因为汽蚀作用面积很小，与合金结构、晶粒大小相近，所以汽蚀的机制可以近似地看作为表面微观腐蚀疲劳。合金的晶粒、夹杂在脉冲打击及介质的腐蚀作用影响下，产生龟裂、裂纹扩展和剥落。

金属材料的抗汽蚀性能与下列因素有关。

a）材料具有的强度、弹性极限愈高，则材料的抗剥落能力愈高。即提高材料的强度及塑性对提高抗汽蚀能力是有益的。只是单一的高塑性对抗汽蚀能力并不是最有利的。因而有人认为在保证材料一定塑性条件下，提高材料的硬度是改善抗汽蚀性能的重要措施之一。

b）细晶粒的材料具有高的抗汽蚀性。

c）降低氮、氢、氧和非金属夹杂在材料中的含量，提高铸件的表面质量，均能显著提高抗汽蚀性能。

d）汽蚀与材料化学成分、热处理情况有很大关系。含碳量相同、含有 Cr、Ni 一类合金元素的钢比碳素钢的抗汽蚀性好，双相组织比单相组织的材料抗汽蚀性能好。

e）介质的影响汽蚀破坏的严重性按下列顺序增加。即蒸馏水——自来水——天然河流水——海水。

海水中的汽蚀破坏程度比蒸馏水中的约高 0.5～1.0 倍。

③ 汽蚀与冲蚀的联合作用　在液流中存在着砂粒，则两种磨蚀互相促进，加速机械构件的破坏，如水轮机叶片的破坏的原因是：

a）砂粒由于具有不平整的表面而易附着气体，使汽蚀的汽核数目增加。

b）砂粒在水流中运动时，在砂粒后方形成压力下降区，易形成汽泡。

c）汽泡破灭造成水流的局部强烈扰动，使水流中的泥砂颗粒获得极大的加速度，提高了砂粒的冲击动能。所以一旦汽蚀呈针孔，随后泥砂磨损成凹坑，凹坑又加剧汽蚀作用，而形成较大的鱼鳞坑。

（3）黏着磨损

1）黏着磨损的理论概述　1953 年 J.F.Archard 提出了一个简化的黏着磨损定律。他假设接触是由半径均为 a 的一些相同微凸体所形成的。每个接触点的面积为 πa^2，每个接触点所支承的载荷为 $\sigma_y \pi a^2$（σ_y 为流变应力）当滑动距离为 $2a$ 时，表面能够完全擦过每个微凸体，因而可以假设每个微凸体上所产生的磨屑为半球形，其体积为 $2/3\pi a^2$。

则单位滑动距离的总磨损量为

$$w = \sum \frac{\frac{2}{3}\pi a^2}{2a} = \frac{1}{3}\sum \pi a^2 = \frac{\pi}{3}a^2 n \quad (5.3\text{-}21)$$

式中，n 为接触点的总数，而每个接触点所承受的载荷为 $\sigma_y \pi a^2$。

Archard 的黏着磨损理论广泛为各国学者引用，在某种程度上描述了磨损的一般规律，把磨损与材料性能联系了起来。但该理论假设也存在许多不足之处，应进一步研究完善。

2）机械构件的磨损特点和分类

组成摩擦副的机械构件在正常的运行条件下，其磨损过程一般分为三个阶段：跑合阶段，稳定磨损阶段，剧烈磨损阶段。

① 跑合阶段（又称磨合阶段）　新的摩擦副表面由于机械加工造成了粗糙度，真实接触面积较少，相对运动过程会使表面磨平而变得光滑，真实接触面积逐渐增大，磨损速度减缓进入稳定磨损阶段。新设备或车辆的正常运行之前都需要一个跑合阶段，为构件稳定磨损创造条件。选择合理的跑合规程，采用适当的摩擦副配对材料及加工工艺，使用含活性添加剂的润滑油等方法，都能缩短跑合期。跑合结束应重新换油。

② 稳定磨损阶段　稳定缓慢的磨损阶段。

③ 剧烈磨损阶段　磨损急剧增加，机械效率下降，精度降低，产生异常噪声及振动，摩擦副温度升高，最终使构件失效。

有时也会发生：①转入稳定磨损阶段后，长时间内磨损甚微，无明显的剧烈磨损阶段，构件寿命较长。②跑合阶段和稳定磨损阶段无明显磨损，当表层达到疲劳极限后，产生剧烈磨损。③摩擦条件恶劣，跑合阶段后，立即转入剧烈磨损阶段，机械无法常运转。

人们通常按照摩擦表面损坏程度，把黏着磨损分为五类，如表 5.2-13 所示。

3）影响黏着磨损的因素

① 材料特性的影响

a）脆性材料比塑性材料的抗黏着能力高，塑性材料黏着破坏常常发生在离表面一定的深度部位，磨损下来的颗粒较大，而脆性材料的黏着磨损产物多数呈金属磨屑碎片状，破坏深度较浅。

b）互溶性大的材料所组成的摩擦副（包括相同金属或相同晶格类型的金属，以及相近的晶格间距、电子密度、电化学性能的金属），其黏着倾向大。互溶性小的材料所组成的摩擦副（异种金属或晶格结构不相近的金属），其黏着倾向小。

c）从金相结构考虑，多相金属比单相金属的黏着倾向小。金属中化合物比单相固溶体黏着倾向小。金属与非金属材料（如石墨、塑料等）组成的摩擦副比金属间组成的摩擦副的黏着倾向小。

d）周期表中能与铁不相溶或能形成化合物的 B 族元素，由它们组成的摩擦副，其黏着倾向小，而铁与 A 族元素组成的摩擦副，其黏着倾向大。

e）采用表面处理工艺（如电镀、表面化学处理、表面合金化沉淀、表面热处理、喷镀或堆焊等工艺）使摩擦副表面生成互溶性小、多相带有化合物组织，或采用非金属涂层来避免同种金属相互摩擦，均可防止黏着磨损发生。

② 压力的影响　一般当压力增大到某一临界值后黏着磨损量便急骤增加。当载荷超过材料硬度的 1/3 时，磨损量增加，严重时发生咬死现象。因此设计中选择许用压力必须低于材料硬度的 1/3，才有可能减轻或不发生黏着磨损。

③ 表面粗糙度的影响　一般说来，摩擦副表面粗糙度愈低（即愈光滑），则抗黏着磨损能力愈大；另一方面，过高地提高摩擦副表面粗糙度，因润滑剂不能储存于摩擦表面上，又会促进黏着磨损，在生产中，多是采用润滑方法来避免黏着磨损，所以表面状态对黏着磨损有十分密切的关系。如能使表面润滑剂很好地储存，便可大大降低摩擦力。比如机床导轨刮研花纹的排列形状对油的储存量及储存的时间有重要影响，它可使其摩擦力发生 5～10 倍的变化，这对减少摩擦功的消耗十分有利。另外工程上新机器采用逐渐加载跑合的方法，也是减少早期产生黏着磨损的有效措施之一。

④ 温度的影响　在摩擦过程中所产生的热量，可使局部微区的温升达到润滑油的“闪点”的温度。当摩擦表面温度升高到一定程度后，轻者破坏油膜，重者使材料因回火而软化，甚至发生熔化。有人认为，摩擦所消耗的功绝大部分以热的形式表现出来。通常，用热电偶测定出表面温度，是

真实接触面积上产生的温升向周围表面传播的结果，应视为平均温度，它比结合点的真实温度低很多。

⑤ 润滑油、脂的影响 润滑状态对黏着磨损的影响比较大。边界润滑状态下，黏着磨损值大于流体动压润滑，而流体动压润滑又大于流体静压润滑。有关润滑方面的理论请参见其他有关专著。

(4) 疲劳磨损

1) 表面疲劳磨损的起因 由前面的叙述可知，磨料磨损是由于磨料颗粒与构件表面间的相互作用而造成的金属流失，黏着磨损是摩擦表面间因直接接触而发生的金属流失。如果有润滑油将两表面隔开，同时又消除磨料颗粒的作用，则上述两类磨损可大为减少，但仍可能发生疲劳磨损。这是由于摩擦表面通过润滑油膜传递而承受着很大的应力。有油膜可使接触应力均匀，并起缓冲作用，可以减轻疲劳磨损的发生。产生疲劳磨损的构件的表面特征是深浅不同，大小不一的痘斑状凹坑，或较大面积的表层剥落。

2) 疲劳磨损的表示方法 滚动接触疲劳磨损的特征是在达到临界转数后开始滚动及滑动形成大块的磨屑，在此临界转数之前产生的磨屑可以忽略不计。因而滚动轴承采用正常工作到出现磨屑分离的时间来表示轴承的有效寿命。所以用磨损失重来表示疲劳磨损就没有意义了。

大批试验表明：轴承寿命 n 与外加载荷 P 的三次方成反比。即

$$nP^3 = 常数 \tag{5.3-22}$$

有些疲劳磨损也采用失重法测定的数据作为参考。但是由于些数值与构件本身质量相比非常微小，常常在误差范围内，所以绝大多数易产生疲劳磨损的构件，都是以寿命表示其耐磨性。

3) 影响疲劳磨损的因素及抗磨措施 从疲劳磨损的机理可知，疲劳磨损与裂纹的形成及其扩展有关。因此，凡是能够阻止裂纹形成及其扩展的方法都能减少疲劳磨损。

① 材质的影响

a) 钢的冶炼质量对构件抗疲劳磨损的能力有极为显著的影响。钢中的非金属夹杂物，特别是脆性的带有棱角的氧化物、硅酸盐及其他各种复杂成分的点状、球状夹杂物，它们破坏了基体的连续性，对疲劳磨损有严重的影响。显微镜的观察表明，裂纹几乎都以尖角形状夹杂物或化合物为源，特别是在交变应力作用下，由于尖角处的应力集中，使材料很快进入塑性变形，材料在加工硬化过程中，产生裂纹，导致早期破坏，缩短了轴承寿命。

b) 材料的组织状态也是重要的影响因素。轴承钢中未溶碳化物含量应该控制在 6.5%以下，否则容易形成粗大晶粒及带状组织等缺陷，造成钢中基体碳含量不均匀，使钢中马氏体基体在强度上产生差异。加之碳化物带状组织中往往附有非金属夹杂物，这些都会降低疲劳磨损抗力。对于未溶碳化物，要通过适当的热处理，使其趋于小、匀、少、圆为好。在未溶碳化物状态相同的条件下，马氏体含碳量控制在 0.5%左右为好。钢中固溶体的含碳量过多，使马氏体粗大，脆性增加，残余奥氏体含量相应增多，从而使疲劳寿命降低。若固溶体含碳量过低，会降低钢的基体强度与硬度，从而减弱基体的疲劳抗力。

② 渗碳钢的渗碳层对疲劳磨损的影响 渗碳钢表面脱碳或贫碳，使奥氏体稳定性差，寿命降低。外加载荷一定的条件下，心部强度越高，产生疲劳裂纹危险性越小。合理地提高渗碳钢心部硬度，能有效地提高表面的抗疲劳磨损性能。但是不能无限制地提高心部硬度，这会降低韧性而导致脆断。

另外，厚些的渗碳层可以提高抗疲劳磨损能力，这样可以使疲劳裂纹产生在硬化层内。渗碳层过浅，疲劳裂纹将产生在硬化层与心部的过渡区，容易使表层压碎剥落，但多厚的渗碳层比较合适，还要根据不同的使用要求来定。

③ 表面硬度的影响 轴承钢的硬度为 62HRC 时，抗疲劳磨损能力最大，硬度过高或过低，寿命均会明显下降。对齿轮来说，齿面硬度在 58 ~ 62HRC 的范围内为最佳。对于承受冲击的齿，其硬度取下限。一般要求小齿轮硬度大于大齿轮，跑合之后，使接触应力分布均匀。齿轮的硬度匹配很重要，直接影响接触疲劳寿命。

④ 表面粗糙度的影响 降低表面粗糙度，可以有效地提高抗疲劳磨损的能力。对于滚动轴承，表面粗糙度从 $R_a = 0.32 \sim 0.63$ 变到 $R_a = 0.16 \sim 0.32$，则寿命可提高 2 ~ 3 倍；由 $R_a = 0.16 \sim 0.32$ 变到 $R_a = 0.08 \sim 0.16$，寿命可提高 1 倍，由 $R_a = 0.08 \sim 0.16$ 变到更高级，寿命可提高 0.4 倍，粗糙度的值更小时对寿命的影响很小。接触应力大小不同，对表面粗糙度要求也不同。一般接触应力大，要求较小的粗糙度。表面硬度愈高的轴承和齿轮，要求粗糙度也愈小。

降低表面粗糙度的加工工艺也很重要。如呆表面在抛光、磨光过程中产生裂纹则很容易产生剥落坑。

⑤ 残余应力的影响 一般说来，当表面层在一定深度范围内存在有利的残余压应力时，不仅可提高弯曲、扭转疲劳抗力，而且能提高接触疲劳抗力，减少疲劳磨损。表面渗碳、淬火、表面喷丸、滚压处理等，都可以使表面产生压应力。

残余压应力对应变疲劳磨损效果不大。

⑥ 其他因素的影响 对装配应有严格的要求，应保证装配精度。每对齿轮的装配，要防止和减轻齿面的对角接触，保证接触印痕总长不少于齿宽的 60%，并且接触印痕处在节圆附近。这就有可能避免出现早期剥落坑。

润滑油的黏度愈高，接触部分的压力愈接近平均分布，对防止疲劳磨损愈有利。油的黏度低，容易使油渗入裂纹中，加速裂纹的扩展，使疲劳寿命降低，润滑油中含水量过多，也促使点蚀的发生，必须严格控制润滑油中的含水量。

(5) 微振磨损

微振磨损指两个配合表面之间叫小振幅滑动所引起的微小振幅滑动所引起的一种磨损形式。

微振和往复滑动的重要区别仅仅在于每次往复的距离不同，微振也是一种往复运动，但振幅很小，一般处于几微米至上百微米之间。振幅的范围目前还没有统一的认识。微振磨损的“临界振幅”是指不能造成材料磨损的位移下限。在临界振幅之下称为微观滑动，微观滑动只在接触面上造成塑性变形，但不产生磨屑。

1) 特点 和往复滑动相比，微振磨损有下列特点。

① 由于振幅小，滑动的相对速度低。例如一个振幅为 20 μm，频率为 50 Hz 的微振幅，平均速度仅为 2 mm/s。

② 由于振幅小，微振表面的绝大部分总是保持接触，因此磨屑逸出的机会很少。

③ 磨损产物如钢上磨损产物是红棕色粉末，而铝或铝合金为黑色粉末。

2) 微振磨损发生过程 关于微振磨损的实验研究和机理的探讨，已提出了各种理论模型，但尚无一个公认的、能解释所有微振磨损现象的理论。尽管如此，大家都认为微振磨损是一个复杂的过程，包含黏着、氧化、磨粒和疲劳等的综合作用，只不过对上述四个过程的顺序和相对重要性存在分歧，争论的焦点集中在氧化及氧化物在微振磨损中的作用问题。

3) 主要影响因素

① 微振振幅的影响 实验证明，存在着临界振幅。不同的材料和不同的试验条件，临界振幅的大小也不同。

② 循环次数的影响 最初的跑合阶段磨损速度高，经

过一定循环数后成为线性关系的稳定磨损阶段。

③ 环境温度的影响　一般认为较高的环境温度有助于减轻微振磨损。但环境温度对不同材料的影响是不一样的，有的材料加速磨损，有的材料减弱磨损。

④ 环境气氛的影响　室温下在惰性氩气中和在空气中试验，微振磨损量接近。也有人在惰性气体和真空中进行试验，和一般滑动磨损试验结果不同，其磨损速率比在氧化性气氛中要低得多。这证明氧在室温条件下的不利影响。微观分析证明，在非氧条件下，加工硬化加强了接触表面，阻止磨损颗粒的形成。

⑤ 垂直载荷和频率　如果滑动振幅不变，平面钢样微振副之间的磨损量随垂直载荷的大小成线性增加。但很高的载荷可降低滑动振幅，甚至完全阻止微振。防止微振磨损的措施之一，就是增大构件间的垂直接触载荷，从而阻止微振磨损的发生。

频率对微振磨损的影响，不如其他参量明显。在固定循环数的条件下，频率愈低则磨损量愈大。在非氧化条件下，则频率对磨损量没有影响。

从上述各参量对微振磨损的影响可以看出，还存在着许多矛盾的现象和结论，说明对微振磨损的研究仍比较欠缺。

4）判断是否因微振磨损造成构件失效

① 了解构件的工况条件，是否存在能引起紧密配合的表面间微小滑动的振动源或交变应力。

② 分析接触表面的形貌，观察匹配表面上是否出现麻点坑或小划痕，以及残留的磨屑。

③ 分析与接触表面垂直剖面的亚表层特征，观察是否出现微小裂纹，这些裂纹是否平行表面或与表面成一定角度发生扩展。

④ 磨屑是判断发生微振磨损的重要依据。钢的磨屑要比普通铁锈红得多，极易团聚，而铝的磨屑则是黑色的（氧化铝通常是白色的）。

1.5　非金属构件的失效预测

由金属材料及其构件失效造成的灾难性事故已引起人们的高度重视，而对非金属材料及构件失效的研究开展的较少。然而，随着非金属构件的广泛应用，由此而产生的失效与日俱增。非金属制件因品种多，材料的性质各异，用途广泛，有的材料在常温下属于脆性材料（如有机玻璃），有的是高弹性材料（如橡胶），有的比强度特别高（如复合材料）等等，情况复杂，同时，非金属材料由于其本身的物理性能和化学性能与金属材料有很大差别，如橡胶不透明、不反光、不导电；而有机玻璃透明、不导电、反光性能差，对温度、有机溶剂敏感等，因此非金属材料及构件失效的形式、种类以及分析技术等方面与金属材料有所不同。本节仅简要介绍了非金属构件失效的基本特点。

1.5.1　非金属材料的基本特性

非金属材料由高分子链组成，因此也称为高分子材料。与金属材料相比，非金属材料除具有一般不导电、反光性差、对有机溶剂敏感和易于老化等特点外，还具有结构上的不均匀性及表现出其他材料不可能出现的高弹性及突出的黏弹性。

（1）非金属材料与制件在结构上的不均匀性

非金属构件所用的高分子材料，有些是化学均匀的单相材料，有些是非均质的多相材料。就构件而言，有些是由一种材料制成的，有些是由多种性能差异比较大的材料（包括非高分子材料）宏观复合而成的复合材料。因此，航空非金属材料本身及构件上存在着种种不均匀性。就化学均匀的单相材料如有机玻璃来说，材料的不均匀性包括以下几个方面：分子量大小不均匀；分子链间的物理交联点不同，分子链末端的结构往往与分子链的重复结构单元不同，而且其周围的自由体积也较大；分子链段在材料表面的堆砌密度低于材料内部。

对非均质多相材料如纤维增强复合材料、金属与非金属层板复合材料等，其不均匀性表现为：各相的物理化学性质差别很大，分散相的形状和尺寸不均匀，它们在基体中的分布和取向不均匀。

对于由两种或两种以上性质不同的材料组合而成的复合材料，材料的不均匀性则表现得更为突出。层板之间的差异以及开孔切割断部分纤维，从而使开孔处的应力集中现象比金属材料要严重。

在构件的制备过程中，由于受热和力学等因素的影响，构件内结构的不均匀性可能会进一步增加。

（2）非金属的黏弹性

理想的弹性材料的应力－应变关系服从胡克定律：

$$\sigma = E\varepsilon \qquad (5.3\text{-}23)$$

式中，σ 为应力；ε 为应变；比例系数 E 为材料的弹性模量。理想的黏性材料的应力－应变关系服从牛顿定律：

$$\sigma = \eta \mathrm{d}\varepsilon/\mathrm{d}t \qquad (5.3\text{-}24)$$

式中，$\mathrm{d}\varepsilon/\mathrm{d}t$ 为应变速率；比例系数 η 为材料的黏度。大多数高分子材料的力学行为则既不服从虎克定律又不服从牛顿定律，而是弹性与黏性的线性或非线性组合。具体地说，黏弹性材料的应变明显落后于应力，因此使材料产生一定的应变所需的应力，与应变速率的关系极大。

高分子材料具有突出的黏弹性的根本原因是长而柔软的高分子链具有多重热运动组元：键长键角、侧基、链节、链段和整个高分子链。其中除键长键角的变化不需要克服分子间的内摩擦力外，其他各种单元的运动均或多或少地需要克服分子间的内摩擦力。克服分子间的内摩擦力就是黏性的本质。

高分子材料黏弹性的表现形式很多。蠕变、应力松弛以及应力应变行为强烈地依赖于应变速率和温度也是黏弹性的表现。正是由于高分子材料具有突出的黏弹性，其破坏机理也必然与温度及应变速率密切相关。

1.5.2　非金属构件失效的基本类型

非金属材料及构件种类繁多，使用环境各异，因而失效形式也复杂多样，但根据所受的载荷和环境条件，非金属构件失效的主要类型可分为两种。一类是因制件材料本身在加工、储存和使用过程中受各种环境因素的作用而性能逐渐下降导致最终失效，这一类称之为老化。另一类是制件在使用中长期受机械力和环境的作用而丧失规定的功能，这一类习惯上称为失效。

根据非金属材料或构件所受的载荷条件或环境条件，非金属材料及构件发生机械失效的模式又可分为以下五类。

1）直接加载下的断裂—材料或构件在拉伸、压缩、剪切或冲击等载荷作用下发生变形直至发生断裂。材料断裂时对应的应力叫做断裂强度。

2）疲劳断裂—材料或构件在远低于其断裂强度的交变应力作用下发生损伤直至断裂的过程。

3）蠕变断裂—材料或构件在远低于其断裂强度的恒定应力作用下发生变形直至断裂的过程。

4）环境应力开裂—材料或构件在腐蚀性环境（包括溶剂）和应力的共同作用下发生开裂。

5）磨损磨耗—两种不同材料在摩擦过程中，其表面材料以小颗粒的形式断裂下来。

上面述及的是非金属材料制件基本的老化失效和各种机械失效方式，而实际上这些失效形式往往不是单一出现的，而常常是两种或两种以上的失效模式交织出现，呈错综复杂的模式。

1.5.3　直接加载下的瞬时断裂

作为构件，决不希望它们在使用中所受直接载荷使它们发生断裂破坏。因为构件承受的设计应力水平远远低于材料在使用温度和应变速率范围内的断裂强度或屈服强度。但是，研究非金属构件在直接加载条件下典型的断裂形貌及其特征，有助于分析非金属构件在实际工作状态下发生的破坏。

非金属构件在直接加载下的断裂分为脆性断裂和韧性断裂两大类，为方便起见，我们仅以拉伸断裂为例。

(1) 脆性断裂

聚苯乙烯塑料、非定向有机玻璃、热固性塑料等刚性高分子材料在室温下的断裂都属脆性断裂。其基本特征为断裂前应力与应变之间呈良好的线性关系，即基本服从理想弹性体的胡克定律；断裂应变小（$<5\%$），断裂能低，即应力应变曲线与应变坐标轴之间包围的面积小，断裂后，试样几乎无残余应变；断口表面与拉伸方向基本垂直。

脆性断裂过程基本上可分为三个阶段，即断裂源首先在材料的最薄弱处形成，一般是主裂纹通过单个银纹扩展。随着裂纹扩展和应力水平的提高，主裂纹不再是通过单个银纹扩展，而是通过多个银纹扩展，因而断面转人雾状区即裂纹扩展的第二阶段当裂纹扩展到临界长度时，断裂突然发生，即进入裂纹快速扩展的第三阶段。这三个阶段在断口上留下不同的形貌，粗略地说，第一阶段在断口上留下光滑的镜面区，第二阶段的断口称为雾状区或平坦区，而第三阶段在断口上留下粗糙区。

(2) 韧性断裂

各种橡胶材料和某些工程塑料如尼龙和双轴拉伸定向有机玻璃等在室温下的拉伸断裂属韧性断裂。韧性断裂的主要特征是断裂应变很大（达百分之几十至百分之几百），断裂能很高。但在橡胶的韧性断裂和塑料的韧性断裂之间又有形式和本质上的差别。

橡胶材料的拉伸应力-应变行为表现为软而韧的特点，即弹性模量很低（$10^{-1}\sim1$ MPa），弹性应变很大（最高达1 000%），卸载后试样几乎无残余应变。这类材料的断裂称为高弹性断裂。高弹性断裂过程虽然也分为裂纹的慢速扩展和快速扩展两个阶段，但与脆性断裂的断口形貌相反，高弹性断裂过程中裂纹的慢速扩展阶段在断口上留下的是粗糙区，而裂纹的快速扩展阶段在断口上留下的却是光滑区。这是由于橡胶的高弹形变本质上是卷曲的高分子链（或交联橡胶中的网链）通过链段运动沿拉伸方向的伸展取向。在橡胶拉伸时，高弹形变中分子链的伸展取向能使那些垂直于拉伸方向的缺陷的尖端钝化，从而在一定程度下一直其扩展。当橡胶的高弹应变足够大时，应力在取向分子链上分布的不均匀性越来越大。那些受拉伸应力特别大的分子链将首先被拉断或滑移，形成断裂源。在裂纹的慢速扩展阶段，取向的分子链被一个一个或一束一束地相继拉断。由于分子链或链束中的薄弱点是随机分布的，未必都位于与应力方向垂直的某个平面附近，因而断口上的形貌比较粗糙。而在裂纹的快速扩展阶段，则因为是许多分子链同时被拉断，因此断口上的形貌反而比较光滑。

与橡胶材料不同，一些工程塑料在拉伸断裂过程中则表现出硬而韧的特点。在这种形变断裂过程中，最重要的特征是初始弹性模量高、能出现屈服点并在屈服后能继续产生很大的形变，断裂应变大、断裂能高。与一些金属材料的应力-应变曲线相似。在初始阶段键被拉伸并且能量以完全可逆的形式储存起来。当卸去载荷后，形变恢复，符合胡克弹性（亦称线性黏弹性）。测试材料的弹性模量和泊松比等在此范围内进行。在应力进一步增加，特性曲线呈非线性，除去应力后，形变仍然恢复，但是很缓慢（非线性黏弹性）。随着应力增加，产生屈服，伴随不可逆形变的产生。在这种情况下，能量以形变黏滞的形式消耗。应力继续增加，通常正应力稍微有些下降，并且伸长很容易发展。在连续经受拉伸载荷下，逐渐形成缩颈区，断裂失效可能发生。在缩颈开始以后，分子取向，推迟断裂（发生应变硬化），并且实际应力增加。拉伸时，缩颈和非缩颈的截面积都基本保持不变；但缩颈段长度不断扩展，非缩颈段长度不断减小。当整个工作段全部转变为缩颈后被均匀拉伸至断裂，这一点与金属材料不同。如果试样在拉断前卸载，或试样因拉断而自动卸载，则拉伸中产生的大形变除少量可回复之外，大部分形变都将残留下来：这种拉伸形变过程称为冷拉。

塑料除了能出现上述宏观屈服和塑性形变之外，还可能在微小区域内出现局部屈服和塑性形变，这在高分子材料中称为银纹现象。银纹不同于裂纹：裂纹内部是空的，而银纹内部却只有约40%的体积是空穴。银纹的两个银纹面之间有银纹质。银纹质实际上是银纹面之间高度取向的分子链构成的微纤束。产生银纹的机理是材料在拉伸应力作用下局部薄弱区域发生屈服冷拉，但由于其周围的材料并未屈服，局部冷拉中所需要的横向收缩受到限制，因此在微束之间留下大量的空穴。

尽管银纹的形成是材料局部韧性的表现，但它又是高分子材料脆性断裂的先决条件。在高分子材料的脆性断裂过程中，裂纹是通过银纹扩展的。这是由于在一定的应力作用下，高分子材料首先在薄弱区产生局部的塑性形变，形成银纹。随着应力水平的提高，银纹尖端的应力集中区进一步通过局部塑性形变将本体高聚物转化为银纹质，从而使银纹向前扩展。同时，在应力作用下，银纹底部进一步张开，其中的微纤束进一步被拉伸。当应力水平增加到足以将银纹底部高度取向的微纤束拉断时，部分银纹便转化为裂纹，但裂纹前缘仍保留有一个银纹区。正是由于裂纹扩展中其前缘始终有一个或多个通过局部塑性形变而扩展的银纹，断裂过程中需吸收较多的能量，因此，即使高分子材料发生脆性断裂，其断裂韧性也比无机玻璃高得多。

(3) 韧-脆转变

以上讨论了塑料和橡胶在室温条件下的脆性与韧性断裂。然而，由于高分子材料具有突出的黏弹性，它们的应力-应变行为为严重地受温度和应变速率的影响。以有机玻璃在变温拉伸为例，随着温度的升高，有机玻璃的模量和强度下降，断裂伸长率增加。在4℃时，有机玻璃是典型的刚而脆的材料，而到66℃时，已变成典型的刚而韧的材料。韧-脆转变温度约在20～30℃。习惯上把这个转变温度称为脆化温度，以 T_b 表示。应变速率对高分子材料应力-应变行为的影响可概括为，应变速率的升高相当于温度的下降。

此外，一些高分子材料呈现很强的缺口敏感性，缺口之所以对有些高分子材料有脆化作用，根本的原因在于三向应力的作用使缺口尖端的屈服强度提高。

1.5.4　疲劳断裂与蠕变断裂

(1) 非金属构件的疲劳断裂

当受交变载荷作用时，高分子材料构件在应力作用下首先在最薄弱或应力集中较大的区域产生银纹。随后在每一循环应力的应力上升期间，银纹中的微纤束被进一步拉伸，银纹/本体界面由于本体材料的微纤化而变宽，同时随银纹前缘本体材料的微纤化而增长。在循环应力的应力下降期间，由于微纤的弹性回缩和屈曲，银纹闭合。在循环应力的反复作用下，银纹底部的微纤变得愈来愈细长。当在某个循环的应力上升期间，银纹中的应力足以使微纤断裂或使微纤/本体界面分离时，银纹便转化为裂纹。

疲劳裂纹一旦形成，便开始进入裂纹的扩展阶段。在裂纹的慢速扩展阶段，裂纹在每一循环应力的作用下扩展一个

微小的量。

各种高分子材料疲劳裂纹的扩展速率与它所受的应力强度因子范围之间的关系不像金属材料那样具有规律。但对于像有机玻璃、聚苯乙烯和聚氯乙烯等非晶态塑料，一般可用Paris公式表达。

在疲劳裂纹的初始阶段，裂纹是通过单一的一个银纹扩展的，但当应力强度因子范围超过临界值后，裂纹尖端有更多的能量用来引发银纹，因此在裂纹前缘便会出现由多个银纹组成的银纹束。此后，疲劳裂纹便不再是通过单个银纹，而是通过银纹束扩展了。

疲劳裂纹通过单个银纹扩展时，在断口表面留下光滑的镜面区，当疲劳裂纹通过银纹束扩展时，在断口表面留下粗糙区。与静载断口相似，疲劳断口也存在三个区域，只不过是疲劳断口形貌最大的特征是在镜面区中有许多以主裂纹源为中心的环状平行条带，条带间距反映出裂纹每次向前扩展的距离。不过在像炭黑增强橡胶之类的多相体系中，疲劳条带有时不易分辨。许多航空非金属件在实际使用中常受交变载荷或脉冲载荷的作用。这类制件，特别是当它们在加工或装配过程中已存在缺陷，比较容易出现疲劳损伤甚至疲劳断裂的现象。

（2）蠕变断裂

材料在恒定应力（应力水平低于材料的断裂强度）作用下应变随时间逐渐增加，最后发生宏观断裂，这种现象称做蠕变断裂，也可称做静态疲劳。蠕变是高分子材料黏弹性的典型表现之一，也是该类材料制件机械失效的重要模式之一。

与金属材料类似，高分子材料从蠕变开始（即从受到恒定应力作用的时刻开始）直到断裂所需的时间一般符合下式所示的规律：

$$t_f = A\mathrm{e}^{-B\sigma}$$

式中，t_f 为材料从蠕变开始直至断裂的时间；σ 为应力，A 和 B 在一定应力水平范围内是常数。

研究表明，高分子材料的蠕变断裂有以下几个特征：①材料在高应力水平下发生的蠕变断裂为韧性断裂，断裂应变大；而在低应力水平相反。在一定的应力水平范围内，蠕变断裂发生韧-脆转变。②在韧性蠕变断裂过程中，材料会出现“发白”现象。“发白”的原因是材料内部出现了许多空穴。在像聚苯乙烯和有机玻璃之类的脆性蠕变断裂过程中，材料内部必定产生许多应力银纹。应力水平愈高，材料内的银纹密度愈高。银纹中绝大多数不会发展为宏观裂纹，因为相邻银纹尖端的应力场相互干扰，使银纹的扩展受到抑制。只有少数银纹能通过扩展和彼此合并，形成贯穿试样或构件的宏观裂纹。可见，制件在长期使用中出现“发白”或应力银纹是蠕变断裂的先兆。

1.5.5 环境应力开裂

环境应力开裂是非金属构件在使用中因特殊介质（腐蚀性介质、溶剂和某种气氛）和应力的共同作用产生许多小裂纹，甚至发生断裂的现象。这类失效模式的基本特点如下。①裂纹始于制件表面，裂纹长度方向与拉伸应力方向垂直。②使材料产生环境应力开裂的应力水平比该种材料的断裂强度低得多，产生环境应力开裂所需要的最低应力值，称为临界应力。当材料所受的应力水平低于这个临界值时，不可能发生环境应力开裂。使材料或构件产生环境应力开裂的应力可以是外力，也可以是内部存在的残余应力。③材料在环境应力开裂中产生的许多小裂纹，大多因邻近裂纹的抑制作用而不易扩展，只有少数裂纹互相贯穿导致材料的断裂。④对于一定的材料，通常只对某些环境介质敏感。

表征材料抗环境应力开裂能力的指标是该材料在单轴拉伸和接触某种介质的条件下直至断裂所需的时间。它与介质的性质、应力水平、温度和材料本身的结构因素等有关。引起高分子材料或构件发生环境应力开裂的介质包括有机溶剂、水、某些表面活性剂和臭氧等。有机溶剂容易促进塑料，特别是非晶态塑料的环境应力开裂，例如有机玻璃在苯、丙酮、乙酸乙酯和石油醚中，聚碳酸酯在四氯化碳中。水和表面活性剂容易引起聚乙烯发生环境应力开裂。臭氧容易使不饱和碳链高聚物，尤其是不饱和碳链橡胶发生环境应力开裂，例如天然橡胶只要在微量臭氧和5%的应变条件下就能开裂。

各种高分子材料抗环境应力开裂的能力是不同的。例如，线形聚乙烯在水和表面活性剂环境中很容易开裂，但辐照交联聚乙烯的抗开裂能力大大提高，聚丙烯塑料则几乎不出现环境应力开裂现象。就有机玻璃和聚碳酸酯塑料而言，材料中低分子量成分愈多，愈容易发生环境应力开裂。例如，有机玻璃老化后出现溶剂-应力银纹的临界应力比新玻璃的要低得多。老化严重时，临界应力值甚至可降到零。国产聚碳酸酯塑料之所以容易开裂，就是因为分子量分布宽，且包含较多的低分子量成分的缘故。

高分子材料的环境应力开裂既不同于高分子材料的老化失效，又与之密切相关。老化会降低材料的抗环境应力开裂能力，环境-应力的共同作用又会促进材料的老化，因此在分析制件失效原因时，不能把两者完全割裂开来。

1.5.6 磨损磨耗

有些高分子材料制件，例如橡胶轮胎和塑料传动零件（齿轮、齿条、轴承等）是在摩擦条件下使用的。制件受摩擦时，表面的材料以小颗粒的形式断裂下来，称为磨损磨耗。很难说磨损磨耗的机理纯粹是材料的断裂过程，因为制件在摩擦中产生的热使材料升温。温度过高时，会引起制件材料的局部熔化、降解和氧化反应等。不过，制件在摩擦中表面材料以碎屑形式掉落下来毕竟意味着断裂是磨损磨耗的主要机理。

在像喷砂或用泵汲送悬浊液等冲击摩擦条件下，高分子材料的耐磨性远远超过金属材料，但在滑动摩擦条件下，除一些摩擦系数很小的自润滑塑料如聚四氟乙烯、尼龙、聚甲醛等之外，高分子材料的耐磨性一般不如金属材料。因为高分子材料是热的不良导体，摩擦中产生的热量容易积聚起来，从而引起材料发生除机械断裂之外的其他破坏过程，如熔化、降解甚至热爆破等。

在目前所用的各种航空非金属件中，橡胶轮胎是在摩擦条件下使用的最重要的制件，此外，一些输油软管的内壁也不断受到液体介质的摩擦和冲刷作用。

橡胶制件在粗糙表面上滑动时，在磨损磨耗中材料的断裂主要是拉伸断裂，而在光滑表面滑动时，则主要是疲劳断裂。橡胶制件在磨损磨耗中表面常常形成粗糙不平的花样，其结果使橡胶的摩擦因数提高，而摩擦因数的提高又会进一步加速其磨损磨耗的速率，除非经常更换摩擦方向才能有效地降低橡胶的磨损磨耗速率。对于橡胶轮胎或输油软管来说，要经常更换摩擦方向似乎是做不到的。另外，橡胶的耐磨性与温度有关，温度越高，耐磨性越差。

1.5.7 高分子材料及构件的老化

与金属材料不同，高分子材料本身在加工、储存和使用过程中由于对一些环境因素较为敏感而导致性能逐渐下降，即发生高分子材料或构件的老化。

引起高分子材料老化的环境因素有物理因素（包括热、光、高能辐射和机械应力的作用）、化学因素（如氧、臭氧、水和酸、碱、油等的作用）和生物因素（如微生物和昆虫的作用）。在这些环境因素作用下，高分子材料性能下降，例如，有机玻璃发黄、发雾、出现银纹甚至龟裂。汽车轮胎和橡胶软管出现龟裂、变硬、变脆；油漆涂层失去光泽甚至粉

化、龟裂、起泡和剥落；玻璃钢制品起毛、变色、强度下降等。高分子材料在老化过程中性能下降的主要原因是分子链发生降解和交联反应。降解反应导致分子链断裂，即分子量下降，从而使材料变软、发黏甚至丧失机械强度；交联则往往使高分子材料变脆或失去弹性。

各种高分子材料老化的难易程度与高分子链的结构直接相关。一般来说，杂链高分子容易受化学的侵蚀，而碳链高分子往往对化学试剂比较稳定，但容易在物理因素和氧的作用下老化。

(1) 老化的基本类型

按引起老化的外界条件分类，老化主要可归结为以下五类：由热、光、高能辐射、氧化及生物降解引起的老化。

1) 热老化　高分子材料在热的作用下发生的老化称为热老化。高分子材料是否发生老化，除了热的作用，还取决于材料的热稳定性，而热稳定性与高分子链上化学键的键能有关。化学键能愈高，热稳定性愈好。高分子材料的热稳定性通常用半分解温度 $T_{1/2}$ 来表征，即为高分子材料在真空中加热 30 min 后损失一半质量所需要的温度。高分子材料的热稳定性和高分子链结构之间的关系有下述基本规律：①高分子链中靠近叔碳原子和季碳原子的键较易断裂；②高分子主链中有—O—O—，—N══N—键时，热稳定性低；③高分子链中存在氯原子之类的取代基将形成弱键而降低热稳定性；④C—H 键中的氢完全被氟原子所取代而形成 C—F 键时，可大大提高热稳定性；⑤当高分子主链中含有较大比例的环状结构时，一般热稳定性都较高；⑥主链由—Si—O—键组成的有机硅塑料和硅橡胶具有很高的热稳定性。

大多数航空非金属件都是在远低于其热分解温度的环境中使用的，因此纯粹由于热引起的老化速率并不高，但某些安装在发动机附近的部件和一些在工作条件下容易发生摩擦而生热的部件（如轮胎），则因工作温度较高，热老化不容忽视。

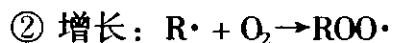
2) 光老化　高分子材料在光的作用下发生的老化叫光老化。光是一种电磁波，当阳光通过大气层到达地面时，波长范围为 $3\times(10^2\sim10^4)$ nm 叫不同波长的光具有不同的能量，其中波长为 300~400 nm 的近紫外光的能量为 300~400 kJ/mol，一般共价键断裂所需的能量为 160~420 kJ/mol，因此太阳光中的近紫外光可能引起共价键为主的高分子物质的化学键断裂。不过，光要在高分子材料中引发反应，首先必须被高分子物质吸收。一些只含单键的高分子物质，如饱和聚烯烃及其衍生物，一般不吸收波长大于 300 nm 的光。按理，地球表面的紫外光不应引起这类物质的光老化，但由于这类高分子材料在合成、加工和储存中往往与氧发生反应形成羟基、过氧化氢基或双键，加上某些添加剂和催化剂残留物也可能吸收紫外光，因此实际上这类高分子材料也会发生光老化。高分子材料在光老化过程中，既可能发生降解反应，也能发生交联反应。

3) 高能辐射老化　α、β、γ、X 射线，快中子、慢中子和离子辐照等均为高能辐照，高能辐照引起的高分子物质的化学变化，有的以辐照降解为主，有的以辐照交联为主。

4) 氧化老化　高分子材料与空气中的氧和臭氧发生反应而引起高分子的降解或交联称为氧化老化。高分子物质的氧化反应在室温和避光条件下进行得十分缓慢，但在受到热或光的照射时，氧化老化速率则大大提高，从而导致材料迅速老化。高分子材料在热和氧的共同作用下发生的老化称作热氧老化，同样在光和氧的共同作用下发生的老化称做光氧老化。光氧老化是引起高分子材料大气老化的主要原因，某些在高温条件下使用的材料则容易发生热氧老化。热氧老化和光氧老化是自动加速过程，具有自由基型链锁反应的特点，一般可以用下列方程描述热氧老化和光氧老化的反应过程：

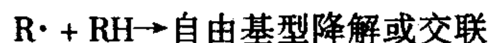
光或热


① 引发：RH→R· + H·

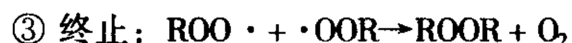

② 增长：R· + O_2→ROO·


ROO· + HR→ROOH + R·


ROOH→RO· + ·OH


R· + RH→自由基型降解或交联


HO·


③ 终止：ROO · + ·OOR→ROOR + O_2


R· + ·R→R—R


ROO · + ·R→ROOR

此外，某些微量过渡金属（如 Fe，Cu，Mn 或 Ni）的离子会加速氧化反应。高分子材料与氧反应时有一个吸氧诱导期，诱导期一过就转入自动加速氧化阶段。高分子材料是否容易发生氧化，首先取决于它是否容易吸氧。一般地说，氧化老化是碳链高聚物的特征，杂链高聚物一般对氧比较稳定。在碳链高聚物中，不饱和双烯类高聚物如天然橡胶、顺丁橡胶等比较容易氧化老化，因为它们主链上的双键和 α 碳原子容易吸氧。这类橡胶氧化时往往同时发生氧化降解和氧化交联。氧化降解的结果使其分子量下降，从而使之发黏并丧失原有的强度；氧化交联的结果又使其分子量和交联度增加，从而失去弹性、变脆并出现龟裂。比较起来，饱和碳链高聚物的氧化速率低得多，例如聚苯乙烯在 100℃时长期暴露在空气中变化很小，但在紫外光存在下，于 60℃时就会氧化，形成羰基和羟基，表面发黄。除化学结构之外，高分子的聚集态结构对高分子材料的氧化速率有影响。例如，对于像聚乙烯、聚丙烯之类的结晶高聚物来说，结晶程度愈高，则氧化速率愈低，因为氧在分子链密集堆砌的晶相中的扩散速度低于在分子链无规排列的非晶相中的扩散速率。

最后要指出的是，臭氧对高聚物的氧化能力比氧更强，因为臭氧极不稳定，易分解出原子氧。微量的臭氧能直接氧化天然橡胶，使之产生裂缝。这类臭氧引发的裂缝常常是橡胶机械失效的起源。

5) 生物降解　在自然界中，微生物为生存和繁衍后代而需要能源，而能源正是通过对高分子的分解、氧化和消化得到的。一般地说，微生物对聚烯烃的作用甚微，但天然纤维、木材、丝、毛和天然橡胶等却是微生物的传统食粮。微生物主要通过破坏高分子主链及消化高分子材料中的增塑剂和其他添加剂而使高分子材料老化的。

(2) 水在老化中的作用

水对高分子材料的老化作用包括化学作用和物理作用两方面，化学作用通常是指水引起高分子的水解，例如尼龙中的酰胺基也可被水解为羧基和胺基；聚酯中的酯基可被水解为羧基和羟基；如果容易水解的基因分布在高分子链的主链上，则由于水解引起高分子链的断裂，对材料的性能影响很大。反之，如果容易水解的基团位于高分子链的侧链（基）中，则由于水解对高分子的平均分子量影响不大，所以对材料性能的影响比较小。水对高分子材料的物理作用包括以下几个方面：①溶胀增速作用，使材料刚度和强度下降。②脆化作用，使材料刚度和脆性提高。研究表明，当有些极性高分子材料的含水量超过某个临界值后，水分子在材料中的分布是不均匀的。在某些高分子链间的空穴中或极性基团周围密集的缔结水分子，可以像刚性填料一样使高分子材料的刚度和脆性增加。③在材料表面引入张应力，当高分子材料吸水后干燥时，失水过程中表面处于受张应力作用的状态。

以上对引起高分子材料老化的各种因素分别作了讨论。

对于一个具体的高分子材料制件来说，其老化失效不一定是某个单一因素造成的，而往往是多个因素，如热与氧、光与氧、湿与热或热-光-氧-湿等共同作用的结果。具体作用因素取决于制件的工作环境。高分子材料老化的结果，主要不是在制件中形成局部缺陷，而是引起整个制件材料的性能下降，以致丧失使用价值。一般地说，材料因老化而引起的性能下降速率比较缓慢，而且在很多情况下都是可观察得到的。因此，只要掌握构件老化的规律，及时更换老化失效制件，一般不至于造成飞行中的灾难性事故。目前尚没有一项灾难性事故可仅仅直接归咎于构件材料的老化。然而，与金属和陶瓷材料相比，高分子材料的老化又是一个突出的问题。航空非金属件的老化现象是普遍存在的，座舱盖有机玻璃的强度下降，防弹玻璃的脱胶，低温滑石封严层的掉块，各种橡胶制品的变硬、变脆等都是由于制件老化引起的部分或完全失效。

2　以工程力学为基础的失效预测技术和方法

机械零件采用的强度判据需要根据它的失效模式来确定，即失效分析是机械强度设计的根据。下面分别介绍基于经典力学、断裂力学、疲劳力学和损伤力学的失效分析预测技术和方法。

2.1　以经典力学为基础的失效预测预防技术和方法

2.1.1　经典力学基础

从广义上来说，零件或构件失效的原因有设计、材料、制造和使用（包括环境）等多方面的问题，其中设计则是首要问题。一个零件或构件的设计主要是指构件功能设计、机械强度设计（包括材料的选用），外观设计等，而其关键则是机械强度设计。在常规的零件强度设计中，首先要分析构件中所承受的载荷和它所产生的应力，其次是根据失效类型确定其许用应力，最后要通过计算使得构件的实际工作应力不超过许用应力。这就是保证所设计的机械可以安全运行的判据，并据此设计计算出零件的合理尺寸与结构。

2.1.2　受简单应力零件的失效分析技术和方法

这里首先介绍零件承受简单应力的失效和抗力指标，其中包括受拉压、受剪、扭以及受弯、弯冲情况下的失效和失效抗力指标。

(1) 受拉、压零件的失效和失效的抗力指标

在工程实际当中，承受简单的拉压载荷的杆件是很常见的，这些杆件共同的受力特点是，作用在杆端外力的合力作用线与杆件的轴线重合，其变形特点是沿轴线方向伸长或缩短，故可称为轴向拉伸或轴向压缩。

通常受轴向拉伸杆件的变形和断裂可以用单轴拉伸的试样模拟。其变形和断裂过程用拉伸图或应力-应变图来描述，如图5.3-3拉伸图。

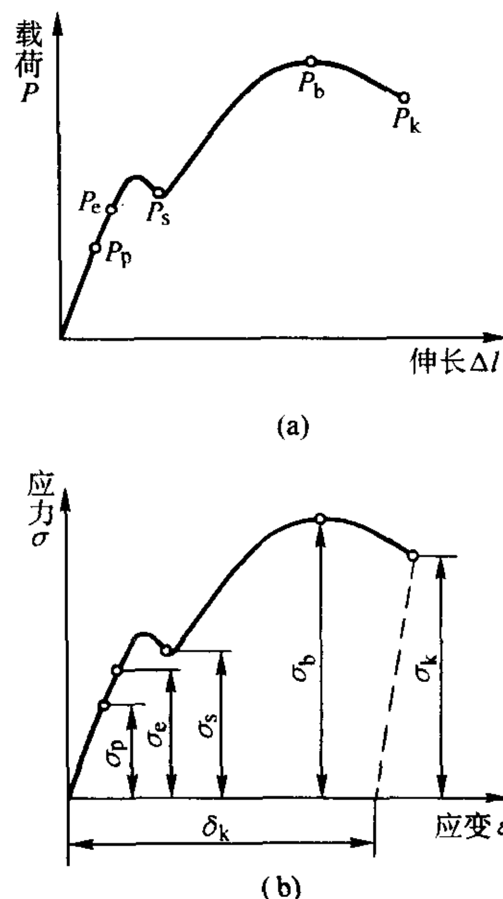


图5.3-3　拉伸图

图5.3-3a为退火低碳钢的拉伸图，应力-应变曲线的形状与拉伸图相似，如图5.3-3b所示。在应力应变曲线上可以直接读出材料的力学性能。当外力达到材料的某种力学性能时，即发生拉伸（或压缩）弹性畸变、塑性畸变或断裂失效。下面就这些力学性能，即抗力指标进行简单介绍。

比例极限 σ_p，是应力与应变成比例关系的最大应力，即在应力-应变曲线（图5.3-3b）上开始偏离直线时的应力，即有当 $\sigma \leqslant \sigma_p$ 时，$\sigma = E\varepsilon$。

弹性极限 σ_e，是材料由弹性变形过渡到弹-塑变形的应力。当构件承受的实际应力超过材料的弹性极限后，便开始发生塑性变形。

屈服点 σ_s，是应力不增加或开始下降，试样的应变还继续增加所对应的应力值。它是具有屈服现象的材料特有的强度指标。但在工程上只有少数材料有明显的屈服现象，因此工程上规定产生0.2%残余伸长的应力作为屈服强度，以 $\sigma_{0.2}$ 表示。

抗拉强度（强度极限）σ_b，是试样拉断前的最大载荷所决定的条件临界应力，即试样所能承受的最大载荷除以原始面积，

$$\sigma_b = \frac{P_b}{A} \tag{5.3-25}$$

式中，P_b 为拉断前试样所承受的最大载荷；A 为试样的原始截面积。

断裂强度 σ_k，是拉断试样的真实应力，它等于拉断时的载荷除以断裂后缩颈处的截面积。

下面简单介绍一些压杆的失效及其失效抗力指标。

压杆的失效形式与拉杆的失效形式相同，即同样存在压缩性失稳（畸变）和压缩断裂。此外，压杆的设计计算除了强度设计外，还应考虑刚度设计。压杆在压缩时同样可以得到应力-应变曲线，如图5.3-4所示，图中虚线表示低碳钢的拉伸应力-应变曲线，实线表示低碳钢的压缩应力-应变曲线。可见在屈服前，两条曲线重合，但低碳钢（塑性材料）压缩时，试件越压越粗，并不发生断裂，因此测不出它的抗压强度极限。而对于脆性材料来说，拉、压的应力-应变曲线都没有明显的直线部分，也不存在屈服极限（图5.3-5）。脆性材料拉伸时的强度极限低，塑性差，但其抗压能力强，因此脆性材料多用作承压构件。

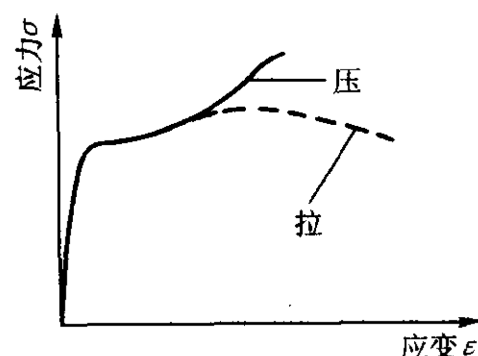


图5.3-4　塑性材料拉、压应力-应变曲线

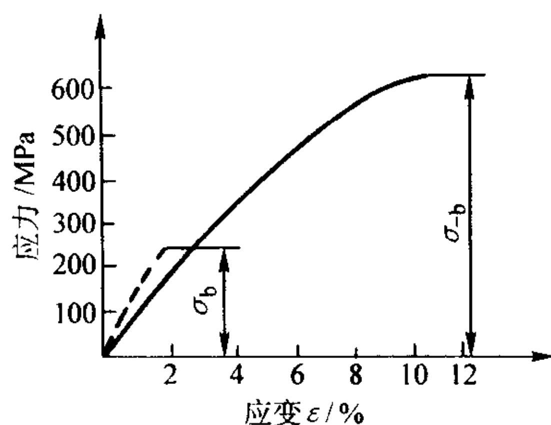


图5.3-5 脆性材料拉、压应力-应变曲线

(2) 受剪、扭零件的失效和失效抗力指标

许多零件是在切应力作用下工作的，如键、齿轮等。对塑性材料来说受剪、扭的变形和断裂过程中一般均先发生明显的剪切变形，当剪切面上的应力达到某一临界值大小时，则沿受剪面或沿扭矩轴的横截面扭断；对于脆性材料来说，其扭转变形很小，并最后沿轴的最外层与轴线约成45°倾角的螺旋形曲面拉断。

受剪零件的失效方式一般为两种，一是过量的剪切变形，发生过量剪切变形的零件的中心线呈现明显的错位；另一种是断裂，一般发生在受剪面上，并往往有摩擦的痕迹。

受扭零件的失效也有两种，一是过量的扭转变形，另一种是扭断。

材料抗剪切或抗扭转变形和断裂的力学性能指标可通过专门的扭转试验机或在万能材料试验机上获得。扭矩 M 与切应力 τ 成正比，扭转角 φ 与切应变 γ 成正比，因此，$M-\varphi$ 图与 $\tau-\gamma$ 图的形状一致，均称为扭转图，即图5.3-6。

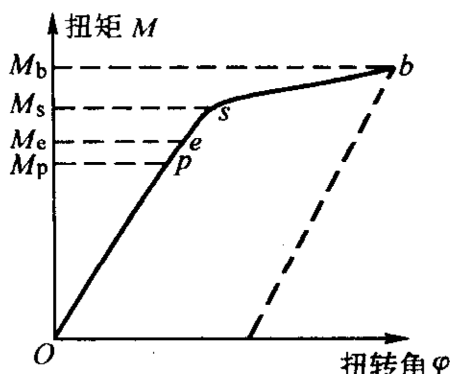


图5.3-6 扭转图

在扭转变形的初始阶段，M 与 φ 成直线关系，此时切应力 τ 和切应变 γ 存在下述关系，

$$\tau = G\gamma \tag{5.3-26}$$

式中，G 为切变模量。由材料力学的知识可知，在单项拉伸时，E（从拉伸图上得到的弹性模量）和 G 存在以下关系，

$$G = \frac{E}{2(1+\mu)} \tag{5.3-27}$$

式中，μ 为泊松比。

图5.3-6中 p、e、s、b 点分别为扭转时的比例极限 τ_p、弹性极限 τ_e、屈服极限 τ_s 及强度极限 τ_b 等。

(3) 受弯、弯冲零件的失效和失效抗力指标

在机器或结构中，如机车的轮轴、电机轴等杆件，它们的受力特点是：在轴线平面内受到力偶或垂直于轴线方向的外力。

其变形特点是：杆件轴线弯曲成曲线。

一般以弯曲为主要变形的杆件，习惯上称为梁。该类零件失效的特点是过量的弯曲塑性变形和弯曲断裂。材料的容许弯曲应力可以按下式求得：

$$[\sigma] = \frac{\sigma_{bb}}{n} \tag{5.3-28}$$

式中，$[\sigma]$ 为容许弯曲应力；σ_{bb} 为材料的抗弯强度，可以通过弯曲试验测得；n 为安全系数。

用标准弯曲试样在万能材料试验机上做弯曲试验时，可得载荷 P-挠度 f 曲线和弯矩 M-挠度 f 曲线（这两种曲线形状相同），如图5.3-7所示。

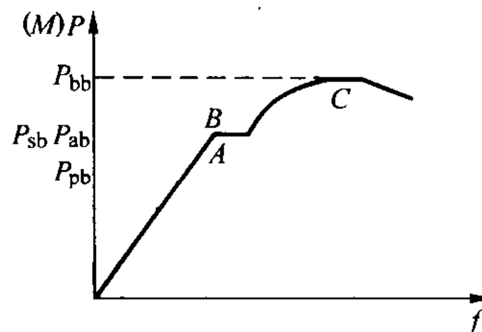


图5.3-7 载荷 P-挠度 f 曲线

图5.3-7中 A 点所对应的 P_{pb} 为弯曲时的比例极限载荷；B 点所对应的 P_{sb}，P_{ab}，为弯曲时的弹性极限和屈服极限载荷；C 点对应的为弯曲时强度极限载荷 P_{bb}。弯曲时试样下表面所受的最大的应力 σ 为：

$$\sigma = \frac{M}{W_z} \tag{5.3-29}$$

式中，M 为最大弯曲力矩；W_z 为抗弯截面模量。

2.1.3 受复杂应力零件的失效分析技术和方法

前面讨论了构件的拉、压扭转、剪切和弯曲等几种变形时的强度计算问题和失效问题。但实际上大多数构件在外力作用下，常常是包括两种或两种以上的变形，这种情况称为组合变形，如复合弯曲、拉伸（压缩）与弯曲的组合、扭转与弯曲的组合等。

人们在大量观察和分析研究了各种类型的材料在不同受力条件下的破坏情况后，得出一个概念，即认为材料受外力作用而发生破坏失效时，不论其失效的表面现象如何复杂，其破坏形式总是几种。若能确定引起某种破坏形式的原因，就可以通过个别应力状态下的试样结果，来推测材料在复杂应力状态下的强度，从而建立相应的强度条件。可以认为在复杂应力状态下，构件的破坏是由危险点处的下列因素之一所引起的：即最大正应力、最大线应变、最大切应力和比能等，而关于材料强度破坏决定因素的各种假设称为强度理论。

一般认为材料的破坏可分为脆断和屈服两大类，因此与之相对应的强度理论也分为两大类：一是揭示材料脆断失效的强度理论，其中有最大拉应力理论和最大拉应变理论；另一类是解释材料塑性屈服破坏的强度理论，其中有最大切应力理论和均方根切应力理论。通过这些强度理论，即可对受复杂应力零件的失效情况进行分析。

(1) 最大拉应力理论（第一强度理论）

脆性材料的破坏方式都是拉应力作用而造成的分裂。因而该理论认为决定材料产生断裂失效的主要原因是单元体的最大拉应力 σ_1，不论它是单向应力状态还是复杂应力状态，只要单元体中的最大拉应力 σ_1 达到材料在轴向拉伸的极限应力值 σ_f，就将发生断裂失效，即有

$$\sigma_1 = \sigma_f \text{，在工程上近似用 } \sigma_1 \cong \sigma_b \tag{5.3-30}$$

将抗拉强度极限 σ_b 除以安全系数，可得许用应力 $[\sigma]$，因此有按第一强度理论建立的强度条件是

$$\sigma_1 \leqslant [\sigma] \tag{5.3-31}$$

但这一理论没有考虑其他两个主应力对材料断裂破坏的影响，且对于单向压缩、三向压缩等应力状态也无法应用。

(2) 最大拉应变理论（第二强度理论）

当作用在构件上的外力过大时，其危险点处的材料会沿最大伸长线应变的方向发生脆性破坏。于是该理论认为，材

料发生断裂破坏的主要原因时单元体的最大拉应变 ε_1，即不论它是单向应力状态还是复杂应力状态，只要单元体中的最大拉应变 ε_1 达到材料在单轴拉伸情况下的拉应变的极限值 ε_f，就将发生断裂失效。由胡克定律，$\varepsilon_f=\sigma_b/E$，发生断裂破坏的条件是

$$\varepsilon_1=\varepsilon_f=\frac{\sigma_b}{E} \tag{5.3-32}$$

由广义胡克定律有

$$\varepsilon_1=\frac{\sigma_1}{E}-\mu\left(\frac{\sigma_2}{E}+\frac{\sigma_3}{E}\right) \tag{5.3-33}$$

代入式（5.3-32）有 $\sigma_1-\mu(\sigma_2+\sigma_3)=\sigma_b$

将 σ_b 除以安全系数，可得许用应力［σ］，按第二强度理论建立的强度条件即为，

$$\sigma_1-\mu(\sigma_2+\sigma_3)\leqslant[\sigma]$$

在单向拉伸的情况下，该理论与第二强度理论完全一致，但是上面公式所用的［σ］是材料在单向拉伸时的许用拉应力，只对在单向拉伸时沿横截面发生脆断破坏的材料才适用，而对塑性材料来说，不可能通过单向拉伸试样得到材料在脆断时的极限值 ε_f 的，因此该理论对塑性材料不适用。

(3) 最大切应力理论（第三强度理论）

当作用在构件上的外力过大时，其危险点处的材料会沿最大切应力所在的截面滑移而发生屈服失效。于是该理论认为，材料发生塑性屈服失效的主要原因是单元体的最大切应力 τ_{max} 达到单轴拉伸下发生塑性屈服时的极限切应力值 τ_s 时，材料即发生塑性屈服失效。

单向拉伸的情况下，当横截面上的拉应力达到极限应力 σ_s 时，与轴线成45°的斜截面上相应的极限切应力为 $\tau_s=\frac{\sigma_s}{2}$，发生屈服破坏的条件是

$$\tau_{max}=\tau_s=\frac{\sigma_s}{2} \tag{5.3-34}$$

在复杂应力状态下的最大剪应力为

$$\tau_{max}=\frac{\sigma_1-\sigma_3}{2}$$

代入上式，有　　$\sigma_1-\sigma_3=\sigma_s$　　(5.3-35)

第三强度理论实际上是一种塑性变形判据，而不是断裂判据。该理论能够合理的解释塑性材料出现塑性变形的条件和现象。

(4) 均方根切应力理论（第四强度理论）

这一理论认为，决定塑性材料屈服失效的主要原因是单元体的均方根切应力 τ^*。它在数值上与单元体的三个切应力 $\tau_{12}=\frac{\sigma_1-\sigma_2}{2}$，$\tau_{23}=\frac{\sigma_2-\sigma_3}{2}$，$\tau_{31}=\frac{\sigma_3-\sigma_1}{2}$ 有关，即

$$\tau^{12*}=\sqrt{\frac{\tau_{12}^2+\tau_{23}^2+\tau_{31}^2}{2}}=\sqrt{\frac{1}{2}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2]} \tag{5.3-36}$$

当一个单元体上的均方根切应力达到单向拉压危险状态时的均方根剪应力 τ_s^* 时，材料发生塑性屈服失效，即其塑性屈服失效的条件为：

$$\tau^*=\tau_s^*$$

于是，其强度条件为，

$$\sqrt{\frac{1}{2}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^3+(\sigma_3-\sigma_1)^2]}\leqslant[\sigma]$$

由于从形状改变比能理论或从八面体面上的切应力理论推导的结果与上式相同，因此，均方根切应力理论亦可称为形状改变比能理论或称八面体面上切应力理论。

以上四个强度理论所建立的强度条件可以统一写为，

$$\sigma^*\leqslant[\sigma]$$

式中 σ^* 是根据不同强度理论所得到的危险点处的三个主应力的组合。四个强度理论的相当应力表达式归纳如表5.3-10所示。

表5.3-10　四个强度理论的相当应力表达式

强度理论的名称及分类		相当应力表达式
第一类强度理论（脆断破坏的理论）	第一强度理论——最大拉应力理论	$\sigma_1^*=\sigma_1$
	第二强度理论——最大拉应变理论	$\sigma_2^*=[\sigma_1-\mu(\sigma_2+\sigma_3)]$
第二类强度理论（屈服破坏的理论）	第三强度理论——最大切应力理论	$\sigma_3^*=\sigma_1-\sigma_3$
	第四强度理论——均方根切应力理论	$\sigma_4^*=\sqrt{\frac{1}{2}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2]}$

(5) 各种强度理论的使用范围

在三向拉伸应力状态下，不论是脆性材料或是塑性材料，都会发生脆性断裂，因此宜采用最大拉应力理论。

对于脆性材料，在二向拉伸应力状态下应采用最大拉应力理论。

对于低碳钢这类的塑性材料，在除了三向拉伸应力状态以外的各种复杂应力状态下，材料都会发生屈服现象。因此宜采用均方根切应力理论，但是有时也可采用最大切应力理论，这是因为最大切应力理论的物理概念较为直观，按此理论所计算也偏于安全。

在三向应力状态下，不论是塑性材料还是脆性材料，通常都发生屈服破坏，故一般采用均方根切应力理论。

这里应当指出，在不同的情况下究竟如何选用强度理论，这并不单纯是力学问题，而与有关工程技术部门长期积累的经验，以及根据这些经验制定的一整套计算方法和规定的许用应力值等都有关系。

2.1.4　应力集中情况下的失效分析技术和方法

一般情况下，等截面杆受轴向拉压时，其截面上应力分布应当是均匀的。但在实际情况下，由于构件中常存在一些油孔、切槽等，在截面突变处有应力突然增加的现象，此种现象称为应力集中。图5.3-8，图5.3-9分别表示在存在圆孔和半圆角时截面被削弱的内拉伸应力分布情况。应力集中程度以应力集中系数来说明，当静态应力不超过弹性极限时，缺口应力集中程度可用理论（或弹性）应力集中系数 K_t 来表示，其中拉伸理论应力集中系数用 $(K_t)_b$ 来表示，即有

$$K_t=\frac{\sigma_{max}}{\sigma_n},(K_t)_b=\frac{\sigma_{max}}{\sigma_n}$$

式中，σ_{max} 为应力集中区的最大应力；σ_n 为名义应力，且有 $\sigma_n=\frac{P}{S_{min}}$，其中 P 为杆件所承受的轴向应力，S_{min} 为被削弱截面的面积，称为净面积。

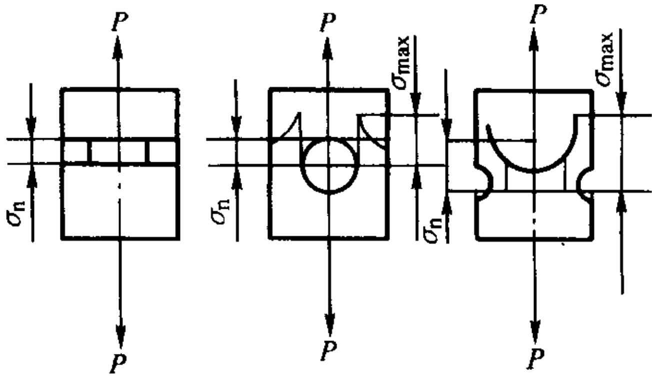


图5.3-8　应力集中情况示意图

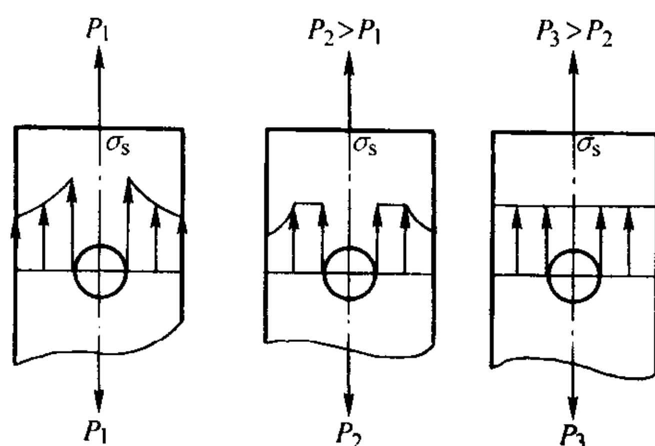


图5.3-9　塑性材料缺口根部应力集中分布示意图

系数 K_t 的取值取决于构件的受力情况、几何形状和缺口的形状等。

有时在零件的一种缺口上又叠加了另一种形状的缺口，例如在圆角上又刻有划痕，这时其最终的复合理论应力集中系数 $(K_t)_{复合}$ 为

$$(K_t)_{复合}=(K_t)_{圆角}(K_t)_{划痕}$$

对于脆性材料，当其静应力不超过弹性极限时，实际的应力集中系数基本上与理论的应力集中系数相等，因此可以 $\sigma_{max}=K_t\sigma_n$ 来估算缺口根部的最大应力。当 σ_{max} 达到该材料的强度极限时，零件则会在此处开裂，而此时零件的心部仍处于较低的应力状态，因此脆性材料在其断裂时的平均应力比光滑试样的抗拉强度 σ_b 低，因此应力集中对脆性材料来说是特别危险的。

而对于塑性材料来说，当外力大到使缺口根部达到屈服极限时，缺口根部的材料就沿着平行轴向载荷方向发生塑性伸长，通过塑性变形而发生应力松弛。当外力增加时，处在弹性应变的其他部位的应力继续增加，直至整个截面达到屈服极限。且因为塑性变形要保持试样的体积不变，所有缺口根部的材料也要进行横向收缩，但是试样心部仍处在弹性应力状态部分的限制，就形成其他方向上的约束，并使得缺口附近的材料处在三向拉应力状态，且此三向拉应力状态提高了材料的屈服做需要的轴向应力，即由于缺口根部的应变集中，使材料实际的屈服极限得到了提高。

以上为应力集中对受拉零件的影响，同理应力集中对受剪、扭和弯的零件的失效的影响也可用相应的应力集中系数来表示。

如扭转时，在应力集中区的最大应力为，

$$\tau_{max}=(K_t)_\tau\tau_n$$

式中，τ_n 为按材料力学方法计算出的标称应力；$(K_t)_\tau$ 为扭转时的应力集中系数，亦可称为理论扭转应力集中系数。

而弯曲时在截面尺寸和形状急剧变化的区域产生的应力集中系数可用 $(K_t)_t$ 来表示，

$$(K_t)_t=\frac{\sigma_{max}}{\sigma_n}$$

式中，σ_{max}为应力集中处的最大应力；σ_n 无应力集中情况下的弯曲应力。

在一般情况下，当零件的几何形状相同时，拉伸理论应力集中系数 $(K_t)_b$ 最大，弯曲理论应力集中系数 $(K_t)_t$ 次之，扭转理论应力集中系数 $(K_t)_\tau$ 最小。

2.1.5　残余应力的失效分析技术和方法

残余应力是指机械构件在无外加载荷的作用下，存在于整个机械构件内部或在较大尺寸的宏观范围内均匀分布并保持平衡的一种内应力。当内应力存在整个或在较大的范围内保持，则此内应力叫做宏观应力或残余应力。残余压应力能够提高疲劳极限和抗腐蚀能力。但有些残余应力也可能使零件在使用过程中形状尺寸发生变化，致使其失效。据调查分析表明由于残余应力的影响而引起的机械构件的失效达到50%，因此对残余应力的失效分析技术和方法应当给予详细的分析和研究。

(1) 残余应力的产生

机械构件经过冷热加工（如切削、磨削、铣削等加工、装配、冷拔以及热处理）后，其内部或多或少都存在一定的残余应力，具体分为下列几种：

1）焊接引起的残余应力；

2）热处理引起的残余应力；

3）表面化学处理引起的残余应力；

4）切削加工引起的残余应力。

(2) 残余应力的检测

宏观内应力即残余应力的测定方法有很多种，其中包括电阻应变片法、光弹性复膜法、脆性涂料法、X射线法以及声发射检测法等。这些方法实质上都不外乎首先检测其应变，再通过弹性力学定律由应变计算出应力的数值。

目前还有一种残余应力的测定方法，即X射线应力测定法。该方法是利用布拉格定律：

$$2d\sin\theta=n\lambda$$

即内应力的存在使金属的晶格发生变化，从而造成晶粒中晶面间距 d 增大或减小，最终导致衍射峰宽化。

该应力测定方法有以下优点：不损坏零件；它所测定的仅仅是弹性应变，而不含塑性应变（因为零件发生塑性变形时，其晶面间距并不改变）；X射线照射被测零件的截面可以小到1 mm直径，因而该方法能够研究特定小区的局部应力和突变的应力梯度，而其他测定法所测定的大都是较大区域的应力平均值；同时它还可以用剥层方法测定沿层深分布的应力。

但对一些复杂形状的零件（造成X射线入射、反射困难，如深孔等）还存在测定准确度不高或不能测试的问题。

(3) 有害残余应力的消除

对于有害的残余应力可以通过加热、加静载荷或动载荷的方法进行消除（或降低）。

1）去应力退火　即把材料加热到某一温度，使残余应力超过材料的屈服强度 σ_s，从而使材料发生塑性变形。

2）回火或自然时效处理　回火时由于组织的变化以及加热时临界切应力的降低从而消除残余应力或将其重新分配。

3）加载处理　加静载可以使残余应力部位发生屈服，从而产生残余应力松弛；加动载（如振动或锤机、喷丸等）亦可消除一定的残余应力。

4）火焰烘烤法　即在焊缝两侧用火焰加热，并随即喷水冷却，造成一个热应力场叠加在焊接残余应力场上，使残余应力减小。

当然，在加工装配当中由于装配不当也会产生一定的附加应力，此应力亦需尽量避免。

2.1.6　热应力的失效分析技术和方法

除了前面所提的机械应力外，温度场、电场、磁场等也可以引起受约束的变形，并产生内应力。而其中热应力或温差应力在近年来越来越受到重视。

(1) 热应力的来源和特点

通常，当物体的温度发生变化时，由于它和不能自由伸缩的其他物体之间或物体内部各部分之间相互约束产生的应力称为热应力。一般，只考虑外力作用引起的应变，但实际上温度的变化引起的热胀冷缩受到约束时，也会产生应力和应变，例如组成构件的材料具有不同的膨胀性能，或构件的各部分的温度分布不均匀等。

外载引起的机械应力和受约束引起的热应力之间有许多共同之处：①都可分为宏观和微观应力和应变；②两种变形均可能是弹性和塑性的；③都可分为静力加载和冲击加载等。但二者之间又有明显的区别：

1）热应力几乎总是变形引起的，总具有强烈的松弛特

性；而机械应力则可以是完全不变的；

2）同恒温变形相比，在不均匀温度场下，加载和变形过程可能使材料性能和物理性能发生很大的变化；

3）温度变化，特别是周期性的强度变化，可能显著加速组织变化；

4）机械载荷和热载荷都产生宏观和显微应力，但在热载荷下，显微应力的作用通常比机械载荷大。

(2) 简单情况下的热应力计算

热应力的计算离不开材料的线胀系数 α，其物理意义是单位长度的物体，当温度升高 1℃时的伸长量，单位是 1/℃。它随温度和材料的不同而不同。

假设有一边长为 1 cm 的各向同性的立方体，因均匀受热而在长、宽、高三个方向是产生同样的膨胀伸长，这种应变状态没有切应变而仅有纵向应变。当温度由 T_1 升为 T_2 时，各方向上的应变均为，

$$\varepsilon = \alpha(T_2 - T_1) = \alpha\Delta T \qquad (5.3\text{-}37)$$

由上式可见热应变 ε 与材料的线胀系数 α 和温度差 ΔT 成正比。下面给出几种不同情况下的热应力计算方法。

1）全约束棒的热应力计算　将长为 L，直径为 d 的圆棒固定在间距为 L 的两壁之中，由 T_1 到 T_2 冷却时，其热应力值为，

$$\sigma = 4P/\pi d^2 = \alpha E(T_1 - T_2) = \alpha E\pi$$

2）两根长度相同、温度不同的棒相互约束的热应力计算　设两根长度相同的棒相互约束，棒的初始长度为 L，温度为 T_0，用下标 1，2 区别各棒的热应力，S_1，S_2 分别为棒 1，2 的横截面积，$\overline{K}$ 为约束系数，当 $\overline{K}>0$ 时，σ_1 为压应力，而 σ_2 为拉应力，若 $\overline{K}<0$，则相反，有

$$\sigma_1 = -\frac{\alpha_1 E_1[(T_1 - T_0) - \alpha_2/\alpha_1(T_2 - T_0)]}{1 + (S_1E_1/S_2E_2)} = -\overline{K}\alpha_1 E_1(T_1 - T_0)$$

$$\sigma_2 = -(S_2/S_1)\sigma_1$$

式中 $$\overline{K} = \frac{1 - \alpha_2[(T_2 - T_0)/\alpha_1(T_1 - T_0)]}{1 + (S_1E_1/S_2E_2)}$$

3）上下表面有温差周边固定板的热应力计算　设上下表面的温度差为 ΔT，D 为板的弯曲刚度，最大弯曲应力为

$$\sigma_{max} = \frac{\alpha\Delta TE}{2(1-\mu)}$$

(3) 热应力引起的塑性变形和蠕变

尽管热应力容易引起松弛和衰减，但热应力仍可能引起构件的塑性变形和蠕变。

材料在工作条件下会同时出现宏观热应力和显微热应力。构件的宏观热残余变形是二者综合作用的结果。在显微热应力和宏观热应力下，重复加热和冷却时，残余变形沿一定的方向增加。即使整体缓慢均匀的加热或冷却，在非立方晶格的金属，如锌、锡等也可能发生随循环次数增加而增加的塑性切变（如图 5.3-10 所示）；而在个别情况下，即使时间很短的高温作用也会明显增加构件的蠕变速率（如图 5.3-11 所示）。

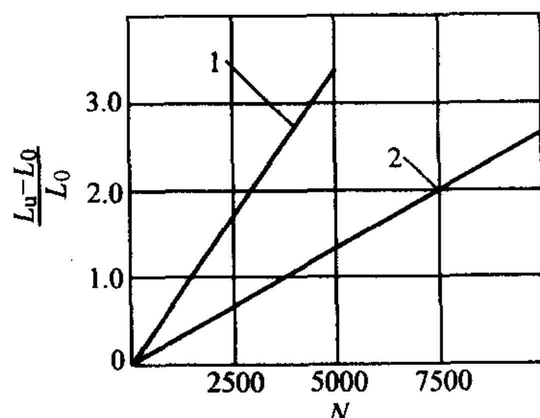


图 5.3-10　相对伸长（$L_u - L_0$）/L_0 和热循环次数的关系

1—Zn，10℃⇌210℃；2—Co，10℃⇌140℃

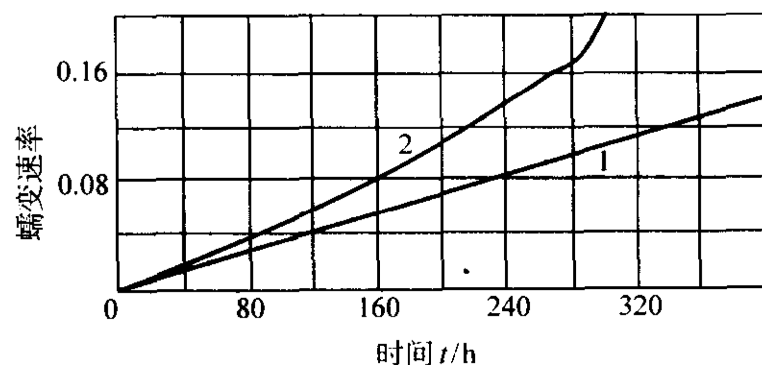


图 5.3-11　耐热合金钢在 122 MPa 应力和不同温度的蠕变量

1—816℃；2—900℃，2 min；816℃，12 h

若热应力是压应力，则可能产生压缩塑性变形，即产生所谓的失稳现象。该现象是许多构件发生重大事故和失效的原因之一。

若在热应力作用下计算弹性状态的构件，主要是计算其临界温差，临界温差 ΔT 可由下式估算：

$$\alpha E\Delta T \leqslant \sigma_s(T)$$

式中，$\sigma_s(T)$ 为发生最大温差下工作时的屈服强度；α 为材料的线胀系数；E 为拉伸弹性模量。

当构件在超过临界温度下工作时，就应该尽可能减少其热应力。

(4) 热冲击

构件在迅速加热或冷却过程中温度急剧的变化，由此而产生热应力，该现象称为热冲击。同机械冲击一样，在热冲击时，变形的不均匀性、惯性抗力和波动过程可能产生一定的作用。

热冲击对脆性材料尤为重要，因为脆性材料在达到断裂应力前，应力和应变基本上保持比例关系，于是一次热载荷可能会引起破坏或失效；而对塑性材料，通常一次热冲击的危害并不十分明显，因为当应力超过屈服强度时，就会迅速被热塑性形变所松弛，导致热应力不可能过多地超过屈服强度。

一般情况下，热冲击问题需联立导热性方程和热弹性方程，从而定量地求解热冲击强度，但由于解的形式过于复杂，难以在实际中应用，因此往往要用简化的理论概念，并通过实验来确定热冲击抗力。

(5) 热疲劳

通常认为，一次热冲击应力会得到不同程度的松弛，因而不会有开裂的危险。但多次急冷应力则难以松弛，因此塑性材料受到重复加热和冷却作用时，也会由于疲劳产生龟裂而导致破坏，这种现象称为热疲劳。

热疲劳和机械疲劳即有相似之处又有区别。相似之处如：①随循环次数的增加，在应力最大的地方产生局部塑性变形和裂纹；②增加应力或应变幅断裂循环周次将减少。其二者的区别之处为：（Ⅰ）在热载荷下，给定的变形决定其应力水平。在热循环时，特别是在大振幅热循环界限温差值有较大的作用，如在循环上限温度（T_{max}），停留时间与塑性变形有很大的关系，因此常伴有显著的蠕变和松弛；而在下限温度（T_{min}）时，塑性变形与时间段关系很小，甚至完全没有关系。若 $T_{max} >> T_{min}$，则在循环的上限温度发生热变形，并且晶间变形和沿晶断裂可能强烈发展。这是因为在 T_{max} 停留期间，扩散、氧化和其他过程引起蠕变和持久断裂的结果。而在 T_{min}，则主要产生晶内变形和穿晶断裂。（Ⅱ）物体内可能出现温度梯度，这将增大试样不同区域变形和断裂的宏观不均匀性。（Ⅲ）热疲劳时，可能显微热应力会起很大的作用。（Ⅳ）热疲劳时应力和应变值与材料的导热性、传热和热膨胀性能有很大的关系；而在机械疲劳时，却没有此关系。总之，机械疲劳和热疲劳的区别随热振幅的增加而增加。

2.2　以断裂力学为基础的失效预测技术和方法

2.2.1　断裂力学基础

防止断裂原来是材料力学的主要任务之一。根据材料力学的结论，对于每种材料要求测定四项机械性能指标：即屈服点 σ_s，抗拉强度 σ_b，延伸率 δ_y 和冲击韧度 a_k，有时还需要加上一项断面收缩率 ψ；σ_s 和 σ_b 是强度指标，而 δ_y、a_k 和 ψ 都是韧性指标。材料力学已经清楚地说明，材料光有足够的强度是不行的，必须同时具有足够的韧性，方可避免断裂的发生。局部的应力集中可能造成数千倍于平均应力的峰值应力。对脆性材料，此项峰值应力由于超过了材料的强度极限，会立即造成断裂；但对韧性材料则峰值应力造成局部的屈服，从而使应力松弛而重新分配，峰值应力被抑制不超过材料的屈服极限，避免了断裂。

以上述四项机械性能指标为依据的设计方法称为常规设计方法。常规设计方法规定平均应力设计值不超过某一“许用应力”值，而延伸率和冲击韧度则不低于某些规定值。根据常规设计方法分析认为是安全的设计，但近半个世纪以来，特别是第二次世界大战以来，随着现代工业的迅速发展和焊接技术的普遍运用，以及设备使用条件的严酷化（温度、介质、载荷波动、原子辐射等），使设备断裂事故增多，一些按照常规的强度理论（材料力学）进行设计并满足强度要求的设备（如压力容器等）会在服役期间甚至在试压期间发生灾难性的破坏（如断裂事故等）。

这些事故的发生促使人们对断裂现象进行了更深入的研究。研究发现材料中存在很多宏观和微观的缺陷如裂缝、夹渣、气泡等，远不是材料力学所假设的各向同性的均匀介质。从这个意义上说，作为常规设计指是由近于理想连续体的小试样得到的屈服点 σ_s 和抗拉强度 σ_b，与实际构件的破坏并无直接关系。因此，常规强度理论并不能完全保证构件的安全服役。

而断裂力学不同于常规强度理论，它是以材料总不可避免地存在有裂纹等缺陷这一事实作为前提的。为此，人们经常把断裂力学称为裂纹（体）力学。也可以说断裂力学是应用连续介质力学理论，研究含有裂纹缺陷的材料和构件的破坏本质，并用定量的方法来确定承载体扩展规律及产生失效条件的工程强度科学。任何工程结构都不可避免地存在有裂纹等缺陷是断裂力学存在的前提。

断裂力学的两大重要分支是线弹性断裂力学和弹塑性断裂力学。线弹性断裂力学主要研究裂纹尖端完全弹性变形及小范围屈服的断裂问题（脆性材料、高强度金属材料）；弹塑性断裂力学主要研究裂纹尖端大范围屈服的断裂问题（中、低强度材料）。

2.2.2　以线弹性断裂力学基础的失效分析技术和方法

线弹性断裂力学理论是断裂力学中最简单也是最基本的理论。其研究基础是线弹性理论，其研究对象是理想的线弹性体即服从胡克定律的材料。它主要从两个角度分析裂纹体的力学性能：一个是通过分析含有裂纹体的应力应变场，得到表征裂纹尖端应力应变场强度的特征参数——应力强度因子 K；另一个是从能量的观点出发，考察裂纹扩展过程中能量的变化，得到表征裂纹扩展的能量变化的参数——能量释放率 G。

(1) 裂纹的基本开裂类型

由于裂纹受力情况不同，会产生三种独立的基本开裂类型，见图 5.3-12。

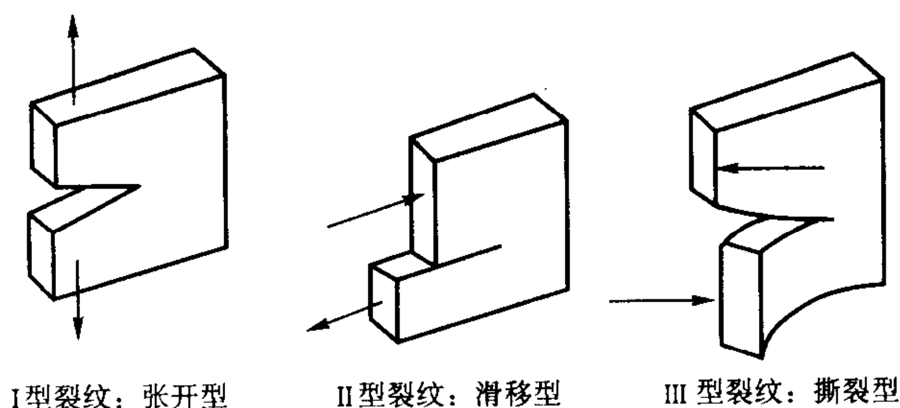


图 5.3-12　三种基本裂纹开裂类型

(2) 裂纹尖端应力场

图 5.3-13 所示为一具有中心穿透裂纹的无限板。板的长度和宽度都为无限大，中心穿透裂纹的长度为 $2a$，裂纹位于 x 轴上，在离裂纹足够远处沿 x 方向和 y 方向作用有均布的拉应力 σ。这就是 Westergaard 在 1939 年提出的关于脆性材料断裂问题的第一力学模型。

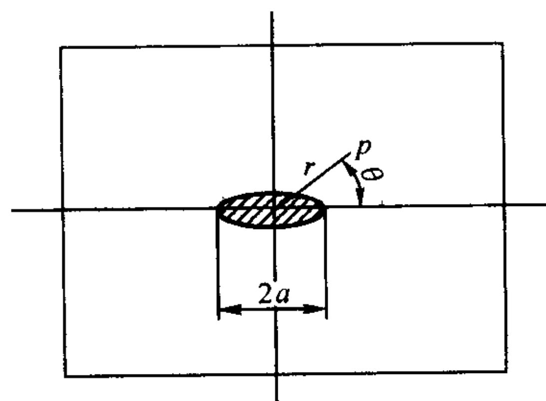


图 5.3-13　带中心穿透裂纹的无限板

应用弹性力学的基本方程，采用 Westergaard 应力函数，并联系该力学模型的边界条件就可以解出该力学问题的应力解。结果表明，在裂纹尖端附近（$r \ll a$）任一点 $p(r,\theta)$ 处的应力近似解对应三种基本开裂类型分别为

Ⅰ型

$$\begin{cases} \sigma_x = \dfrac{K_{\mathrm{I}}}{\sqrt{2\pi r}}\cos\dfrac{\theta}{2}\left(1-\sin\dfrac{\theta}{2}\sin\dfrac{3\theta}{2}\right) \\ \sigma_y = \dfrac{K_{\mathrm{I}}}{\sqrt{2\pi r}}\cos\dfrac{\theta}{2}\left(1+\sin\dfrac{\theta}{2}\sin\dfrac{3\theta}{2}\right) \\ \tau_{xy} = \dfrac{K_{\mathrm{I}}}{\sqrt{2\pi r}}\cos\dfrac{\theta}{2}\sin\dfrac{\theta}{2}\cos\dfrac{3\theta}{2} \\ \sigma_z = \begin{cases} \dfrac{K_{\mathrm{I}}}{\sqrt{2\pi r}}\nu\cos\dfrac{\theta}{2} & (\text{平面应变}) \\ 0 & (\text{平面应力}) \end{cases} \end{cases} \tag{5.3-38}$$

Ⅱ型

$$\begin{cases} \sigma_x = -\dfrac{K_{\mathrm{II}}}{\sqrt{2\pi r}}\sin\dfrac{\theta}{2}\left(2+\cos\dfrac{\theta}{2}\cos\dfrac{3\theta}{2}\right) \\ \sigma_y = \dfrac{K_{\mathrm{II}}}{\sqrt{2\pi r}}\sin\dfrac{\theta}{2}\cos\dfrac{\theta}{2}\cos\dfrac{3\theta}{2} \\ \tau_{xy} = \dfrac{K_{\mathrm{II}}}{\sqrt{2\pi r}}\cos\dfrac{\theta}{2}\left(1-\sin\dfrac{\theta}{2}\sin\dfrac{3\theta}{2}\right) \\ \sigma_z = \begin{cases} -2\nu\dfrac{K_{\mathrm{II}}}{\sqrt{2\pi r}}\sin\dfrac{\theta}{2} & (\text{平面应变}) \\ 0 & (\text{平面应力}) \end{cases} \end{cases} \tag{5.3-39}$$

Ⅲ型

$$\begin{cases} \tau_{xz} = -\dfrac{K_{\mathrm{III}}}{\sqrt{2\pi r}}\sin\dfrac{\theta}{2} \\ \tau_{yz} = \dfrac{K_{\mathrm{III}}}{\sqrt{2\pi r}}\cos\dfrac{\theta}{2} \\ \sigma_x = \sigma_y = \sigma_z = \sigma_{xy} = 0 \end{cases} \tag{5.3-40}$$

式中，K_{I}、K_{II}、K_{III} 分别为Ⅰ、Ⅱ、Ⅲ型加载下裂纹尖端的应力强度因子；ν 为泊松比。

(3) 应力强度因子理论

1) 应力强度因子　由Ⅰ、Ⅱ、Ⅲ型裂纹尖端的应力场表达式（5.3-38）~式（5.3-40）可以看出，裂纹尖端的应

力场可由两部分来描述：一部分是关于裂纹尖端应力场分布的描述，这部分与所考虑点的位置有关，即是点的坐标的函数；另一部分是关于应力场强度的描述，反映应力场的强弱，这部分与裂纹尺寸及所受的应力有关，而与位置无关，简称为应力强度因子，K_{I}、K_{II}、K_{III}，单位为 $\mathrm{MPa \cdot mm^{1/2}}$。

对于一般的裂纹问题，裂纹附近的应力场可用下式表示：

$$\sigma_{ij} = Kf_{ij}(\theta)/\sqrt{2\pi r} \tag{5.3-41}$$

式中，K 为应力强度因子，可用下式表示：

$$K = Y\sigma\sqrt{\pi r} \tag{5.3-42}$$

式（5.3-42）即为一般裂纹体的应力强度因子表达式，其中 Y 为与裂纹体的几何边界有关的修正系数，一般 $Y \geqslant 1$，通过计算或查阅有关手册而得。几种常见的 Y 因子值见表 5.3-11。

表 5.3-11 几种常见的 Y 因子

Y	情况
1.0	无限宽板，中心穿透裂纹，远处均匀拉伸
1.12	半无限宽板，边缘裂纹，埃尔富特处均匀拉伸
$\frac{2}{\pi}$	无限大体，半径为 a 的内埋圆盘裂纹，远处均匀拉伸
$\sqrt{\sec\frac{\pi a}{W}}$	有限宽板，中心穿透裂纹，远处均匀拉伸

对具有中心穿透裂纹的无限大平板中的Ⅰ型裂纹、Ⅱ型裂纹、Ⅲ型裂纹，由表 5.3-11 可见 $Y=1$，故它们的应力强度因子表达式分别为：

$$\begin{cases} K_{\mathrm{I}} = \sigma\sqrt{\pi r} \\ K_{\mathrm{II}} = \tau\sqrt{\pi r} \\ K_{\mathrm{III}} = \tau\sqrt{\pi r} \end{cases} \tag{5.3-43}$$

2）按应力强度因子建立的判据　由式（5.3-38）~式（5.3-40）可以看出，随着 r 的减小，各应力分量值 σ_x，σ_y，τ_{xy}便随之增大。当 $r \to 0$ 时，σ_x，σ_y，τ_{xy}就趋向于无穷大。这就说明，裂纹端部区域的应力场具有奇异性。而该奇异性只有尖裂纹（纹尖端的曲率半径等于零）的情况下才成立，而实际的裂纹都不是尖裂纹。但是也有很多裂纹，其尖端的曲率半径很小，完全可以近似地把它们看成为尖端裂纹。因此，尖裂纹仍是线弹性力学的合理模型，由此得出的应力强度因子（奇异性强度）理论也是正确的，实践已经充分证明了这一结论。

同时随着应力强度因子的增加，各处的应力也随之增大，即应力强度因子 K 控制着裂纹尖端整个应力场应力的大小。由于三种基本裂纹开裂类型中，以Ⅰ型断裂最常见、最基本、也最危险，因此下面以Ⅰ型断裂为主进行讨论。

当应力强度因子 K_{I} 增大到某一临界值 K_c 时，裂纹即发生失稳扩展。反之，当 K_{I} 小于临界值 K_c 时，裂纹就不会发生失稳扩展。因此，若取 K_c 的低限 $(K_c)_{\min}$ 作为临界值，并将 $(K_c)_{\min}$ 记作 K_{Ic}，那么按应力强度因子建立的断裂判据即为

$$K_{\mathrm{I}} = K_{\mathrm{Ic}} \tag{5.3-44}$$

由于 K_{Ic}是表征材料抗断裂性能的一个常数，因此称之为断裂韧度。在一定的条件（温度、板厚、变形速度等）下，各种材料都有确定的 K_{Ic}值，可由实验测定之。K_{Ic}的单位也为 $\mathrm{MPa \cdot mm^{1/2}}$。关于断裂韧度的测定，各国都已制定出各自的规范。表 5.3-12 中给出了几种常用材料的 K_{Ic}值。通过断裂判据即可对构件进行以线弹性断裂力学为基础的断裂失效预测分析。

表 5.3-12 几种常用材料的 K_{IC}值

材料	$K_{\mathrm{Ic}}/\mathrm{MPa \cdot mm^{1/2}}$
12Ni 马氏体时效钢	79~87
Cr15NiCo 不锈钢	43~16
30CrNi2MoV 钢	106
34CrNi3Mo 钢	78
34SiMnCrMoV 钢	96~101
60Si2Mn 钢	85~91
14MnMoNbB 钢	155~171
45# 钢	63~65
45# 钢（在温度 40℃）	122
40# 钢	62~71

3）应力强度因子的确定　在线弹性断裂力学分析中，应力强度因子的计算是一个最困难的问题，其求解方法也已达几十种，大体可分为两大类：一类是理论计算方法（包括解析法和数值方法），另一类是实验方法。下面给出几种常用的应力强度因子公式。

① 无限体内埋入椭圆片状裂纹（图 5.3-14）的应力强度因子

$$K_{\mathrm{I}} = \frac{\sigma\sqrt{\pi a}}{\Phi} \tag{5.3-45}$$

式中，Φ 是第二类椭圆积分：

$$\Phi = \int_0^{\pi/2}\sqrt{1 - \frac{c^2 - a^2}{a^2}\sin^2\phi}\,\mathrm{d}\phi$$

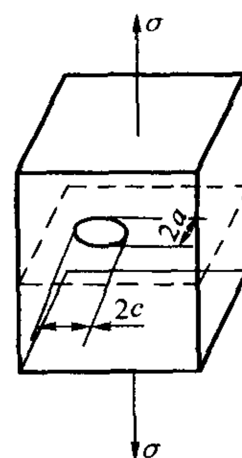


图 5.3-14 无限体内埋入椭圆片状裂纹

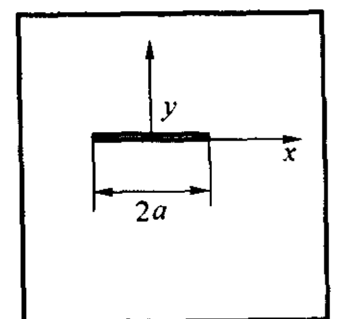


图 5.3-15 无限板上穿透裂纹

② 无限板上穿透裂纹（图 5.3-15）的应力强度因子由复变函数直接计算得到Ⅰ、Ⅱ型复合加载下的复应力强度因子

$$K = \lim_{z \to 0} 2\sqrt{2\pi z}\varphi'(z) \tag{5.3-46}$$

式中，$z = x + iy$，$\varphi'(z)$为可根据问题的边界条件得出的应力函数。对于纯Ⅰ型加载的单向均匀受拉的无限板中心穿透裂纹的应力强度因子为

$$K_{\mathrm{I}} = \sigma\sqrt{\pi a} \tag{5.3-47}$$

③ 表面半椭圆裂纹的应力强度因子　表面裂纹是最为常见的裂纹，但由于表面裂纹问题是一个三维问题，数学上处理起来非常困难，因此很多方法都是以图 5.3-16 所示的

表面半椭圆片状裂纹作为表面裂纹来进行研究。这里介绍比较常见的 Paris - Sih 公式。该公式考虑了前表面（自由表面）和后表面的影响，他们提出的 $K_{\rm I}$ 计算式为

$$K_{\rm I}=M_1M_2\frac{\sigma\sqrt{\pi a}}{\Phi}\tag{5.3-48}$$

式中，$M_1=1+0.12\left(1-\frac{a}{c}\right)$ 是考虑前自由表面影响的修正系数，当 $\frac{a}{c}\to 0$ 时，$M_1=1.12$；$M_2=\sqrt{\frac{2B}{\pi a}\tan\frac{\pi a}{2B}}$ 是考虑后表面影响的修正系数。

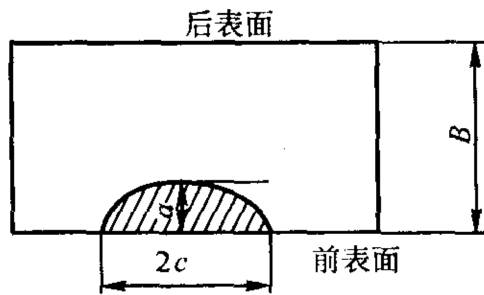


图 5.3-16　无限体表面半椭圆片状裂纹

此式的缺点是未考虑塑性区的修正。若再考虑塑性区的修正，则式（5.3-48）即可以写成

$$K_{\rm I}=M_1M_2\frac{\sigma\sqrt{\pi a}}{\sqrt{Q}}$$

式中，$Q=\Phi^2-0.212\left(\frac{\sigma}{\sigma_s}\right)^2$ 为表面裂纹的形状参数。

4）塑性区的修正及线弹性断裂力学的适用范围　如前所述，随着 r 的减小，各应力分量值 σ_x，σ_y，τ_{xy} 便随之增大。当 $r\to 0$ 时，σ_x，σ_y，τ_{xy} 就趋向于无穷大。然而对金属材料来说，裂纹尖端会出现在一个小的塑性区，应力并不会无限增大。但该塑性区会对线弹性断裂力学分析中的公式的精度与适用范围产生影响，因此需要对确定线弹性断裂理论的适用范围，并对线弹性断裂力学公式进行塑性区修正。

① 塑性区的形状和尺寸　由 Von Mises 屈服准则，材料在三向力状态下的屈服条件为，

$$(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2=2\sigma_s^2$$

式中，σ_s 为材料的屈服强度，σ_1，σ_2，σ_3 为主应力，可按下式求出，

$$\left.\begin{aligned}\sigma_1&=\frac{1}{2}(\sigma_x+\sigma_y)+\frac{1}{2}[(\sigma_x-\sigma_y)^2+4\tau_{xy}^2]^{1/2}\\ \sigma_2&=\frac{1}{2}(\sigma_x+\sigma_y)-\frac{1}{2}[(\sigma_x-\sigma_y)^2+4\tau_{xy}^2]^{1/2}\\ \sigma^3&=\begin{cases}\nu(\sigma_1+\sigma_2) & (\text{平面应力})\\ 0 & (\text{平面应变})\end{cases}\end{aligned}\right\}\tag{5.3-49}$$

式中，ν 为泊松比。将式（5.3-38）的 σ_x，σ_y 和 τ_{xy} 代入式（5.3-49）即可得到 Ⅰ 型裂纹端塑性区的形状方程，如图 5.3-17 所示。

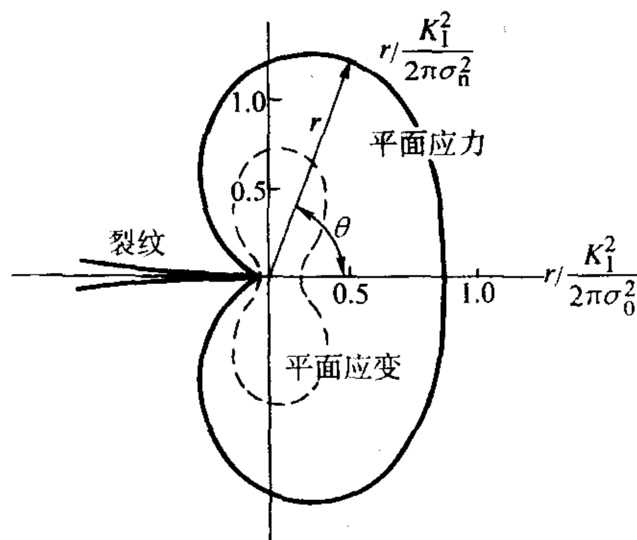


图 5.3-17　裂纹尖端的塑性区

$$r(\theta)=\begin{cases}\frac{K_{\rm I}^2}{2\pi\sigma_s^2}\cos^2\frac{\theta}{2}\left(1+3\sin^2\frac{\theta}{2}\right)(\text{平面应力})\\ \frac{K_{\rm I}^2}{2\pi\sigma_s^2}\cos^2\frac{\theta}{2}\left[3\sin^2\frac{\theta}{2}+(1-2\nu)^2\right](\text{平面应力})\end{cases}\tag{5.3-50}$$

以上是裂纹尖端塑性区的一次近似解。若再考虑应力松弛和材料硬化等因素则塑性区的尺寸 $r(\theta)$ 扩大到 R_y 为：

$$R_y=\begin{cases}\frac{K_{\rm I}^2}{(1+n)\pi\sigma_s^2} & (\text{平面应力})\\ \frac{K_{\rm I}^2}{(1+n)\pi\sigma_s^2}(1-2\nu)^2 & (\text{平面应力})\end{cases}\tag{5.3-51}$$

式中，n 为材料的硬化指数。

由于这里所给出的公式是通过弹性力学方法得到的，因而所得的裂纹端部塑性区的尺寸是一个粗略的估计值。

② 应力强度因子的塑性区修正方法及适用范围　Irwin 提出了一种现为人们接受的修正方法。他认为塑性区使裂纹的刚度减小，对 $K_{\rm I}$ 的影响相当于增加了裂纹的长度（增加了半个塑性区长度）。对于一些复杂的情况，应力强度因子的修正值 $K'_{\rm I}$ 很难得到解析式。而对一些相对简单的情况可以通过近似处理求出解析式，如对具有中心穿透裂纹的无限大平板中问题，其应力强度因子的修正值 $K'_{\rm I}$ 可表示为

$$K'_{\rm I}=\frac{\sigma\sqrt{\pi a}}{\sqrt{1-(\sigma/\sigma_s)^2/2}}\quad(\text{平面应力})\tag{5.3-52}$$

$$K'_{\rm I}=\frac{\sigma\sqrt{\pi a}}{\sqrt{1-(1-2\nu)^2(\sigma/\sigma_s)^2/2}}\quad(\text{平面应变})\tag{5.3-53}$$

线弹性断裂理论的适用条件是处于线弹性状态的裂纹尖端附近，即只适用于弹性区而不适用于塑性区，并且仅在裂纹尖端附近的一个区域内适用。即 $r(\theta)$ 值有上下限，其下限值为塑性区尺寸 R_y；其上限值为 $r(\theta)\ll a$ 的一个值。对中心穿透裂纹的无限大平板而言，平面应变条件下的适用范围为

$$\begin{aligned}a&\geqslant 2.5\left(\frac{K_{\rm Ic}}{\sigma_s}\right)^2\\ b=B-a&\geqslant 2.5\left(\frac{K_{\rm Ic}}{\sigma_s}\right)^2\end{aligned}\tag{5.3-54}$$

式中，b 为试样的韧带尺寸，$b=B-a$；B 为试样的宽度。

2.2.3　以弹塑性断裂力学为基础的失效分析技术和方法

如上所述，裂纹在扩展前，其尖端会出现一个塑性区，线弹性断裂力学只适用于该塑性区很小，即远小于裂纹尺寸的情况。若该塑性区很大甚至超过裂纹本身的尺寸时，断裂即不属于小范围屈服断裂，因此线弹性断裂理论将不再适用，而需要采用弹塑性断裂力学理论来进行分析。此外，工程中还存在另一类断裂问题——全面屈服问题。在范围屈服和全面屈服均属于弹塑性断裂范畴。解决这类问题的失效分析的理论基础即为弹塑性断裂理论。

目前用于弹塑性断裂力学研究的方法主要有 COD 法（裂纹张开位移法）和 J 积分法两种。

（1）COD 法

COD 法即裂纹张开位移（Crack Opening Displacement）法。COD 理论是由 Wells 于 1963 年首先提出的。该分析方法主要是建立在经验基础上的，它在工程界已得到较广泛的应用。

1）COD 定义　所谓 COD，就是裂纹尖端的裂纹表面张开的位移量。对于延性较好的金属材料来说，当其内部裂纹受载荷作用时，由于裂纹尖端的应力集中，致使该区域的材料发生塑性滑移，从而导致裂纹尖端的钝化（即圆化），裂

纹表面张开同，见图5.3-18。20世纪70年代以来，用电子显微镜对COD试样进行断口形貌分析也证实了裂纹尖端钝化现象的存在。一般可以把COD值作为裂纹端部应力、应变场的度量，用符号δ来表示。实验证明，随着载荷的增大，裂纹尖端的δ值也随之增大。当δ达到某一临界值δ_c时，裂纹开始扩展。因此Wells提出用裂纹尖端张开位移δ作为断裂判据，即

$$\delta = \delta_c \tag{5.3-55}$$

式中，δ_c即为临界COD，是断裂开始时裂纹尖端的张开位移，可采用其作为材料对裂纹扩展阻力的度量，即材料的断裂韧度。

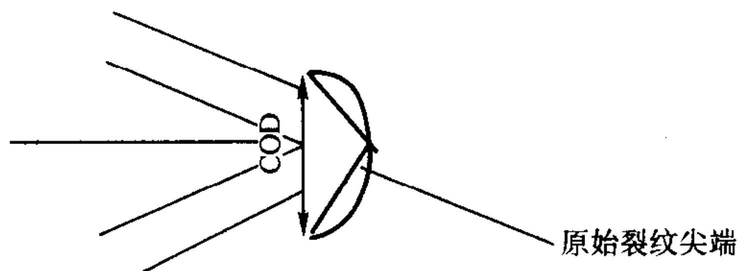


图5.3-18　裂纹尖端钝化模型

试验测量结果表明δ_c是一个不随试样尺寸而变的材料常数。

2）两种典型的裂纹模型的COD计算公式

① BCS公式　从具有贯穿裂纹的大型薄板的拉伸试验中，可以观察到裂纹尖端的塑性区具有扁带状的特征。D.S. Dugdale曾对此问题作过研究，并建立了D－M模型。该模型将弹塑性问题转化成了线弹性问题。通过该模型并应用弹性力学的方法可得出BCS公式，

$$\delta = \frac{8\sigma_s a}{\pi E}\ln\sec\frac{\pi\sigma}{2\sigma_s} \tag{5.3-56}$$

式中，σ_s为屈服强度。

② Wells公式　对于全屈服区内小裂纹的COD分析，在理论上至今尚无完善的解答。Wells根据宽板试验结果，归纳出一个全屈服区内小裂纹的COD计算公式，即

$$\delta = 2\pi ae \tag{5.3-57}$$

（2）J积分理论

COD法虽然也能较有效地解决一些工程实际问题，但它毕竟不是一个直接而严密的应变场参量，涉及裂纹尖端位移场的定义、分析和直接测量也都比较困难。因此，Rice于1968年提出了J积分的概念。J积分定义明确、理论严密的应力、应变场参量，它的实验测定也比较简单可靠。且它还具有与积分路线无关的特点，可以避开裂纹尖端处的极其复杂的应力、应变场。同时它不仅适用于线弹性区域，也适用于弹塑性区域。但由于J积分是一个平面积分，它只能用于二维问题的分析。

1）J积分的定义　J积分有两种定义。一是回路积分定义，另一个是形变功率定义。

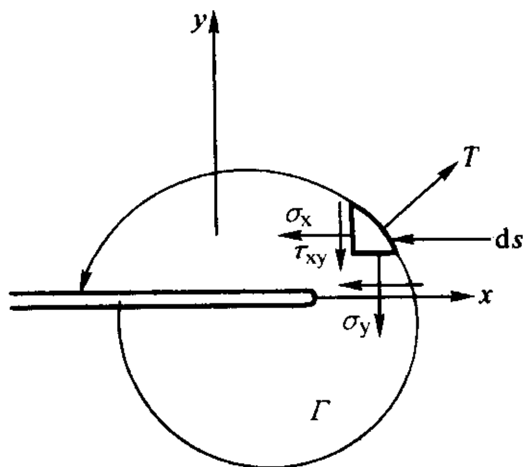


图5.3-19　J积分定义

如图5.3-19所示，对于平行于x轴的穿透裂纹，J积分的定义可由一线积分（回路积分）给出：

$$J = \int_\Gamma \left(W\mathrm{d}y - \vec{T}\frac{\partial\vec{u}}{\partial x}\mathrm{d}s\right) \tag{5.3-58}$$

式中，W为任一点（x，y）的应变能密度，

$$W = \int_0^{\varepsilon_{min}} \sigma_{ij}\mathrm{d}\varepsilon_{ij} \tag{5.3-59}$$

式中，σ_{ij}，ε_{ij}分别为应力分量和应变分量；$\vec{T}$为积分回路上的应力矢量；$\mathrm{d}s$为Γ上的弧元；$\vec{u}$为力$\vec{T}$作用点（或$\mathrm{d}s$弧元上）的位移矢量。

J积分的回路积分定义的物理意义不明确，也不便于用实验来进行直接测定。

形变功率定义：

$$J = -\frac{1}{B}\frac{\mathrm{d}U}{\mathrm{d}a} + \oint_C T_i\frac{\mathrm{d}u_i}{\mathrm{d}a}\mathrm{d}s \tag{5.3-60}$$

式中，B为试样厚度；a为裂纹尺寸；U为总变形功；T_i为应力矢量的分量；u_i为位移矢量的分量；其他符号含义同前。

由于积分$\oint_C T_i\frac{\mathrm{d}u_i}{\mathrm{d}a}\mathrm{d}s$是沿着试样的边界进行的，其值也只与边界上力矢量的分量T_i及位移分量u_i的微商$\frac{\mathrm{d}u_i}{\mathrm{d}a}$有关。因此按此定义计算$J$积分比直接按回路积分定义计算要方便得多，也便于实验标定。

2）两种典型的弹塑性断裂问题的J积分

① D－M模型的J积分　如上所述，D－M模型是一个弹性化了的模型，且消除了弹性奇点，整个区域的应变量都是小应变。因而J积分的路径无关性是准确成立的。平面应力状态下D－M模型的J积分表达式为：

$$J = \frac{8a\sigma_s^2}{\pi E}\ln\sec\left(\frac{\pi\sigma}{2\sigma_s}\right) \tag{5.3-61}$$

② 位于全屈服区内的小裂纹的J积分　全屈服区内小裂纹问题的特点是整个板材已经全面屈服，只有裂纹的自由表面两表面两侧还存在一个小弹性区，因而D－M模型的方法已不再适用。通常应用应力、应变集中理论和J积分理论来分析这类问题，可得其计算式为：

$$J = \frac{2K_I^2}{\sigma}\int\sigma\mathrm{d}e = 2\pi y^2\int\sigma\mathrm{d}e \tag{5.3-62}$$

式中，$y = \frac{K_I}{\sigma\sqrt{\pi a}}$为裂纹几何形状因子。

3）J积分判据　只要J主导区存在，裂纹尖端部区域的应力应变场强度就能由J积分惟一确定，J积分可以作为断裂判据。当裂纹端部的应力场达到J_{cr}（临界值）时，裂纹开始扩展。因此断裂判据可表示为，

$$J = J_{cr} \tag{5.3-63}$$

当$J_{cr} = J_{Ic}$时，上式为平面应变条件下的断裂判据，J_{Ic}为材料的平面应变断裂韧性，它是与试样尺寸无关的材料性能参数；若$J_{cr} = J_c$，则上式为平面应力条件下的断裂判据，J_c为材料的平面应力断裂韧性，它与试样的厚度有关。在线弹性状态下，J积分就是弹性应变能释放率G_I。J积分和G_I及应力强度因子K_I之间的关系为：

$$J = G_I = \frac{K_I^2}{E_1} \tag{5.3-64}$$

式中，$E_1 = \begin{cases} E & \text{（平面应力状态）} \\ \dfrac{E}{1-\mu^2} & \text{（平面应变状态）} \end{cases}$。通过COD和$J$积分判据即可进行以弹塑性断裂力学为基础的失效分析。

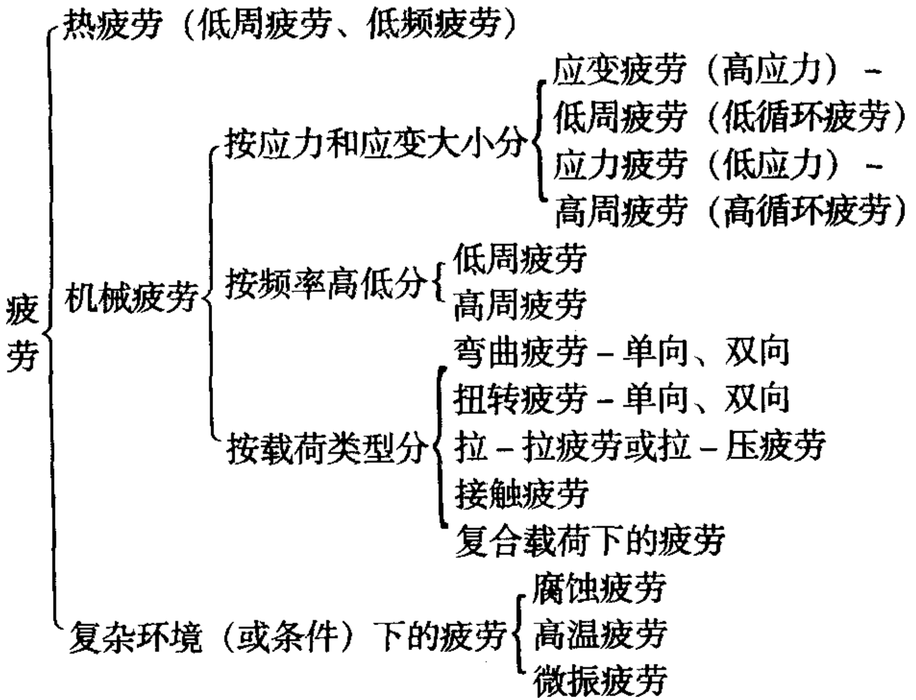
2.3　以疲劳力学为基础的失效预测技术和方法

在工程中有许多构件，如轴、齿轮等都是在交变循环应

力下工作的。所谓“交变应力”是指应力的大小、方向或大小和方向都随时间作周期性的改变的应力。这种改变可能是规律性的，也可能是随机的。构件在这种交变应力作用下的失效称为疲劳失效。

金属的疲劳失效或疲劳破坏具有很大的危害性，据统计，在损坏的零件中，60% ~ 80%是由疲劳引起的。因此，需要对疲劳失效进行深入的研究。

根据应力、应变及环境等的不同，可将疲劳进行如下分类



2.3.1　疲劳力学基础

（1）基本概念

为了清楚地看出应力的变化规律，人们将应力随时间的变化规律绘成如下图形（图 5.3-20、图 5.3-21）。

根据交变应力的对称情况不同，应力循环可以分为五种：

1）当 $R=-1$ 时，对应应力循环为对称循环；

2）当 $R=0$ 时，对应应力循环为脉动拉伸循环；

3）当 $R=\infty$ 时，对应应力循环为脉动压缩循环；

4）当 $0<R<1$ 时，对应应力循环为波动拉伸循环；

5）当 $-1<R<0$ 时，对应应力循环为波动拉压循环。作为特例，可以把静载荷看成是 $R=+1$ 的交变载荷。

图 5.3-20　应力循环图

图 5.3-21　应变循环图

图 5.3-20 中，σ_{max} 为最大应力，$\sigma_{max}=\sigma_m+\sigma_a$；$\sigma_{min}$ 为最小应力，$\sigma_{min}=\sigma_m-\sigma_a$；$\sigma_m$ 为平均应力，$\sigma_m=1/2(\sigma_{max}+\sigma_{min})$；$\sigma_a$ 为应力幅值，$\sigma_a=\frac{1}{2}(\sigma_{max}-\sigma_{min})$；$\Delta\sigma$ 为应力范围，$\Delta\sigma=2\sigma_a$；R 为应力比，$R=\sigma_{min}/\sigma_{max}$。

研究表明在交变载荷作用下，金属承受的最大交变应力 σ_{max} 越大，则至断裂的应力的交变次数 N 越少；反之，σ_{max} 越小，N 越多。将所加的应力 σ_{max} 和其对应的断裂周次 N 绘成图，即可得到图 5.3-22 所示的曲线。该曲线称为疲劳曲线（即 $S-N$ 曲线）。由此图可以看出，当应力低于某值时，材料或构件可以承受无穷多次的应力循环而不会发生疲劳断裂，该值称为材料或构件的疲劳极限，可以用 σ_r 来表示。而发生破坏时的应力循环次数或从开始承受应力直至断裂所经历的时间称为疲劳寿命，通常以 N_f 来表示。通常横轴采用对数坐标表示寿命，纵轴用均匀坐标或对数坐标表示最大应力或应力幅值来绘制的疲劳试验中获得的 $S-N$ 曲线。

图 5.3-22　疲劳曲线示意图

对于低碳钢材料来说，$S-N$ 曲线有一水平渐近线（如图 5.3-23 所示），若这类钢经过 10^7 此循环仍不破坏，即认为它可承受无限次的循环。而铝合金等材料的 $S-N$ 曲线则没有这样的渐近线（如图 5.3-24）所示，因此常采用对应于一定的破坏的循环次数（如 2×10^7）的应力作为“条件疲劳极限”。

图 5.3-23　钢的 $S-N$ 曲线

图 5.3-24　铝的 $S-N$ 曲线

(2) 疲劳抗力指标及其影响因素

在评定金属的疲劳抗力时通常采用以下几个抗力指标：疲劳极限、过载荷持久值、过载荷损害界和疲劳缺口敏感度等。

1) 疲劳极限　上述材料的疲劳极限（或条件疲劳）极限通常是在旋转对称弯曲疲劳试验机上测定的，因为 $R=-1$，所以常用 σ_{-1} 来表示。具体的试验测定方法请详见有关标准。

对于不对称应力循环下的 $S-N$ 曲线如图5.3-25所示，图上的曲线是在不同的应力比 R 值下得到的。当 R 值越正，即 σ_m 越高，则其疲劳极限越高。

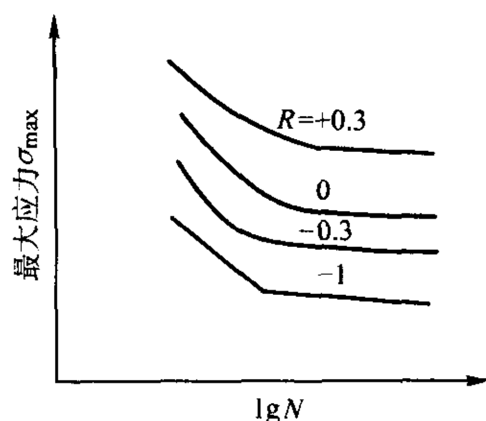


图5.3-25　不对称应力循环下的 $S-N$ 曲线

研究表明，在工程实际当中有很多的因素能够影响材料或构件的疲劳极限，如应力状态、应力循环的对称度、加载频率、工作温度、环境介质、表面损伤、表面粗糙度、材料种类、组织类型、表面处理情况等。因此在测定过程中，疲劳曲线不是一条线，而是一分布带，所绘的疲劳曲线是一组试样结果的统计平均值。

同时研究发现对称弯曲疲劳极限 σ_{-1} 与抗拉强度 σ_b 之间有一定的关系，如图5.3-26所示。

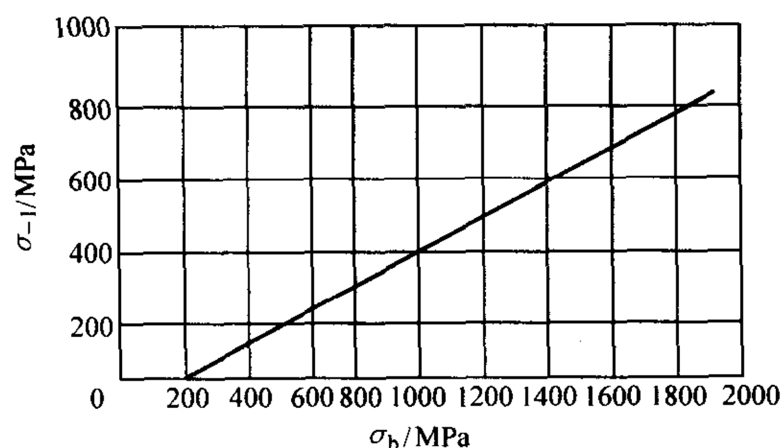


图5.3-26　钢的抗拉强度 σ_b 与对称弯曲疲劳极限 σ_{-1} 之间的关系

对钢来说，可以用下列公式来估算对称弯曲疲劳极限 σ_{-1}。

$$\sigma_{-1}=0.285(\sigma_s+\sigma_b) \qquad (5.3\text{-}65)$$

式中，σ_s 为钢的屈服强度；σ_b 为抗拉强度极限。

对铸钢、可锻铸铁和铜合金有

$$\sigma_{-1}=(0.3\sim0.4)\sigma_b$$

对灰口铸铁有

$$\sigma_{-1}=(0.4\sim0.55)\sigma_b$$

需要注意的是，上述估算公式是通过统计得到的近似关系，因此它们只可以作为粗略的分析和大致的强度核算，不能作为精确的定量计算的依据。

2) 过载荷持久值和过载荷损害界　许多零件经常会在短时高于疲劳极限的情况下工作，如汽车紧急刹车、猛然启动，飞机俯冲拉起时机翼在飞行中受到突风冲击等。这些构件的偶然过载运行对疲劳寿命的影响程度一般是通过载荷损害来衡量的。

材料的过载荷损害则是由实验来确定的。首先求出一条完整的疲劳曲线，找出疲劳极限 σ_{-1}，然后对试样采用任一高于 σ_{-1} 的应力进行疲劳试验，经过一定的循环次数 N 后，再在疲劳应力下运行，观察过载荷是否影响了疲劳寿命（N_0），若不影响寿命，说明过载荷没有造成损害，若寿命缩短，则说明造成了损害。在每个过载荷应力下，经过不同的 N 次循环，寻找开始损伤的周次 a、b、c 点等，连接 a、b、c 点就可得出过载荷疲劳损害界。图中阴影区即为过载荷损害区。若过载荷下的循环次数落入此区，将导致疲劳寿命的缩短。因此该区域越窄，则材料的抗过载荷能力越好。

疲劳曲线上的斜线称为过载荷持久值，它表示在超过疲劳极限的应力下直到断裂所能经受的最大的应力循环周次。此斜线越陡直，则表示在相同的过载荷下能经受的应力循环周次越多，即其过载荷抗力越高。

3) 疲劳缺口敏感度　在实际工作中，由于使用的需要常在构件上钻一些小孔、开键槽和过渡面等。而这些将由可能引起应力集中，疲劳裂纹通常会从这类缺口或几何形状突变处开始。因此可以说由于这些表面缺口的存在，会显著降低材料在交变载荷作用下的疲劳抗力。实践证明，仅由光滑试件的 σ_{-1} 数据已经不能满足工程需要，必须针对带有缺口的试样或构件的疲劳强度进行研究。

为了表征缺口对疲劳抗力的影响，首先可以采用循环应力下的有效应力集中系数 K_f 来表示。

$$(K_f)_\sigma=\sigma_{-1}/\sigma_{-1N}\text{和}(K_f)_\tau=\tau_{-1}/\tau_{-1N} \qquad (5.3\text{-}66)$$

式中，$(K_f)_\sigma$，$(K_f)_\tau$ 分别表示弯曲和扭转交变应力下的有效应力集中系数；σ_{-1} 和 τ_{-1} 分布表示为光滑试样的对称弯曲疲劳极限和对称扭转疲劳极限；σ_{-1N} 和 τ_{-1N} 分布表示为缺口试样的对称弯曲疲劳极限和对称扭转疲劳极限，它们是随零件的几何形状和缺口的几何形状而变化的。

由于工程零件的缺口几何形状千变万化，人们无法通过试验来求出不同的 σ_{-1N} 和 τ_{-1N} 值，因此需要寻找一个能衡量疲劳缺口敏感程度的指标，它仅与被试验材料的性质有关，而与缺口几何形状无关。这个指标可采用疲劳缺口敏感度 q 来表示。

$$q_\sigma=\frac{(K_f)_\sigma-1}{(K_t)_\sigma-1}\text{和 }q_\tau=\frac{(K_f)_\tau-1}{(K_t)_\tau-1} \qquad (5.3\text{-}67)$$

式中两个分子项为材料性质和缺口几何形状的函数；分母项为缺口几何形状的函数，表示 q 与缺口圆角半径 r 的关系。当半径 r 很小时，即缺口很尖锐的时候，K_t 值很大，但 K_f/K_t 的比值反而随缺口的尖锐而减小，致使 q 随 r 的减小而减小。而当缺口的圆角半径超过5.0～7.5 mm时，不同材料有不同的稳定的 q 值，不再与缺口的几何形状有关，因此，q 值即可表征材料疲劳缺口敏感度。

对疲劳缺口敏感度 q 的影响因素除了缺口半径外，还包括绝对尺寸、晶粒尺寸及应力梯度等。金属的疲劳缺口敏感度对疲劳强度的计算和疲劳失效的分析是十分重要的性能参量。

2.3.2　疲劳强度理论及其应用

在工程构件的疲劳设计和疲劳失效分析中，常常需要对构件的疲劳强度进行校核。校核时首先要计算构件在危险点的最大工作应力；然后根据构件的疲劳极限简化图确定构件的疲劳极限，接下来计算构件的疲劳极限与最大工作应力的比值，从而得到工作的安全系数。由于构件所承受的应力情况各有不同，因此疲劳强度理论也有所差别，下面就两种应力情况的疲劳强度理论及其应用进行介绍。

(1) 单向应力下的疲劳强度理论

所谓受单向应力是指零件只承受单向的正应力或切应

力。承受对称循环应力的构件，其疲劳极限 σ_{-1}^{*} 等于材料的疲劳极限 σ_{-1} 或 τ_{-1} 乘以相应的系数，如下式所示，

$$\sigma_{-1}^{*}=\sigma_{-1}\times\frac{\varepsilon_{\sigma}\beta}{(K_{f})_{\sigma}} \tag{5.3-68}$$

式中，ε_{σ} 为尺寸系数，它是光滑大试件的疲劳极限 $(\sigma_{-1})_{d}$ 与同样几何形状的光滑小试样（$d_0\leqslant 10$ mm）的疲劳极限 σ_{-1} 之比，即 $\varepsilon_{\sigma}=\frac{(\sigma_{-1})_{d}}{\sigma_{-1}}$，相应的受扭转的尺寸系数为 ε_{τ}，它们都是小于1的数；β 为加工质量系数，$(K_{f})_{\sigma}$ 为疲劳有效应力集中系数。

对于承受对称拉压循环应力的构件，其实际疲劳极限 σ_{-1p}^{*} 为

$$\sigma_{-1p}^{*}=\sigma_{-1p}\times\frac{\varepsilon_{\sigma}\beta}{(K_{f})\sigma} \tag{5.3-69}$$

对于承受对称扭转循环应力的构件，其实际构件的疲劳极限 τ_{-1}^{*} 为

$$\tau_{-1}^{*}=\tau_{-1}\times\frac{\varepsilon_{\tau}\beta}{(K_{f})_{\tau}} \tag{5.3-70}$$

若采用安全系数形式的强度理论，就要求构件对于疲劳破坏的实际安全系数不小于规定的安全系数，于是有弯曲、拉压和扭转下疲劳强度校核公式为：

$$\text{对称弯曲时，}n_{\sigma}=\frac{\sigma_{-1}}{\sigma_{max}}\geqslant[n] \tag{5.3-71}$$

$$\text{对称拉压时，}n_{\sigma}=\frac{\sigma_{-1p}}{\sigma_{max}}\geqslant[n]$$

$$\text{对称扭转时，}n_{\tau}=\frac{\tau_{-1}}{\tau_{max}}\geqslant[n]$$

式中，n_{σ} 和 n_{τ} 为实际安全系数；$[n]$ 为规定安全系数。以上是对称循环的单向应力下的疲劳强度理论校核条件。下面介绍非对称循环应力下的实际安全系数计算方法。

对于非对称循环拉压或弯曲交变应力下，循环特征值 R 保持不变时，构件的实际安全系数为 n_{σ}，强度判据如下式。

$$n_{\sigma}=\frac{\sigma_{-1}}{\bar{\sigma}_{a}\times\frac{(K_{f})_{\sigma}}{\varepsilon_{\sigma}\beta}+\psi_{\sigma}\bar{\sigma}_{m}}\geqslant[n] \tag{5.3-72}$$

式中，ψ_{σ} 为与材料性质有关的常数；$\bar{\sigma}_{a}$、$\bar{\sigma}_{m}$ 为构件工作应力的应力振幅和平均应力。

对于非对称循环扭转交变应力下，构件的实际安全系数为 n_{τ}，强度判据如下式，

$$n_{\tau}=\frac{\tau_{-1}}{\bar{\tau}_{a}\times\frac{(K_{f})_{\tau}}{\varepsilon_{\sigma}\beta}+\psi_{\tau}\bar{\tau}_{m}}\geqslant[n] \tag{5.3-73}$$

式中，ψ_{τ} 为与材料性质有关的常数；$\bar{\tau}_{a}$、$\bar{\tau}_{m}$ 为构件工作应力的应力振幅和平均应力。

(2) 二向应力（弯曲应力和扭转应力）下的疲劳强度理论

在实际工作中构件在大多数情况下承受的是复合交变应力，如弯曲和扭转复合应力等。一般对承受复合交变应力的构件进行疲劳强度校核时，首先要应用单向应力下的疲劳强度理论，再推出复合应力下的疲劳强度。

在弯曲和扭转联合作用下有

$$\left(\frac{\sigma_{a}}{\sigma_{-1}}\right)^{2}+\left(\frac{\tau_{a}}{\tau_{-1}}\right)^{2}=1 \tag{5.3-74}$$

式中，σ_{a}、τ_{a} 分别为弯曲和扭转应力幅；σ_{-1}、τ_{-1} 分别为对称弯曲和扭转疲劳强度极限。

上式是一个椭圆方程，它与试验数据符合的很好。图5.3-27为由三组典型的疲劳试验数据画出的三个椭圆曲线，代表这三个钢种的弯扭合成的疲劳极限图。弯扭复合应力下的疲劳失效或疲劳强度校核就是根据此图来进行，最终可得弯扭合成的二向应力的安全系数为

$$n=\frac{1}{\sqrt{\left(\frac{1}{n_{\sigma}}\right)^{2}+\left(\frac{1}{n_{\tau}}\right)^{2}}} \tag{5.3-75}$$

式中，n_{σ}，n_{τ} 分布如式（5.3-72）和式（5.3-73）所示。

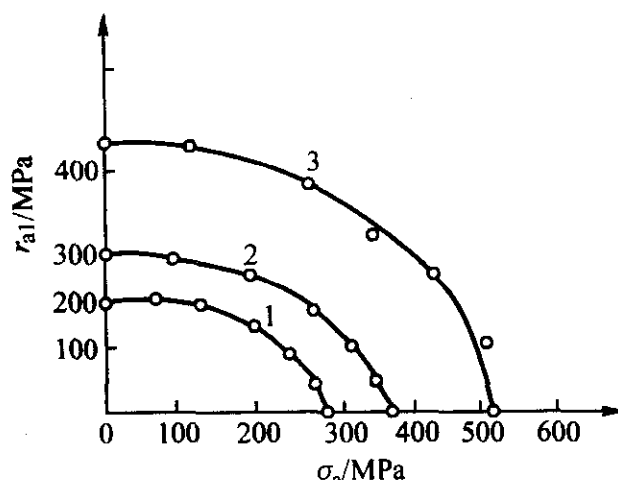


图5.3-27　弯曲扭转合成下，三种钢的疲劳极限
1—0.1%C钢；2—3%Ni钢；3—3.5%Ni-Cr钢

2.3.3　疲劳寿命的估算及其应用

传统的疲劳设计方法是以用光滑试样测得的应力幅——破坏循环次数曲线（即 $\sigma_{a}-N$ 曲线）或虚拟应力幅——破坏循环次数曲线（即 $S_{a}-N$ 曲线）为依据进行疲劳设计的，其中 $\sigma_{a}-N$ 曲线适用于高周疲劳或应力疲劳的场合，而 $S_{a}-N$ 曲线则适用于低周疲劳或应变疲劳的场合。但由于该设计方法没有考虑到构件在制造和使用过程中不可避免地存在有各种缺陷，因此意外事件时有发生。而破损安全设计法恰好弥补了传统疲劳设计的不足。

破损安全设计法认为，任何构件不可避免地存在有裂纹等各种缺陷，在交变载荷的作用下，裂纹将以一定的规律扩展，当裂纹扩展到材料的断裂韧性所允许的构件的临界尺寸时，结构就会产生破坏。根据裂纹扩展规律、构件的裂纹尺寸、材料的断裂韧度以及有关的工作条件就能确定出能保证含裂纹构件安全运行的安全寿命。疲劳寿命的计算对于疲劳设计和疲劳失效分析工作来说是一项十分重要的内容，为了对构件的疲劳寿命进行估算，首先要对疲劳裂纹的扩展规律进行分析。

(1) 疲劳裂纹扩展规律

疲劳裂纹的扩展规律可用疲劳扩展速率图来表示。如图5.3-28所示为一完整的疲劳扩展速率图。图中 da/dN 为裂纹的扩展速率，表示载荷每循环一次所发生的裂纹扩展量，ΔK 表示裂纹尖端应力强度因子变化范围。一般可把疲劳裂纹扩展速率图中扩展的情况分为三个阶段，即阶段Ⅰ，疲劳裂纹萌生与微观缓慢扩展阶段，阶段Ⅱ宏观裂纹扩展阶段（裂纹稳定扩展阶段）和阶段Ⅲ快速扩展阶段。

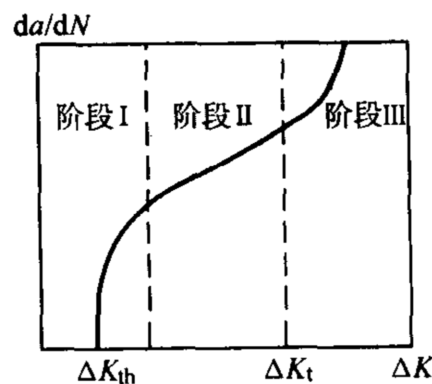


图5.3-28　疲劳扩展速率图

1) 裂纹萌生与微观扩展阶段　如图5.3-28所示，阶段Ⅰ有一个界限值 ΔK_{th}（为疲劳裂纹界限值或称为下门槛

值），低于这个界限值，即 $\Delta K < \Delta K_{th}$ 时，可视为裂纹未扩展，或更确切地说，裂纹扩展量极微小。

在该阶段，裂纹扩展速度比较慢，一般为 10^{-5} mm/r，裂纹尺寸也比较小，通常小于0.05 mm。一般地说，低周疲劳状态，微观裂纹扩展阶段经历的时间占疲劳裂纹总寿命的比例小，而高周疲劳状态，则所占比例较大。因此对高周疲劳破坏问题，测定材料的下门槛值 ΔK_{th} 就有重要意义。因为高周疲劳状态，宏观裂纹一旦形成，仅经历较短时间的宏观裂纹扩展阶段就快速扩展发生断裂。

2）宏观裂纹扩展阶段　该阶段一般指裂纹尺寸从0.05 mm扩展到临界尺寸 a_c 为止，其扩展速度比微观裂纹扩展阶段要大，约为 10^{-3} mm/r。

微观裂纹扩展阶段和宏观裂纹扩展阶段都属于疲劳裂纹的亚临界扩展阶段。对于高强度材料，由于屈服强度高，缺口敏感性大以及内部夹杂和硬颗粒多，因此往往先沿着夹杂物的界面裂开，一开始就进入宏观裂纹扩展阶段而没有成核阶段和微观阶段。

亚临界扩展阶段的理论分析中，Paris应力强度因子理论是目前与实验结果符合最好的一种理论。Paris认为应力强度因子 K 是描述裂纹端部应力、应变场强度的参量，那么就有理由认为，K 值也是控制裂纹扩展速率 $\frac{da}{dN}$ 的主要参量。因此他提出一个半经验公式：

$$\frac{da}{dN} = C(\Delta K)^m \tag{5.3-76}$$

式中，$\frac{da}{dN}$ 为对应于一次循环的裂纹扩展量；ΔK 为载荷循环一次，疲劳裂纹尖端应力强度因子的变化幅度，$\Delta K = K_{max} - K_{min}$；$C$、$m$ 为材料常数。对于金属材料，其 $m = 2 \sim 7$；而绝大多数金属材料的 $m = 2 \sim 4$。

几种材料的 da/dN 与 ΔK 的关系曲线见图5.3-29。

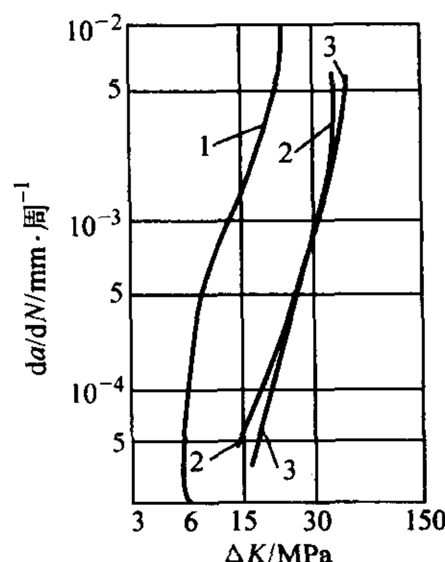


图5.3-29　几种材料的 da/dN 与 ΔK 之间的关系曲线

1—2024-T4；2—Q255-A；3—45钢

由于裂纹扩展具有很多复杂的因素，它们超出了Paris公式所能描述的简单行为。如所考虑的载荷历史是不能由均匀性中导出的；另外若 $\Delta K < \Delta K_{th}$，裂纹将不扩展，而在另一个极端的情况下，将会出现快速扩展。因此可用一些经验公式来考虑一个或多个复杂的因素。如，Donahue等人提出了下式，

$$\frac{da}{dN} = c(\Delta K - \Delta K_{th})^n \tag{5.3-77}$$

就考虑了门槛值的情况。

3）快速扩展阶段　在宏观裂纹扩展阶段后继续加载，裂纹进一步扩展。当裂纹扩展到临界裂纹尺寸 a_c 时，即发生失稳断裂，即疲劳扩展速率图的第三部分。此时 $\Delta K > \Delta K_t$（对应于疲劳扩展曲线的第二个转折点）。变一阶段裂纹扩展速率主要取决于 K_{max}，它对于平均应力、微观组织结构较敏感。

Barson根据实验结果指出，对于断裂韧性好的材料，即当材料的 $K_{Ic} > \sqrt{0.041\sigma_y E}$ 时，K_t 为式（5.3-78），单位是 $MPa \cdot mm^{1/2}$。

$$K_t = \sqrt{0.041\sigma_y E} \tag{5.3-78}$$

对于断裂韧性差的材料，即当材料 $K_{Ic} < \sqrt{0.041\sigma_y E}$ 时，则以 K_{Ic} 作为 K_t。K_t 为快速疲劳扩展的界限值（上门槛值）或称为疲劳断裂韧性。

（2）影响裂纹扩展的因素

应力强度因子幅度 ΔK 是决定疲劳扩展的主要参量，是疲裂纹扩展的原动力。但还有很多复杂的因素对疲劳裂纹扩展也有较大影响，如平均应力、过载峰、加载频率、温度、环境介质等。这里简单介绍几种主要因素对疲劳裂纹扩展的影响。

平均应力的影响：

平均应力的大小可以用循环特性 R 来表示。E.K.Walker考虑了平均应力的影响建议用下式来代替Paris公式，

$$\frac{da}{dN} = c[K_{max}(1-R)^m]^n \tag{5.3-79}$$

式中，c，m，n 皆为材料及介质有关的常数。对于2024-T3铝合金，$m = 0.5$；7075-T6铝合金 $m = 0.425$；对于不锈钢，$m = 0.667$。$[K_{max}(1-R)^m]$ 称为有效应力强度因子幅度，可以用符号 ΔK_e 来表示。

理论分析和实验结果表明，构件表面的残余压应力会使交变载荷的平均应力降低，而表面的残余拉应力则会使平均应力提高。因此，为了降低构件的 $\frac{da}{dN}$，可以采取某些特殊的工艺措施。

过载峰的影响如下。

实际构件所承受的交变载荷，一般都不是单一的、幅度不变的恒幅载荷，而是由各种载荷组成的载荷谱。试验证明，构件在载荷谱作用下的寿命是不同的，相邻的不同幅度的载荷循环之间有很大的相互影响作用。实验结果表明过载峰对疲劳裂纹扩展有延缓或停滞作用。此种延缓效应要持续数千次甚至数万次循环，随后，裂纹的扩展速率又会慢慢恢复至相应裂纹长度下的正常扩展速率。关于过载对疲劳裂纹扩展的延缓作用机理至今尚无完善的解答，曾有几种解释模型，如O.E.Wheeler模型即是其中之一。该模型认为过载峰使裂纹尖端产生比在正常的恒幅载荷更大的塑性区，从而增加了裂纹的扩展阻力，并且使随后在恒幅载荷下的疲劳裂纹扩展速率降低。随着裂纹的继续扩展，一直到穿越过这个在塑性区后，过载峰的延缓作用才消失。

加载频率的影响：

总的说来，在 ΔK 值较低时，加载频率对疲劳裂纹扩展速率的影响很小；但 ΔK 值较大时，特别是在高温下，由于频率与蠕变的交互作用，加载频率对 $\frac{da}{dN}$ 就有明显的影响。随着频率的降低，裂纹扩展速率将有所增大。

关于其他因素对疲劳裂纹扩展的影响此处就不再详述了。

（3）疲劳寿命的估算及其应用

通常根据工作条件的不同，可以将疲劳寿命的估算方法分为三类：高周疲劳寿命的估算方法（包括Basguin Miner累计疲劳寿命损伤理论和用疲劳寿命与应力幅之间的关系来估算）、低周疲劳寿命估算（用Manson-Coffin公式）及断裂力学疲劳寿命的估算（Paris公式）。下面对这几种寿命估算方法予以分别介绍。

1）恒幅载荷下的疲劳裂纹扩展寿命的计算 由式（5.3-76）有

$$dN=\frac{da}{c(\Delta K)^n}$$

$$N=\int_{a_0}^{a_c}\frac{da}{c(\Delta K)^n} \tag{5.3-80}$$

式中，a_0 为初始裂纹尺寸；a_c 为临界裂纹尺寸；ΔK 为应力强度因子幅度，$\Delta K=\gamma\Delta\sigma\sqrt{\pi a}$；$\gamma$ 为裂纹结构的几何形状因子。在很多情况下，γ 与 a 有关。

2）Miner 累积疲劳损伤理论 Miner 累积疲劳损伤理论认为当材料所承受的载荷高于疲劳极限时，将对材料产生一定的损伤，这种损伤是能累积的，当损伤累积到临界值时，构件就会发生破坏。

假设试件所承受的应力为 $\pm\sigma_1$，在该应力水平下达到破坏的总寿命为 N_1，若其循环数为 n_1 则造成的损伤为 n_1/N_1。同理，若设试件所承受的应力为 $\pm\sigma_2$，在该应力水平下达到破坏的总寿命为 N_2，若其循环数为 n_2 则造成的损伤为 n_2/N_2，依此类推可有（如图 5.3-30 所示），

$$\frac{n_1}{N_1}+\frac{n_2}{N_2}+\cdots+\frac{n_t}{N_t}=1 \text{ 或} \sum\frac{n_i}{N}=1 \tag{5.3-81}$$

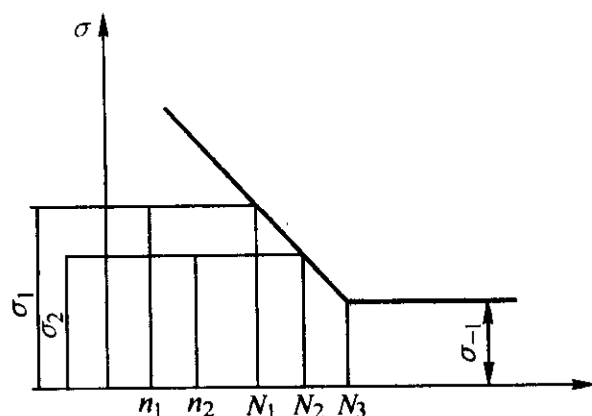


图 5.3-30 累积疲劳损伤理论

该式称为 Miner 公式，或 Miner 定理。在 Miner 公式中没有考虑应力作用的先后顺序对寿命的影响，但实际上，在低应力到高应力的试验中，$\sum\frac{n_i}{N_i}\geqslant 1$，而在高应力到低应力的试验中，$\sum\frac{n_i}{N_i}\leqslant 1$，因此 Miner 公式可认为是一个近似的表达式。

3）Basquin 疲劳寿命与振幅之间的关系及其在疲劳寿命中的应用 图 5.3-30 所示的曲线为双对数曲线，在曲线上有限寿命部分（倾斜部分）为一倾斜的直线，其方程为

$$\lg N+m\lg\sigma=\lg C \text{ 或} \sigma^m N=C=\text{常数} \tag{5.3-82}$$

式中，m 为由材料和加载方式等确定的常数；C 为常数。

若已知应力 σ_1 下的疲劳寿命为 N_1，应力 σ_2 下的疲劳寿命为 N_2，则有

$$\sigma_1^m N_1=\sigma_2^m N_2 \tag{5.3-83}$$

那么在 σ_1 下循环 N_1 次后，在 σ_2 下的剩余寿命 $n_{剩}$ 为

$$n_{剩}=N_2-\left(\frac{\sigma_1}{\sigma_2}\right)^m n_1$$

依此类推，经过（$i-1$）级加载后的使用寿命为

$$n_{剩}=N_i-\left\{N_{i-1}-\left[N_{i-2}-(\cdots)\left(\frac{\sigma_{i-3}}{\sigma_{i-2}}\right)^m\right]\left(\frac{\sigma_{i-2}}{\sigma_{i-1}}\right)^m\right\}\left(\frac{\sigma_{i-1}}{\sigma}\right)^m \tag{5.3-84}$$

式中，m 值对抛光表面为 9～12，对一般加工表面为 6～10，对经过表面强化的表面为 18～20。

4）Manson－Coffin 公式及其在寿命估算中的应用 低周疲劳的塑性应变幅 $\Delta\varepsilon_p$ 与疲劳寿命之间的关系可以写成，如图 5.3-31 所示。

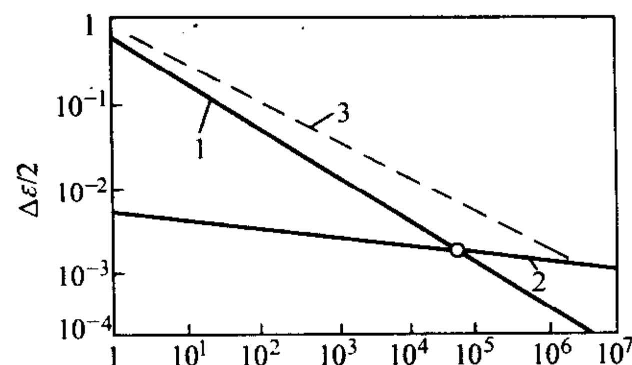


图 5.3-31 对称循环全应变幅与循环数之间的关系
1—$\Delta\varepsilon_p/2-N_f$；2—$\sigma_a-N_f$；3—$\Delta\varepsilon/2-N_f$

$$\Delta\varepsilon_p N^\alpha=C=\frac{\varepsilon_f}{2} \tag{5.3-85}$$

上式称为 Manson－Coffin 定理，其中 α 为材料的塑性指数，一般为 0.3～0.8，C 为在静拉伸断裂真实应变 ε_f 与 $\varepsilon_f/2$ 之间的一个常数，ε_f 与断面收缩率 ψ 之间有如下关系，

$$\varepsilon_f=\ln\frac{100}{100-\psi} \tag{5.3-86}$$

即有若已知 $\Delta\varepsilon_{p1}$，N_1，$\Delta\varepsilon_{p2}$ 则可根据式（5.3-85）求出 N_2，即

$$N_2=\left(\frac{\Delta\varepsilon_{p1}}{\Delta\varepsilon_{p2}}\right)^{\frac{1}{\alpha}}N_1 \tag{5.3-87}$$

$$N_{剩}=\Delta\varepsilon_{p2}^{\alpha}\left(C\frac{1}{\alpha}-\Delta\varepsilon_{p1}^{\frac{1}{\alpha}}n_1\right) \tag{5.3-88}$$

2.4 以损伤力学为基础的失效预测技术和方法

2.4.1 损伤力学基础

（1）损伤力学的基本概念

损伤是材料结构组织在外界因素作用下发生的力学性能劣化并导致体积单元破坏的现象。材料的损伤就是使材料损坏的渐进的物理过程。损伤力学则是研究在各种加载和环境条件下，物体中损伤随着变形或时间而发展导致破坏的规律的学科。它是一门较系统地研究微缺陷以及这些缺陷发展对应力和应变状态影响的科学。

损伤力学和断裂力学是密不可分的。通常认为损伤力学研究微缺陷及其发展对应力和应变的影响，断裂力学研究宏观裂纹对应力和应变状态的影响。而微缺陷是宏观缺陷的前期表现。

损伤力学的研究可分为宏观方法和微（细）观方法。宏观方法即连续损伤力学，它将材料中的应力、应变以及损伤及其分布都采用连续场变量来描述，即采用连续体力学的方法来进行分析。在分析中需要采用力学方法确定损伤参数，作为内变量来建立本构关系。微细观方法即微细观损伤力学，它需要采用单元模型把非均质的微观组织性能转化微宏观材料的性能，建立分析理论。近来，同时采用这两种方法进行损伤分析的情况逐渐增多，且损伤力学与断裂力学的衔接也日趋密切，这必将推动损伤力学的发展。

损伤的发展通常伴随着一定的变形，若按变形情况对损伤现象进行分类可以得出以下结果：

1）脆性损伤，如对脆性材料单调加载时，没有出现明显的不可逆变形；

2）塑性损伤，如在常温或高温情况下，金属的变形；

3）蠕变损伤，它相对于金属在中等或高温条件下的晶间分离；

4）疲劳损伤，指在循环载荷下的损伤等。

当然还有其他形式的损伤，如烧蚀、腐蚀、氧化和辐照等。

（2）损伤变量

在连续损伤力学中，材料的损伤情况是通过“基本单

元”来分析的。“基本单元”经受了弹性、塑性、黏性以致微裂的演变过程，即可由微裂单元聚集成宏观可见的裂纹。下面就通过基本单元体定义可以描述损伤演变的变量，损伤变量。如图 5.3-32 所示，为已受损物体。现把物体剖开，并取出一材料的构元，其尺度等同于前面所述的连续介质力学中基本单元体，且带有一定的损伤。

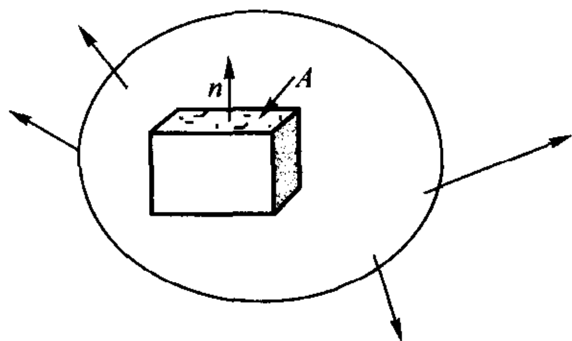


图 5.3-32　受损物体及其中的单元体

若以 A 表示通过单元体外法线 n 的截面的原有面积，$\tilde{A}$ 为受损后的有效面积，则可以将 n 方向上的损伤变量定义为

$$D_n = \frac{A - \tilde{A}}{A} = 1 - \psi \tag{5.3-89}$$

式中，$\psi = \frac{\tilde{A}}{A}$，它也是损伤变量的一种表达方式。

有效应力 $\tilde{\sigma}$ 与名义应力 σ 的关系为

$$\tilde{\sigma} = \frac{\sigma}{1 - D_n} \tag{5.3-90}$$

在通常情况下，物体材料的损伤（微裂纹、孔洞等）是有方向性的。当损伤变量 D_n 与法线 n 相关时，损伤为各向异性损伤；反之，若 D_n 于法线无关，则为各向同性损伤，此时损伤变量为 D，它是一个标量的指标。

受损材料的本构关系在外力的作用下可以采用无损情况下的形式来表示，只要把其中的名义应力换成有效应力即可，这一假设被称为应变等价性假设。根据这一假设有

$$\varepsilon = \frac{\tilde{\sigma}}{E} = \frac{\sigma}{E(1-D)} = \frac{\sigma}{\tilde{E}} \tag{5.3-91}$$

式中，ε 为损伤弹性应变；$\tilde{E}$ 为损伤弹性模量，$\tilde{E} = E(1-D)$；E 为无损时材料的弹性模量。

若进而定义

$$\tilde{\varepsilon} = \varepsilon(1 - D) \tag{5.3-92}$$

式中，$\tilde{\varepsilon}$ 为有效应变。由式（5.3-91）、式（5.3-92）可得

$$\sigma = E\tilde{\varepsilon} \tag{5.3-93}$$

式（5.3-93）称为应力等价性假设，即在外应力下，受损材料的本构关系可采用无损时的形式，只需将其中的应变 ε 换成 $\tilde{\varepsilon}$ 即可。

(3) 损伤的测量方法

为了满足建立损伤本构方程的需要，必须得到一些反映材料损伤性质的常数，而这些常数要能对已经破损的构件的损伤成因提供依据。测量损伤较直接的方法是采用金相学方法直接测定材料中的缺陷，包括位错、空洞等。声发射、声衰减、红外紫外摄像、X射线等都是检验损伤的手段，但这些微观缺陷检测，虽然可以获得直接信息，但难于直接将其结果应用于损伤力学，需作一些宏观尺度下的统计处理。通常损伤的测量是十分困难的，这里简单介绍几种损伤测量的方法。

1）量密度测量法　若损伤是在体内孕育和发展的空洞，则随着损伤的增加，材料在质量不变的情况下，会因为体积的增加而使材料的密度下降，所以可以采用密度的下降来计算损伤情况。如采用 ρ 表示无损伤情况下的材料的密度，用 $\tilde{\rho}$ 表示损伤状态的密度，则可以用它们之间的相对变化来表示表面损伤，即有

$$D = \frac{\tilde{\rho} - \rho}{\rho} \tag{5.3-94}$$

通过上式即可实现由材料受力时的密度变化来表示损伤的变化。

2）弹性模量变化法　由等效应变的概念

$$\varepsilon^{e} = \frac{\sigma}{E(1-D)} = \frac{\sigma}{\tilde{E}} \tag{5.3-95}$$

式中，E 为弹性模量；$\tilde{E}$ 为受损材料的弹性模量或有效弹性模量。由式（5.3-95）可得

$$D = 1 - \frac{\tilde{E}}{E} \tag{5.3-96}$$

只要所测区域内的变形是均匀的即可采用此种方法。它可以用来测量延性塑性、蠕变和疲劳损伤。

3）电阻率测量法　通过电压差的相对变化 $\Delta dV/dV$ 和弹塑性变形来确定损伤的方法称之为电阻率测量法。零件中的电压差可表示为

$$dV = \eta \frac{dl}{dA} dI \tag{5.3-97}$$

式中，η 为电阻率；dl 和 dA 为柱形元素的长度和横截面积；dI 是流经该段元素的电流。

受损材料的有效电流可表示为

$$\tilde{I} = \frac{I}{1 - D} \tag{5.3-98}$$

对于电流的测量，微裂纹比微孔洞要敏感一些。疲劳损伤一般是由微裂纹引起的，因此对于疲劳损伤的测量采用电学的方法和采用力学的方法结果是相近的，而对于蠕变损伤来说，由于它主要是由孔洞造成的，因此用电学的方法测得的损伤量比用力学方法测量的结果一般要小一些。

除了以上的几种测量方法外还由许多其他的测量方法，包括塑性特征测量法、黏塑性特征测量法、超声波法、第三阶段蠕变响应法、微观硬度法等，这里不再作一一介绍，在进行测量时可以参考一些相关文献进行检测。

2.4.2　脆（弹）性损伤失效分析技术和方法

若一个细观裂纹萌生，没有发现宏观塑性应变，则此时的损伤称为脆（弹）性损伤。若要给出一个大小的量级，可以说塑性应变与弹性应变之比小于单位值，即

$$\frac{\varepsilon_p}{\varepsilon_e} < 1$$

这就意味着解理力小于产生滑移的力，且大于脱键力，同时亦意味着损伤局部程度很高。

要进行损伤分析首先要计算损伤率，损伤率可以用应力表示，也可以用应变来表示。

(1) 应力表示的损伤率

损伤率可以简单的认为损伤变量与有效正应力 $\tilde{\sigma}$ 成正比，即有

$$D = \begin{cases} \tilde{\sigma}/K & \tilde{\sigma} > 0 \\ 0 & \tilde{\sigma} < 0 \end{cases} \tag{5.3-99}$$

式中，K 为损伤模量，其量纲与应力的量纲相同，可以通过试验测定来获得。

由式（5.3-90）、式（5.3-93）可得

$$\sigma = (1 - D)\tilde{\sigma} = (1 - D)E\varepsilon \tag{5.3-100}$$

由式（5.3-99）、式（5.3-100）有

$$\sigma = E\varepsilon\left(1 - \frac{E\varepsilon}{K}\right) \tag{5.3-101}$$

当 $K = \infty$ 时，材料（或零件）没有损伤，且 $D = 0$，则式（5.3-101）可以简化为通常的弹性本构关系，当损伤增大到以下条件时认为材料（或零件）失效破坏。

$$\frac{d\sigma}{d\varepsilon} = 0 \tag{5.3-102}$$

材料丧失承载能力，此时有 $D=D_c$，其中 D_c 为临界损伤变量。由式（5.3-101）、式（5.3-102）即可推导出极限 Cauchy 应力

$$\sigma_c = K/4 \tag{5.3-103}$$

由 $\tilde{\sigma}=DK$ 及式（5.3-100）、式（5.3-15）可以推导出临界损伤变量 D_c

$$D_c^2 - D_c + \frac{\sigma_c}{K} = 0 \tag{5.3-104}$$

因此有

$$D_c = 1/2,\ \tilde{\sigma}_c = \frac{K}{2} \tag{5.3-105}$$

式中 $\tilde{\sigma}_c = \dfrac{\sigma_c}{1-D_c}$ 称为极限有效应力。根据上式即可建立结构损伤的判别条件，也可根据载荷的情况确定其损伤的程度。

(2) 应变表示的损伤率

在分析脆性材料的损伤时，人们习惯上采用零件在外载作用下产生的应变来进行判断，且主要采用拉伸应变来确定材料的损伤，如对陶瓷等脆性材料即采用应变准则。下面介绍一个简单的脆性损伤模型，即 Lemaitre 等人提出的模型，

$$dD = \begin{cases} \left(\dfrac{\varepsilon}{\varepsilon_0}\right)^s d\varepsilon & 当\ \varepsilon=\zeta, d\varepsilon = d\zeta > 0 \\ 0 & 当\ \varepsilon < \zeta, 或\ d\varepsilon < 0 \end{cases} \tag{5.3-106}$$

式中，ε_0 和 S 为表征损伤能量强度的材料常数；ζ 为一可变的门槛值，当 $\varepsilon=\zeta$，$d\varepsilon = d\zeta$。

假设初始时，没有损伤的条件为

$$D = \varepsilon = \zeta = 0$$

则可以对式（5.3-106）进行积分，有

$$D = \frac{1}{S+1}\left(\frac{\varepsilon}{\varepsilon_0}\right)^{S+1} = \left(\frac{\varepsilon}{\varepsilon_R}\right)^{S+1} \tag{5.3-107}$$

式中 $\varepsilon_R = [(S+1)\varepsilon_0^{S+1}]^{\frac{1}{S+1}}$，为 $D=1$ 时的破坏应变，由式（5.3-100）最终可得

$$\sigma = E\varepsilon\left[1-\left(\frac{\varepsilon}{\varepsilon_R}\right)^{S+1}\right] \tag{5.3-108}$$

上式说明当损伤构元的应变等于其破坏应变时，材料即不能承受任何载荷。

2.4.3 延性损伤失效分析技术和方法

若损伤发生时，塑性变形大于某一个门槛值 P_D，则此时的损伤称为延性损伤。它是由于夹杂物的存在以及其导致的空洞的存在所引起的，这些空洞由于塑性不十分稳定，导致其不断的增长和合并最终形成损伤。因此，延性损伤的局部化程度与塑性应变的局部化程度是可比的。

为了分析延性损伤，首先要考虑两种应变硬化，一是与位错密度或流动受阻有关的各向同性硬化；另一个是与材料内部微应力集中状态一个的运动硬化（相应的背应力定义了拉压状态下弹性域的中心）。若假设 σ_y 为屈服应力，R 为由各向同性硬化引起的应力，X 为背应力，且 R 和 X 均为塑性应变的函数，则定义屈服极限当前门槛值的一维塑性判据为

$$\sigma = \sigma_y + R + X \quad 或 \quad f = |\sigma - X| - R - \sigma_y = 0 \tag{5.3-109}$$

式中，f 为屈服函数。

发生损伤时，由等价原理屈服函数 f 需由下式来表示

$$f = \left|\frac{\sigma}{1-D} - X\right| - R - \sigma_y = 0 \tag{5.3-110}$$

另外，总应变可写成

$$\varepsilon = \varepsilon_e + \varepsilon_p \tag{5.3-111}$$

塑性应变动力本构方程可以由式（5.3-109）、式（5.3-110）、式（5.3-111）导出

$$若\begin{cases} f=0 \\ \dot{f}=0 \end{cases} 则\ \dot{\varepsilon}_p \neq 0; \quad 若\begin{cases} f<0 \\ \dot{f}<0 \end{cases} 则\ \dot{\varepsilon}_p = 0$$

通过实验结果和方程式（5.3-110）可得

$$\sigma = (\sigma_y + R + X)(1-D) \tag{5.3-112}$$

上式表明，损伤即等同于减少了屈服应力、各向同性应变硬化应力以及背应力。

如前所述，当累积塑性应变 p 达到一定的门槛值 p_D 时，损伤才会发生，即有

$$p < p_D \rightarrow \dot{D} = 0\ 以及\ \varepsilon_p < \varepsilon_{pD} \rightarrow \dot{D} = 0$$

对于完全塑性损伤的情况来说，可以认为应变硬化接近于饱和，此时动力硬化和各向同性硬化也随之饱和，即有

$$p > p_D \rightarrow \begin{cases} \dot{X}^D = 0 \\ \dot{R} = 0 \end{cases}$$

因此塑性判据式（5.3-109）可简化为

$$f = \tilde{\sigma}_t - \sigma_s = 0$$

式中，σ_t 为 VonMises 等效应力；σ_s 为塑性应力门槛值。完全塑性应变和损伤率可由下式来描述

$$当\begin{cases} f = \dfrac{\sigma_t}{1-D} - \sigma_s = 0 \\ \dot{f} = 0 \end{cases} 时，有$$

$$\dot{\varepsilon}^p = \frac{\partial f}{\partial \sigma}\dot{p}(1-D) \tag{5.3-113}$$

当 $p > p_D$ 时，有

$$\dot{D} = \frac{\sigma_s^2 R_v}{2ES(1-D)^2}\dot{p} \tag{5.3-114}$$

式中，$R_v = \dfrac{2}{3}(1+\nu) + \dfrac{3}{1-2\nu}\left(\dfrac{E\varepsilon_H}{\sigma_s}\right)$，$\varepsilon_H$ 为流体静压应变。且其损伤为

$$D = \frac{\sigma_s^2 R_v}{2ES(1-D)^2}(p - p_D) \tag{5.3-115}$$

对于黏塑性损伤，损伤的门槛值为

$$p_D = \varepsilon_{pd}\frac{\sigma_v - \sigma_f}{\sigma_t - \sigma_f} \tag{5.3-116}$$

式中，σ_v 为拉伸应力应变曲线上应力的最大值；σ_f 为疲劳极限应力。

其损伤度亦可由式（5.3-115）来表示。

2.4.4 蠕变损伤失效分析技术和方法

一般在高温下，金属承受载荷时，比如温度高于其熔点的 1/3 时，塑性应变中包含了黏性，当应变足够大时，则产生沿晶开裂，从而引起损伤，并通过蠕变第三阶段使应变率有所增长。蠕变损伤的梯度与黏－塑性应变梯度相似。

一般可以把蠕变曲线（见图 5.3-33）分成三个部分，第一部分（AB 段）为初始部分或第一阶段，该段是变形率从较大的起始值逐渐减小，并趋于稳定状态；第二部分（BC 段）是稳定阶段，这时的应变率接近常数；第三部分（CD 段）为加速段，此时变形加速增长直至破坏。

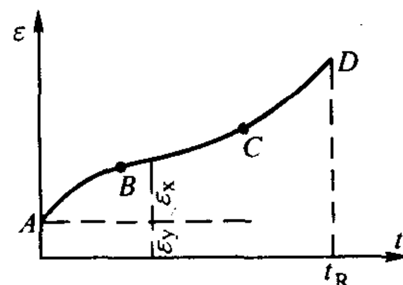


图 5.3-33　蠕变曲线

蠕变损伤率同黏塑性损伤率相似，可由下式表示

$$D = \frac{Y}{S}\dot{p}H(p - p_D) \tag{5.3-117}$$

式中，$Y = \dfrac{\sigma_t^2 R_v}{2E(1-D)^2}$ 为应变能密度释放率，材料参数 S 取

决于温度。

在弯曲黏塑性情况下，即物体应变硬化和屈服应力的状态下，有

$$\dot{p}=\frac{\dot{\lambda}}{1-D}=\left[\frac{\sigma_t}{K_v(1-D)}\right]^N \tag{5.3-118}$$

式中，K_v 和 N 均为材料参数，可以在一些相关的资料中查到，$\dot{\lambda}$ 为塑性乘数。由式（5.3-118）、式（5.3-119）可得

$$\dot{D}=\frac{\sigma_t^2 R_v}{2ES(1-D)^2}\left[\frac{\sigma_t}{K_v(1-D)}\right]^N H(p-p_D) \tag{5.3-119}$$

上式可以简化为，

$$\dot{D}=\left[\frac{\sigma}{A(1-D)}\right]^{N+2} H(p-p_D) \tag{5.3-120}$$

式中$A=(2ESK_v^N)^{1/(N+2)}$，此时可认为 $\sigma=$常数。由于损伤演化是时间的函数，因此下面分析一下损伤随时间的变化情况。

假设 t^* 为达到蠕变 $\epsilon_p=p_D$ 尚未产生任何损伤所需要的时间，则有

$$\mathrm{d}D=\left(\frac{\sigma}{A}\right)^{(N+2)}(1-D)^{-(N+2)}\mathrm{d}t \tag{5.3-121}$$

由式（5.3-121）可推导出蠕变损伤随时间的变化情况，

$$D=1-\left[1-(N+3)\left(\frac{\sigma}{A}\right)^{N+2}(t-t^*)\right]^{1/(N+3)} \tag{5.3-122}$$

2.4.5 疲劳损伤失效分析技术和方法

在疲劳损伤中，又可以分为低周疲劳损伤和高周疲劳损伤。当构件承受较大的应力或应变循环载荷时，在微裂纹形核和扩展前（或潜伏期），损伤随着循环的塑性应变一起发展，此时损伤的局部变化程度明显高于前面所说的塑性（延性）或蠕变损伤。对于金属来说，在滑移带受到抑制后，损伤可能表现为沿晶或穿晶开裂。

疲劳失效的循环数应为形成微裂纹达到 P_D（塑性应变的损伤门槛值）的循环数 N_0 以及损伤发生到形成细观尺度的裂纹的循环次数 N_D 之和，即

$$N_R=N_0+N_D$$

式中，N_R 为断裂循环数。

N_0 必须通过经典的弹塑性本构方程计算，在此期间 $D=0$。考虑到应变为恒幅

$$\Delta\epsilon=\text{常数}\rightarrow\Delta\sigma=\text{常数} \tag{5.3-123}$$

且 $\Delta\epsilon=\Delta\epsilon_e+\Delta\epsilon_p$，以及 $\Delta\epsilon_e=\Delta\sigma/E$ 可得 $\Delta\epsilon_p=$常数，其中 $\Delta\epsilon_e$、$\Delta\epsilon_p$ 为应变幅的弹性部分和塑性部分。则一个循环后的累积塑性应变为

$$\frac{\mathrm{d}p}{\mathrm{d}N}=\int_{\text{一次循环}}|\dot{\epsilon}_p|\,\mathrm{d}t=2\Delta\epsilon_p \tag{5.3-124}$$

式中，p 为累积塑性应变。

则对 N 个循环有

$$p(N)=2N\Delta\epsilon_p \tag{5.3-125}$$

$$N_0=\frac{p_D}{2\Delta\epsilon_p} \tag{5.3-126}$$

若要计算 N_D 需要对一次循环内的动力损伤率积分，然后进行 N 次循环内的积分，即

$$\dot{D}=\frac{\sigma^2}{2ES(1-D)^2}|\dot{\epsilon}_p| \tag{5.3-127}$$

对每次循环的损伤计算有

$$\frac{\mathrm{d}D}{\mathrm{d}N}=\int_{\text{一次循环}}^{\dot{D}}\mathrm{d}t=\frac{\left(\frac{\Delta\sigma}{2}\right)^2}{2ES(1-D)^2}2\int_0^{\Delta\epsilon_p}\epsilon_p\mathrm{d}t=\frac{(\Delta\sigma)^2}{4ES(1-D)^2}\Delta\epsilon_p \tag{5.3-128}$$

若 $\Delta\epsilon_p$ 和 $\Delta\sigma$ 之间存在周期关系，则有

$$\Delta\epsilon_p=\left[\frac{\Delta\sigma}{K_p(1-D)}\right]^M \tag{5.3-129}$$

式中，K_p 和 M 为材料参数，可以通过查阅一些表格得到。由式（5.3-115）有

$$\frac{\mathrm{d}D}{\mathrm{d}N}=\frac{(K_p\Delta\epsilon_p^{1/M})^2}{4ES}\Delta\epsilon_p=\frac{(K_p)^2}{4ES}\Delta\epsilon_p\Delta\epsilon_p^{(M+2)/M} \tag{5.3-130}$$

对式（5.3-130）积分即可得到 N 次循环的损伤 D。

$$D=\frac{(K_p)^2}{4ES}\Delta\epsilon_p\Delta\epsilon_p^{(M+2)/M}(N-N_0) \tag{5.3-131}$$

当 $D=Dc$，且 $N_0=\frac{p_D}{2\Delta\epsilon_p}$ 时有

$$D=\frac{(K_p)^2}{4ES}\Delta\epsilon_p\Delta\epsilon_p^{(M+2)/M}\left(N_R-\frac{p_D}{2\Delta\epsilon_p}\right)$$

$$N_R=\frac{p_D}{2\Delta\epsilon_p}+\frac{4ESD_c}{K_p^2}\Delta\epsilon_p^{-(M+2)M} \tag{5.3-132}$$

式中，N_R 即为达到失效时的循环次数。

当构件承受应力幅值较低的循环载荷时，细观塑性应变通常可以忽略不计，而在微观等级上，个别点或区域的塑性变形可能较高，在这些点处只在一些平面上会出现一些沿晶开裂，最终失效的循环次数很高，此时发生在平面上的损伤称为高周疲劳损伤。

同上述低周疲劳的情形相似，这里首先分析细观尺度时拉压疲劳中的一个周期的情况。首先有 $\Delta\sigma=$常数，$|\dot{\epsilon}|=\frac{|\dot{\sigma}|}{E}$ 则有

$$\dot{D}=\frac{\sigma_f^2}{2E^2S}\left[\frac{2}{3}(1+\nu)+\frac{1}{3}(1-2\nu)\left(\frac{\sigma}{\sigma_f}\right)^2\right]|\dot{\sigma}| \tag{5.3-133}$$

一次循环的损伤为

$$\begin{aligned}\frac{\mathrm{d}D}{\mathrm{d}N}&=\frac{\sigma_f^2}{2E^2S}\int_0^{\frac{\Delta\sigma}{2}}\left[\frac{2}{3}(1+\nu)+\frac{1}{3}(1-2\nu)\left(\frac{\sigma}{\sigma_f}\right)^2\right]\mathrm{d}\sigma\\&=\frac{2\sigma_f^2}{E^2S}\left[\frac{2}{3}(1+\nu)\frac{\Delta\sigma}{2}+\frac{1}{9\sigma_f^2}(1-2\nu)\left(\frac{\sigma}{2}\right)^2\right]\end{aligned} \tag{5.3-134}$$

N 次循环后的损伤为

$$D=\frac{2\sigma_f^2}{E^2S}\left[\frac{2}{3}(1+\nu)\frac{\Delta\sigma}{2}+\frac{1}{9\sigma_f^2}(1-2\nu)\left(\frac{\sigma}{2}\right)^2\right](N-N_0) \tag{5.3-135}$$

式中 N_0 为 $\epsilon_{eq}=p_d$ 时，即当量总应变达到塑性变形门槛值时的循环次数。

$$N_0=\frac{p_D}{2\Delta\epsilon}$$

当 $D=D_c$ 时，有

$$D_c=\frac{2\sigma_f^2}{E^2S}\left[\frac{2}{3}(1+\nu)\frac{\Delta\sigma}{2}+\frac{1}{9\sigma_f^2}(1-2\nu)\left(\frac{\sigma}{2}\right)^3\right](N_R-N_0) \tag{5.3-136}$$

达到失效循环次数为

$$N_R=R_0+\frac{D_c}{\frac{2\sigma_f^2}{E^2S}\left[\frac{2}{3}(1+\nu)\frac{\Delta\sigma}{2}+\frac{1}{9\sigma_f^2}(1-2\nu)\left(\frac{\sigma}{2}\right)^3\right]} \tag{5.3-137}$$

应力幅 $\Delta\sigma$ 随失效循环次数变化的关系图曲线称为 Woehler 曲线，见图 5.3-34，以上几种情况损伤的分析是建立在各向同性的假设上。

2.4.6 复合材料的损伤失效分析技术和方法

复合材料是指由两种或两种以上均质体结合在一起形成的异质混合体，如纤维复合材料、颗粒增强材料、夹层板等。

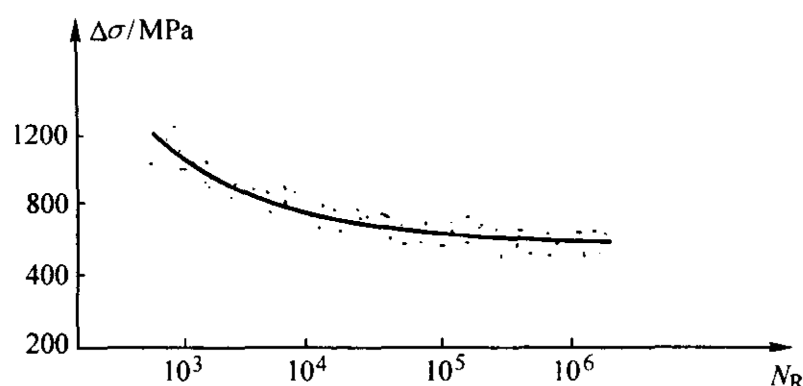


图 5.3-34　AISI316 不锈钢高周疲劳的 Woehler 曲线

由于复合材料具有非常好的性能如高强度比和高刚度比等，因而得到了迅速发展和广泛的应用。

但在其使用过程中，损伤问题是一个比较重要的问题，因为它将会导致复合材料刚度和强度的下降，进而直接影响其使用寿命。同时不同的复合材料具有不同的损伤机制，损伤的因素也很多。在准静载荷或疲劳载荷的作用下，基体中的层内横向开裂、纤维断裂、层间分层（开裂）等都会引起复合材料的刚度和强度下降。

材料有 3 个对称平面，相互正交的轴平行于纤维。采用主材料坐标系。

弹性参量中弹性模量 E_i，和切变模量 G_{ij} 的变化显示出损伤。由于泊松比 ν_{ij} 很小，这里不考虑它的影响；且假设 ν_{ij}/E_i 不变，同前面部分相类似，有

$$\begin{cases}\tilde{E}=E_i(1-D_i)\\ \tilde{G}_{ij}=G_{ij}(1-D_{ij})\end{cases}\tag{5.3-138}$$

弹性本构方程是

$$\begin{cases}\varepsilon_{ij}=\dfrac{\sigma_{ij}}{E_i(1-D_i)}-\Gamma_{ij}\sigma_{ij}-\Gamma_{jk}\sigma_{kk}(i,j,k\text{ 为 }1,2,3\text{ 的排列})\\ \Gamma_{ij}=\nu_{ij}/E_i\\ \varepsilon_{ij}=\dfrac{\sigma_{ij}}{G_{ij}(1-D_{ij})}\end{cases}\tag{5.3-139}$$

拉伸时，性能是线弹性的，压缩时，存在非线性损伤。

对于（1，2）平面上沿着和第一主方向成角 φ 的方向上的应力 σ_{xx} 有，

$$\begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\end{bmatrix}=\begin{bmatrix}\dfrac{1}{\tilde{E}} & & \\ & \dfrac{1}{\tilde{E}_2} & \\ & & \dfrac{1}{2\tilde{G}_{12}}\end{bmatrix}\begin{bmatrix}\cos^2\varphi\\ \sin^2\varphi\\ -\sin\varphi\cos\varphi\end{bmatrix}\sigma_{xx}\tag{5.3-140}$$

和

$$\begin{bmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \varepsilon_{xy}\end{bmatrix}=\begin{bmatrix}\cos^2\varphi & \sin^2\varphi & -2\sin\varphi\cos\varphi\\ \sin^2\varphi & \cos^2\varphi & 2\sin\varphi\cos\varphi\\ \sin\varphi\cos\varphi & -\sin\varphi\cos\varphi & \cos^2\varphi-\sin^2\varphi\end{bmatrix}\begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\end{bmatrix}\tag{5.3-141}$$

弹性的本构关系为

$$\varepsilon_{ij}^{e}=\frac{\sigma_{ij}}{G_{ij}(1-D_{ij})}\tag{5.3-142}$$

由上式即可得出复合材料拉伸时的损伤度。

分散损伤是另一种复合材料损伤的主要形式，它是板中和板与板之间（局部分层）的损伤。这种形式也影响复合材料的总性质。设 V 为 P 点附近损伤实体代表性样本的体积，首先考虑在体积 V 内以面积 S 为界的一个单一的损伤实体的情况。设具有单位法线矢量 $\boldsymbol{n}_i$ 的 S 上一点处的矢量 $\boldsymbol{a}_i$ 代表该点对周围材料的某种影响（如位移）张量。

$$d_{ij}=\int_S a_i n_i\,\mathrm{d}S\tag{5.3-143}$$

上式表示了一个损伤实体，则平均损伤张量为

$$D_{ij}=\frac{1}{V}\sum_k(d_{ij})_k\tag{5.3-144}$$

式中，k 为 V 中损伤实体的个数；D_{ij} 为 P 点损伤的场。由上式即可分析分散损伤对总的性质的影响。

除了上述两种损伤形式外，还有许多机制会引起复合材料的损伤，如纤维失效、基体开裂、界面脱离等，因此需要多个损伤变量来进行描述，且这些变量都属于各向异性的，需要提出新的各向异性指标来替代前面提出的各向同性的损伤指标。因此，复合材料的损伤分析要比前面的分析和计算复杂得多，它是一个目前正在积极探索的研究领域。

3　以数理统计为基础的失效预测技术和方法

3.1　失效分析概率统计基础

通常的材料力学、弹塑性力学和损伤力学失效分析和完整性评估都是确定性的方法，取一定的安全系数，得到安全和不安全两种非此即彼的结论，是防止失效的评定，而不是预计危险程度和安全程度的评定，因而不能用于定量比较构件的危险程度和安全程度。

材料性能的实测数据具有分散性，服役过程中的环境因素有时会使材料性能发生变化，构件的几何尺寸也在公差允许范围内波动，外载荷也具有不同程度的随机性。所以危险点的应力实际上是随机的，该处材料的强度也是随机的，而且这种随机性有时还会随着时间发生变化，因此失效分析和安全评定中要考虑上述随机性，评定作出的判断应该是危险程度或安全程度的概率表示。危险程度或安全程度的概率为定量比较构件危险程度或安全程度提供的重要指标。

本节系统介绍了失效分析概率评定中所涉及的概率统计基础知识。

3.1.1　总体、个体和样本

总体是指具有某些共同特征的研究对象的全体。总体可以是有限集，也可以是无限集。对任何给定的总体，其特性参数都是固定的常数。

总体中的每一个研究对象称为个体。实际问题往往不允许对总体中的每一个个体逐一进行研究。只能从总体中抽出一部分个体进行研究。这一部分个体称为样本（或子样）。样本所包含的个体的数目，称为样本的容量。样本是从总体中抽出的代表，用以推断总体的特征参数。例如为了研究一定工艺下生产的材料性能指标（总体），只能选出部分试件（个体）进行研究，用以推断该工艺条件下生产的材料性能指标。

通常讨论的都是简单随机样本。设样本 x_1，x_2，…，x_n，样本容量为 n，如果①每个 x_k（$k=1$，2，…，n）与总体 X 具有相同的分布；②总体中的每一个个体被抽取到的机会相等，并且 x_1，x_2，…，x_n 相互独立，则称此样本为简单随机样本。

一般将容量 $n>50$ 的样本称为大样本，$n<50$ 的样本称为小样本。但是这不是绝对标准，应视具体问题而定。

根据样本数据进行计算而得到的量称为统计量。它是样本的函数 $f(x_1,x_2,\cdots x_n)$，也是随机变量。

3.1.2　总体的常用概率分布

对于具有分散性的、所研究构件的某些性质（如寿命、强度和裂纹分布等），常常采用正态分布、对数正态分布和威布尔分布进行描述。

(1) 正态分布

1) 正态分布的性质　若随机变量的 X 概率密度为

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\quad(-\infty<x<\infty)\tag{5.3-145}$$

式中，μ、σ ($\sigma>0$) 为常数，则称 X 服从均值为 μ、标准差为 σ 的正态分布或高斯 (Gauss) 分布，记为 $X\sim N(\mu,\sigma^2)$。

正态分布累积概率分布函数为

$$F(x)=P(X\leqslant x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\mathrm{d}x\tag{5.3-146}$$

上式的计算借助于标准正态分布函数表得到。

正态分布是对称分布，其概率密度函数 $f(x)$ 对于直线 $x=\mu$ 是对称函数。正态分布概率密度曲线 $y=f(x)$ 的位置完全由均值 μ 所确定，所以称 μ 为位置参数。标准差 σ 反映随机变量在均值 μ 附近的分散程度，如图 5.3-35 和图 5.3-36 所示。标准差 σ 越小，概率密度函数 $y=f(x)$ 图形变得越尖，分散性越小，故而 X 落在 μ 附近的概率越大。

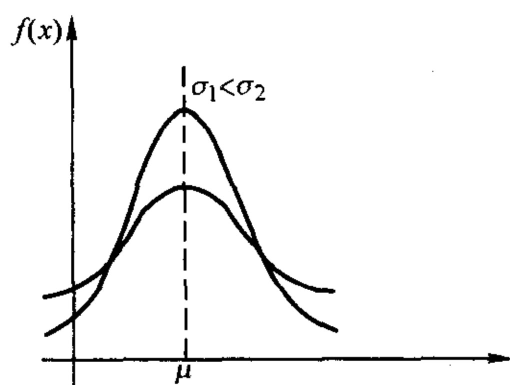


图 5.3-35　正态分布的 $f(x)$ 示意图

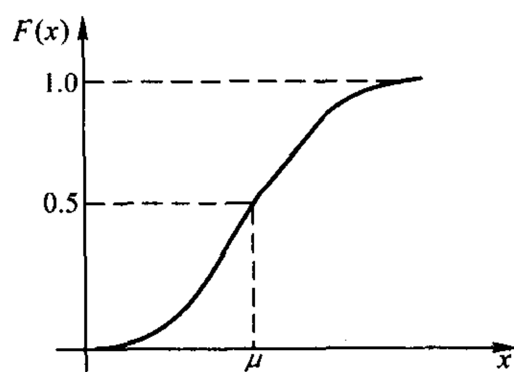


图 5.3-36　正态分布 $F(x)$ 示意图

对于正态随机变量 X 有

$$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)\approx0.997\,4\tag{5.3-147}$$

或

$$P(|X-\mu|>3\sigma)\approx0.002\,6$$

即正态随机变量的值落在区间 $[\mu-3\sigma,\ \sigma+3\sigma]$ 内几乎是肯定的事件；而它的值落在区间 $[\mu-3\sigma,\ \mu+3\sigma]$ 之外的事件是小概率事件。这就是"3σ 规则"。这是异常数据取舍的常用标准，如果 X 落在 $[\mu-3\sigma,\ \mu+3\sigma]$ 之外，则认为是异常数据，应予以舍去。

2) 标准正态分布　若正态分布随机变量 $x\sim N(\mu,\sigma^2)$，其中 $\mu=0$，$\sigma=1$ 时，称随机变量 X 服从标准正态分布，记为 $X\sim N(0,1)$。其概率密度函数和累积概率分布函数分别用 $\phi(x)$ 和 $\Phi(x)$ 表示。

即有

$$\phi(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\quad(-\infty<x<\infty)\tag{5.3-148}$$

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{x^2}{2}\right)\mathrm{d}x\tag{5.3-149}$$

显然，$\Phi(x)=1-\phi(x)$。

一般正态累积分布函数由标准正态分布函数 $\Phi(x)$ 求得。若 $X\sim N(\mu,\sigma^2)$，则 $(X-\mu)/\sigma\sim N(0,1)$，所以对 $X\sim N(\mu,\sigma^2)$ 有

$$F(x)=P(X\leqslant x)=P\left(\frac{X-\mu}{\sigma}\leqslant\frac{x-\mu}{\sigma}\right)=\Phi\left(\frac{x-\mu}{\sigma}\right)\tag{5.3-150}$$

$$P(x_1<X\leqslant x_2)=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right)\tag{5.3-151}$$

3) 中心极限定理　设随机变量 X_1，X_2，…，X_n 相互独立，服从同一分布，且具有相同数学期望和方差：$E(X_k)=\mu,D(X_k)=\sigma^2\neq0(k=1,2,\cdots)$，则随机变量

$$\overline{X}=(X_1+X_2+\cdots+X_n)/n\tag{5.3-152}$$

当 n 充分大时，服从正态分布

$$\overline{X}\sim N(\mu,\sigma^2/n)\tag{5.3-153}$$

这说明，尽管总体的概率分布不一定是正态分布，只要样本容量 n 充分大，样本的均值 $\overline{X}$ 呈现近似的正态分布，其均值等于总体均值 μ，其标准差等于 $\sigma/\sqrt{n}$，σ 是总体的标准差。即 n 充分大时，可以用样本的均值 $\overline{X}$ 推断总体均值 μ，并且 n 越大时估计越准确。这正是正态随机变量得到广泛应用的原因。

4) 描述总体分布的数字特征　总体的概率密度函数大致位置、离散程度和分布特征通常用以下数字特征描述。

① 位置特征参数　一般用均值表示，均值是描述总体分布取值平均位置的参数，用符号 μ 表示。

② 散布特征参数　散布特征参数用来描述总体分布在均值附近取值的分散程度。常用标准差 σ 或变异系数 C 表示。变异系数为

$$C=\sigma/\mu\tag{5.3-154}$$

标准差 σ 或变异系数 C 越小，在均值附近取值的分散程度越小。

③ 分布特征参数　分布特征参数常用的有偏度系数和峰度系数，表示总体概率密度函数相对正态分布（相对同均值 μ 和标准差 σ 的正态分布）的差异。

偏度系数表示总体分布的不对称性，用三阶中心矩表示为

$$\gamma_1=E[(x-\mu)^3]$$

或用量纲为 1 的三阶中心矩表示为：

$$\gamma'_1=E[(X-\mu)^3]/\sigma^2$$

式中，σ 为总体标准差。当 $\gamma_1>0$（或 $\gamma'_1>0$）时为正偏，$\gamma_1<0$ 或（$\gamma'_1<0$）时为负偏，$\gamma_1=0$ 或（$\gamma'_1=0$）时为对称分布——正态分布，如图 5.3-37 表示。

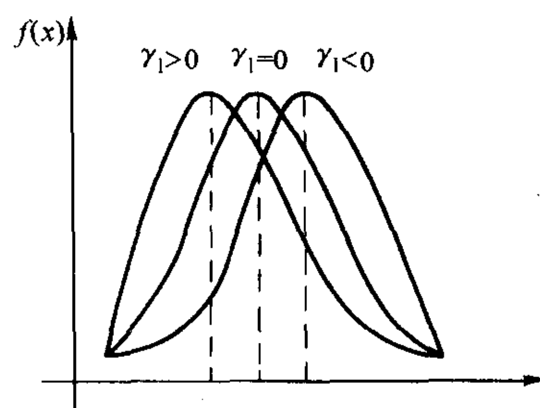


图 5.3-37　偏度系数示意图

峰度系数用来表示总体概率密度函数图形顶峰的凸平度，用四阶中心矩或量纲为 1 的四阶中心矩表示为

$$\gamma_2 = E[(X-\mu)^4] \quad (5.3\text{-}155)$$

或

$$\gamma'_2 = E[(X-\mu)^4]/\sigma^4 - 3$$

量纲为 1 的峰度系数 $\gamma'_2>0$ 表示比对应相同均值 μ 和标准差 σ 的正态分布概率密度图形顶峰更凸出；量纲为 1 的峰度系数 $\gamma'_2<0$ 表示比相同均值 μ 和标准差 σ 的正态分布概率密度图形顶峰低；量纲为 1 的峰度系数 $\gamma'_2=0$ 表示正态分布。如图 5.3-38 所示。

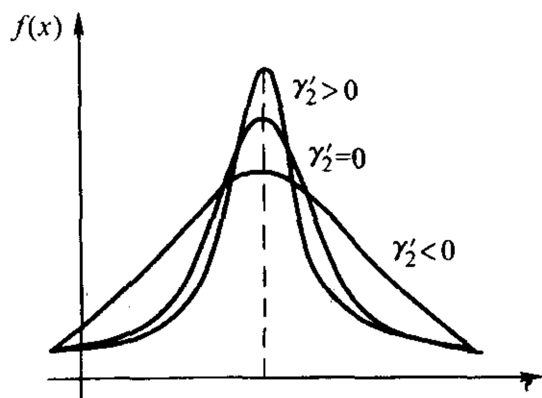


图 5.3-38　量纲为 1 的峰度系数示意图

(2) 对数正态分布

设随机变量 X 的自然对数 $Y=\ln X$ 服从正态分布，即 $\ln X \sim N(\mu,\sigma^2)$，则称 X 服从对数正态分布。它的概率密度函数和累积概率分布函数分别为

$$f(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma x}\times\exp\left[-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right] & (1<x<\infty)\\ 0 & (x\leqslant 0)\end{cases} \quad (5.3\text{-}156)$$

$$F(x)=P(X\leqslant x)=\Phi\left[\frac{\ln x-\mu}{\sigma}\right] \quad (5.3\text{-}157)$$

$$P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1)=\Phi\left(\frac{\ln x_2-\mu}{\sigma}\right)-\Phi\left(\frac{\ln x_1-\mu}{\sigma}\right) \quad (5.3\text{-}158)$$

式中　$\Phi(\cdot)$——标准正态分布函数。

对数正态分布的均值和方差分别为

$$E(X)=\exp\left(\mu+\frac{\sigma^2}{2}\right) \quad (5.3\text{-}159)$$

$$D(X)=\exp(2\mu+\sigma^2)[\exp(\sigma^2)-1] \quad (5.3\text{-}160)$$

对数正态分布的概率密度函数为正偏，其三阶中心矩（偏度系数）为

$$\gamma_1=\exp(3\mu+4.5\sigma^2)-3\exp(3\mu+2.5\sigma^2)+2\exp(3\mu+1.5\sigma^2) \quad (5.3\text{-}161)$$

(3) 威布尔分布

设随机变量 X 服从三参数威布尔分布，则其概率密度函数和累积概率分布函数为

$$f(x)=\begin{cases}\dfrac{\beta}{\alpha}(x-\gamma)^{\beta-1}\times\exp\left[-\dfrac{(x-\gamma)^\beta}{\alpha}\right] & (x\geqslant\gamma)\\ 0 & (x<\gamma)\end{cases} \quad (5.3\text{-}162)$$

$$F(x)=P(X<x)=1-\exp\left[-\frac{(X-\gamma)^\beta}{\alpha}\right] \quad (\gamma\leqslant x<\infty) \quad (5.3\text{-}163)$$

式中，$\alpha>0$，为尺度参数；$\beta>0$，为形状参数；γ 为形状参数。

三参数威布尔分布的均值和方差为

$$\mu=\gamma+\alpha^{1/\beta}\Gamma\left(1+\frac{1}{\beta}\right) \quad (5.3\text{-}164)$$

$$\sigma^2=\alpha^{2/\beta}\left[\Gamma\left(1+\frac{2}{\beta}\right)-\Gamma^2\left(1+\frac{1}{\beta}\right)\right] \quad (5.3\text{-}165)$$

偏度系数用三节中心矩表示为

$$\gamma_1=\alpha^{3/\beta}\left[\Gamma\left(1+\frac{3}{\beta}\right)-3\Gamma\left(1+\frac{2}{\beta}\right)\Gamma\left(1+\frac{1}{\beta}\right)+2\Gamma^2\left(1+\frac{1}{\beta}\right)\right] \quad (5.3\text{-}166)$$

3.1.3　实验数据的数字特征计算

总体的特征参数的估计是通过总体样本数据进行估计的。下面介绍样本数据的处理方法。

(1) 位置特征参数计算

实验数据 $x_1, x_2, \cdots, x_n$ 的位置特征参数是描述实验数据的取值平均位置和特定位置的，其中常用的位置特征参数有均值、最大值、最小值、中位数和众数。

1) 算术均值　算术均值一般用 $\bar{x}$ 表示，是描述实验数据取值平均位置的，如

$$\bar{x}=\frac{1}{n}(x_1+x_2+\cdots+x_n)=\frac{1}{n}\sum_{i=1}^{n}x_i \quad (5.3\text{-}167)$$

2) 最大值、最小值、中位数和众数　设 n 个数据按从小到大的顺序排列为 $x_{(1)}\leqslant x_{(2)}\leqslant\cdots\leqslant x_{(n)}$，则最大值为 $x_{(n)}$，最小值为 $x_{(1)}$，而中位数为

$$M=\begin{cases}x_{\left(\frac{n+1}{2}\right)} & n\text{ 为奇数时}\\ \dfrac{1}{2}\left[x_{\frac{n}{2}}+x_{\left(\frac{n}{2}+1\right)}\right] & n\text{ 为偶数时}\end{cases}$$

众数是指使频数最大的 x_i 值。

(2) 散布特征参数

在实际应用中，常常要求计算以均值 $\bar{x}$ 为中心的散布特征参数，用来描述实验数据在均值 $\bar{x}$ 附近的分散程度。

1) 方差和标准差　样本方差的无偏估计量为

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 \quad (5.3\text{-}168)$$

标准差估计量为：

$$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad (5.3\text{-}169)$$

2) 变异系数　实验数据的变异系数（对总体变异系数的估算公式）为

$$C=S/\bar{x}$$

均值给定时，变异系数越大，标准差就越大，数据的分散程度就越大。变异系数比标准差更好地反映了数据分散的特点。

(3) 分布特征参数

分布特征参数有偏度系数和峰度系数。

1) 偏度系数　偏度系数是用来描述分布的不对称程度的。分为正偏、负偏和对称。量纲为 1 的偏度系数估计式为

$$\gamma_1=\mu_3/S^3 \quad (5.3\text{-}170)$$

$$\mu_3=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3 \quad (5.3\text{-}171)$$

2) 峰度系数　峰度系数是用来描述总体概率密度函数图形顶峰相对同均值和方差的正态分布图形凸平程度。常用下式估计：

$$\gamma'_2=\mu^4/S^4-3 \quad (5.3\text{-}172)$$

$$\mu_4=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4 \quad (5.3\text{-}173)$$

$\gamma'_2>0$ 表示顶峰比正态分布更凸出，$\gamma'_2<0$ 表示顶峰不如正态分布凸出，正态分布时 $\gamma'_2=0$。

由于正态分布的偏度系数和峰度系数皆为零，所以可用偏度检验和峰度检验来检验一个分布是否为正态分布。此时为了消除实验数据量纲的影响和便于进行统计检验，给出标准偏度系数和标准峰度系数分别为

$$g_1=\sqrt{\frac{1}{6n}}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{S}\right)^3 \quad (5.3\text{-}174)$$

$$g_2=\sqrt{\frac{n}{24}}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{S}\right)^4-3\right] \tag{5.3-175}$$

当 n 充分大时，标准偏度系数和标准峰度系数都以标准正态分布为渐近分布。据此对统计量 g_1，g_2 进行 $g_1=0$，$g_2=0$ 的显著性检验。用这种方法样本容量最好大于100。

3.1.4　统计分析中常用的几种分布

在统计分析中网用的分布有 χ^2 分布、t 分布、F 分布和次序统计量分布。

(1) χ^2 分布

若 X_1，X_2，…，X_n 是 n 个相互独立且服从标准正态分布的随机变量，则有统计量

$$\chi^2=X_1^2+X_2^2+\cdots+X_n^2$$

服从自由度为 n 的 χ^2 分布，记为 $\chi^2\sim\chi^2(n)$。

χ^2 分布的概率密度函数为

$$f_n(\chi^2)=\begin{cases}\dfrac{2}{2^{n/2}\Gamma(n/2)}(\chi^2)^{n/2-1}\times\exp\left(-\dfrac{1}{2}\chi^2\right) & (\chi^2\geqslant 0)\\ 0 & (\chi^2<0)\end{cases} \tag{5.3-176}$$

式中，n 为正整数；$\Gamma(\cdot)$ 为 Γ 函数，其通式 $\Gamma(P)=\int_0^\infty t^{p-1}e^{-t}\mathrm{d}t$。当 n 较小时，$f_n(\chi^2)$是不对称的。随着 n 的增大，$f_n(\chi^2)$逐渐趋于对称，当 $n>30$ 后，$f_n(\chi^2)$接近正态分布。

(χ^2) 分布的均值和方差分别为

$$\mu=n\quad(n=1,\ 2,\ \cdots)$$
$$\sigma^2=2n\quad(n=1,\ 2,\ \cdots)$$

χ^2 分布的基本性质如下：

1) 总体 $X\sim N(\mu,\sigma^2)$，X_1，X_2，…，X_n 为该总体的容量为 n 的简单随机样本，若均值为已知，用 $\bar{X}$ 代替。则有：

$$\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2}\sim\chi^2(n) \tag{5.3-177}$$

若 μ 未知，用 $\bar{X}$ 代替，则有

$$\frac{\sum_{i=1}^{n}(x_i-\bar{X})^2}{\sigma^2}\sim\chi^2(n-1) \tag{5.3-178}$$

2) 若总体$X\sim N(\mu,\sigma^2)$，X_1，X_2，…，X_n 是从总体抽出的容量为 n 的简单随机样本，则有：

$$\frac{n-1}{\sigma^2}S\sim\chi^2(n-1) \tag{5.3-179}$$

式中样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$，并且 S^2 和均值 $\bar{X}$ 相互独立。

3) 若 χ_1，χ_2，…，χ_n 为相互独立的随机变量，每个都服从自由度为 n_i 的 χ^2 分布 $\chi_i^2=\chi^2(n_i)$，则

$$\sum_{i=1}^{k}\chi_i^2\sim\chi^2\left(\sum_{i=1}^{k}n_i\right) \tag{5.3-180}$$

此即 χ^2 分布的可加性。

4) χ^2 统计量的概率积分　如图5.3-39所示，对于 χ^2 统计量，如果

$$P(\chi^2\leqslant\chi_P^2)=\int_0^{\chi_P^2}f(\chi^2)\mathrm{d}\chi^2=p \tag{5.3-181}$$

则 χ_P^2 称为 χ^2 分布的 p 下侧分位数。如果

$$P(\chi^2\geqslant\chi_\alpha^2)=\int_{\chi_\alpha^2}^{\infty}f(\chi^2)\mathrm{d}\chi^2=\alpha \tag{5.3-182}$$

则 χ_P^2 称为 χ^2 分布的 α 上侧分位数。其他分布的上侧分位数或下侧分位数定义与此相同。χ^2 分布表在数学手册上可以查到。

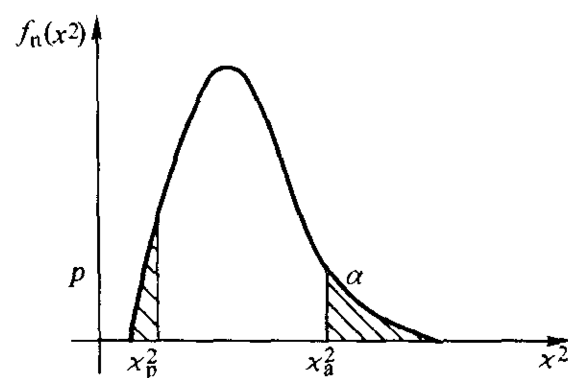


图5.3-39　χ^2 概率密度函数

(2) t 分布

设随机变量 X，Y 相互独立，且 $X\sim N(0,1)$，$Y\sim\chi^2(n)$，则随机变量 $X/\sqrt{Y/n}$服从自由度为 n 的 t 分布，记为

$$\frac{X}{\sqrt{Y/n}}\sim t(n) \tag{5.3-183}$$

自由度为 n 的 t 分布的概率密度函数为

$$f_n(t)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{2\pi}\Gamma\left(\frac{n}{2}\right)}\times\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}$$
$$(-\infty<t<+\infty,n=1,2,\cdots)$$

$f_n(t)$的图形是关于 $t=0$ 对称的，当 $n\to\infty$ 时，t 分布趋近于标准正态分布。

t 分布的均值和方差分别为：

$$\mu=0$$
$$\sigma^2=\frac{n}{n-1}\quad(n\geqslant 3)$$

t 分布的基本性质如下：

1) 若 X_1，X_2，…，X_n 是来自正态总体$X\sim N(\mu,\sigma^2)$的简单随机样本，则统计量

$$\frac{(\bar{X}-\mu)}{S/\sqrt{n}}\sim t(n-1) \tag{5.3-184}$$

式中

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$
$$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

2) 若 X_1，X_2，…，X_{n_1} 是来自正态总体 $X\sim N(\mu_1,\sigma^2)$ 的简单随机样本，Y_1，Y_2，…，Y_{n_2} 是来自正态总体 $N(\mu_2,\sigma^2)$的简单随机样本，且 X，Y 之间相互独立，则

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_{n'}\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2) \tag{5.3-185}$$
$$S_{n'}^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

式中，S_1、S_2 分别为这两个样本的方差的估计。

3) 当 n 充分大时，t 分布近似于标准正态分布。

在 $n>45$ 时，t 分布 α 分位数（上侧分位数或下侧分位数均可）可近似用式（5.3-186）表示。

$$t_\alpha(n)\approx u_\alpha \tag{5.3-186}$$

式中，u_α 为标准正态分布的上侧分位数或下侧分位数。

4) 由于 t 分布关于 y 轴的对称性，t 分布分位数（上侧分位数或下侧分位数均可）具有下述性质。

$$t_p(n)=-t_{1-p}(n) \tag{5.3-187}$$

标准正态分布也具有这种性质，表示为 $u_p=-u_{1-p}$。

另外

$$P[|t|\geqslant t_{\alpha/2}(n)]=1-\int_{-t_{\alpha/2}}^{t_{\alpha/2}}f_n(t)\mathrm{d}t=\alpha \tag{5.3-188}$$

等价于

$$P[t \geqslant t_{\alpha/2}(n)] = P[t \leqslant t_{\alpha/2}(n)] = \alpha/2$$

t 分布表可在数学手册上查到。

(3) F 分布

设随机变量 U，V 相互独立，且 $U \sim \chi^2(n_1)$，$V \sim \chi^2(n_2)$，那么 $\frac{U/n_1}{V/n_2}$ 服从自由度为（n_1，n_2）的 F 分布，记为

$$\frac{U/n_1}{V/n_2} \sim F(n_1, n_2) \quad (5.3\text{-}189)$$

自由度为（n_1，n_2）的 F 分布的概率密度函数为

$$f_{n_1,n_2}(F) = \begin{cases} \dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right)}{\Gamma\left(\dfrac{n_1}{2}\right)\Gamma\left(\dfrac{n_2}{2}\right)} \times \left(\dfrac{n_1}{n_2}\right)^{n_1/2} \times \dfrac{F^{(n_1/2-1)}}{\left(1+\dfrac{n_1}{n_2}F\right)^{\frac{n_1+n_2}{2}}} & (F \geqslant 0) \\ 0 & (F < 0) \end{cases}$$

式中，n_1、$n_2 = 1, 2, \cdots$。

F 分布有两个参数 n_1 和 n_2，是不对称分布。

F 分布的均值和方差为

$$\mu = \frac{n_2}{n_2 - 2} \quad (n_2 > 2)$$

$$\sigma^2 = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)} \quad (n_2 > 4)$$

F 分布的基本性质如下：

1) 若 X_1，X_2，…，X_{n_1} 是来自正态总体 $X \sim N(\mu_1, \sigma_1^2)$ 的简单随机样本，Y_1，Y_2，…，Y_{n_2} 是来自正态总体 $N(\mu_2, \sigma_2^2)$ 的简单随机样本，且它们之间相互独立，则

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1) \quad (5.3\text{-}190)$$

2) 若随机变量 x 服从 $F(n_1, n_2)$ 分布，则随机变量 $1/x$ 也服从 $F(n_2, n_1)$ 分布。

3) 若 $F_\alpha(n_1, n_2)$ 是自由度为 n_1，n_2 的 F 分布的上侧分位数（或下侧分位数），则

$$F_{1-\alpha}(n_1, n_2) = \frac{1}{F_\alpha(n_2, n_1)} \quad (5.3\text{-}191)$$

数学手册上列有 F 分布表，它只是部分 F 分布表，其余可利用式（5.3-191）计算。

(4) 次序统计量分布

设有样本 X_1，X_2，…，X_n，把它们按从小到大的顺序排列为 $x_{(1)} \leqslant x_{(2)} \leqslant \cdots \leqslant x_{(n)}$，则称 $x_{(1)}$，$x_{(2)}$，…，$x_{(n)}$ 为原样本 X_1，X_2，…，X_n 的次序统计量，特别是 $X_{(i)}$ 单个称为"第 i 个次序统计量"，是样本的函数，也是随机变量。

如果 $F(x)$ 是总体 X 的分布函数，$X_{(i)}$ 对应的总体分布函数取值 $F[X_{(i)}]$ 是随 $X_{(i)}$ 变化的，是一个新的随机变量。根据理论证明，无论总体怎样分布，$F[X_{(i)}]$ 的数学期望都是恒定的，称为平均秩，表示为

$$E\{F[X_{(i)}]\} = \frac{i}{n+1} \quad (5.3\text{-}192)$$

工程上常把平均秩作为总体累积分布的估计量。

3.1.5 参数估计

(1) 良好的估计量的性质

由样本 X_1，X_2，…，X_n 构造适当统计量 $\hat{\theta} = \hat{\theta}$（$X_1$，$X_2$，…，$X_n$）来估计总体未知参数 θ 的方法，称为对参数 θ 的点估计。统计量 $\hat{\theta}$ 称为参数的估计量。

估计量 $\hat{\theta}$ 是样本函数，也是一个随机变量，也有它要服从的分布。评价各种点估计量的优劣，常常有三个原则，即无偏性、一致性和有效性。

1) 无偏性　如果估计量 $\hat{\theta}$ 和未知参数 θ 满足

$$E(\hat{\theta}) = \theta \quad (5.3\text{-}193)$$

则称 $\hat{\theta}$ 为 θ 的无偏估计量。

2) 有效性　若 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 是同一参数 θ 的无偏估计量，如果

$$D(\hat{\theta}_1) < D(\hat{\theta}_2) \quad (5.3\text{-}194)$$

则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效。

3) 一致性　若对于任意给定的一个小值 $\varepsilon > 0$，有

$$\lim P(|\hat{\theta} - \theta| < \varepsilon) = 1 \quad (5.3\text{-}195)$$

则称 $\hat{\theta}$ 为 θ 的一致估计量。

(2) 矩法参数估计

1) 正态分布的参数估计　对于正态分布用样本均值 $\overline{X}$ 和样本方差 S^2 作为总体均值 μ 和方差 σ^2 的估计量。

2) 对数正态分布的参数估计　对数正态总体 X 的简单随机变量 X_1，X_2，…，X_n，则 $\ln X_1$，$\ln X_2$，…，$\ln X_n$ 是正态总体 $\ln X \sim N(\mu, \sigma^2)$ 的样本，通过估计 $\ln X$ 正态总体的均值 μ 和方差 σ^2，由式（5.3-159）和式（5.3-160）得到对数正态总体 X 的均值和方差。

3) 威布尔分布的参数估计　若从服从威布尔分布的总体中抽得了一个容量为 n 的样本 X_1，X_2，…，X_n，则可用相应的样本矩去估计总体数学期望 μ、方差 σ^2 和偏度系数 γ_1，即：

$$\hat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

$$\hat{\gamma}_1 = \frac{\hat{\mu}_2}{\hat{\sigma}}$$

$$\hat{\mu}_3 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^3$$

用 $\hat{\mu}$，$\hat{\sigma}^2$ 和 $\hat{\gamma}_1$ 分别代替式（5.3-146）、式（5.3-165）和式（5.3-166）中的 μ，σ^2 和 γ_1，可以得到威布尔分布的三个参数 α，β 和 γ 的估计值。

3.1.6 分布参数的区间估计

对总体未知参数 θ 作估计时，点估计法只是对未知参数给出一个估计数值。不同样本得到不同的估计数值，要将估计量的这种不确定性考虑进去，所采用的方法就是对未知参数 θ 提出一个估值区间。对未知参数给出一个估值区间的方法，称为区间估计法。

假设由样本 X_1，X_2，…，X_n 建立两个统计量 $\theta_1(X_1, X_2, \cdots, X_n)$ 与 $\theta_2(X_1, X_2, \cdots, X_n)$，且 $\theta_1 < \theta_2$，若对给定的很小常数 $\alpha(0 < \alpha < 1)$ 有

$$P[\theta_1(X_1, X_2, \cdots, X_n) < \theta < \theta_2(X_1, X_2, \cdots, X_n)] = 1 - \alpha$$

则称区间（θ_1，θ_2）是未知参数 θ 的置信度为（$1-\alpha$）的置信区间，θ_1 称为置信下限，θ_2 称为置信上限，α 称为显著性水平。

(1) 正态总体均值和区间估计

设总体分布为 $X \sim N(\mu, \sigma^2)$，X_1，X_2，…，X_n 为简单随机样本，则有均值 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 服从正态分布 $N(\mu, \sigma^2/n)$。

1) 方差 σ^2 已知的情况　此时统计量

$$U = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1) \quad (5.3\text{-}196)$$

对于给定的小常数 $\alpha(0 < \alpha < 1)$，使得

$$P\left(\left|\frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}}\right| \leqslant u_{\alpha/2}\right) = 1 - \alpha$$

或

$$P\left(\overline{X}-u_{\alpha/2}\frac{\sigma}{n}\leqslant\mu\leqslant\overline{X}+u_{\alpha/2}\frac{\sigma}{n}\right)=1-\alpha \quad (5.3\text{-}197)$$

$u_{\alpha/2}$为正态分布上侧分位数。

由中心极限定理可知，无论总体 X 服从怎样的分布，只要 n 充分大，随机变量$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，所以只要样本容量 n 足够大（$n>50$），仍然可用式（5.3-197）作为均值的置信区间。

2）方差 σ^2 未知的情况　此时统计量

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1) \quad (5.3\text{-}198)$$

式中，$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 为 σ^2 的无偏估计量。

对于给定显著性水平$\alpha(0<\alpha<1)$，使得

$$P\left[\left|\frac{\overline{X}-\mu}{S/\sqrt{n}}\right|\leqslant t_{\alpha/2}(n-1)\right]=1-\alpha$$

或

$$P\left(\overline{X}-t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\leqslant\mu\leqslant\overline{X}-t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)$$

所以均值 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}},\overline{X}+t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right) \quad (5.3\text{-}199)$$

式中，$t_{\alpha/2}(n-1)$为上侧分位数。

(2) 正态总体方差的置信区间

设正态总体$X\sim N(\mu,\sigma^2)$，X_1，X_2，…，X_n 为简单随机变量，则有

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

对于给定显著性水平 α，使得

$$P\left[\frac{(n-1)S^2}{\sigma^2}\geqslant\chi^2_{1-\alpha/2}(n-1)\right]=1-\frac{\alpha}{2}$$

$$P\left[\frac{(n-1)S^2}{\sigma^2}\geqslant\chi^2_{\alpha/2}(n-1)\right]=\frac{\alpha}{2}$$

于是有

$$P\left[\chi^2_{1-\alpha/2}(n-1)\leqslant\frac{(n-1)S^2}{\sigma^2}\leqslant\chi^2_{\alpha/2}(n-1)\right]=1-\alpha$$

等价于

$$P\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\leqslant\sigma^2\leqslant\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right]=1-\alpha$$

因此，总体方差 σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right] \quad (5.3\text{-}200)$$

或

$$\left[\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\chi^2_{\alpha/2}(n-1)},\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\chi^2_{1-\alpha/2}(n-1)}\right] \quad (5.3\text{-}201)$$

式中，$\chi^2_{\alpha/2}(n-1)$、$\chi^2_{1-\alpha/2}(n-1)$为上侧分位数。

通过以上方法可进行正态分布总体、对数正态分布总体的均值和方差的区间估计以及大样本非正态分布的均值估计。

(3) 方差相容多组数据的方差置信区间

若收集到相近条件下的 m 组实验结果，并且通过 F 检验证明各组方差相同，则可以将这 m 组数据联合起来估计方差σ^2 的置信区间。

对于每一组容量为 n_j 的样本，则有：

$$\frac{(n_j-1)S_j^2}{\sigma^2}\sim\chi^2(n_j-1)$$

式中，$\sigma^2=E(S_j^2)$，并且各组方差相容，$\sigma_1^2=\sigma_2^2=\cdots=\sigma^2$。

根据 χ^2 分布的可知性，可知

$$\frac{1}{\sigma^2}\sum_{j=1}^{m}(n_j-1)S_j^2\sim\chi^2\left[\sum_{j=1}^{m}(n_j-1)\right] \quad (5.3\text{-}202)$$

建立概率条件

$$P\left[\frac{1}{\sigma^2}\sum_{j=1}^{m}(n_j-1)S_j^2>\chi^2_{1-\alpha/2}\right]=1-\alpha/2$$

$$P\left[\frac{1}{\sigma^2}\sum_{j=1}^{m}(n_j-1)S_j^2>\chi^2_{\alpha/2}\right]=\alpha/2$$

所以方差 σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left\{\frac{\sum_{j=1}^{m}(n_j-1)S_j^2}{\chi^2_{\alpha/2}\left[\sum_{j=1}^{m}(n_j-1)S_j^2\right]},\frac{\sum_{j=1}^{m}(n_j-1)S_j^2}{\chi^2_{1-\alpha/2}\left[\sum_{j=1}^{m}(n_j-1)S_j^2\right]}\right\} \quad (5.3\text{-}203)$$

一般 χ^2 分布表只列到 $n=30$ 为止，$n>30$ 时采用近似计算公式

$$\chi^2_{1-\alpha}(n)\approx\frac{1}{2}(u_{1-\alpha}+\sqrt{2n-1})^2 \quad (5.3\text{-}204)$$

式中，$\chi^2_{1-\alpha}$是上侧分位数；$u_{1-\alpha}$是标准正态分布的上侧分位数。

3.1.7　分布参数的假设检验

分布参数的假设检验，主要是指正态分布总体的均值和方差的假设检验。设正态总体$X\sim N(\mu,\sigma^2)$，X_1，X_2，…，X_n 为简单随机样本，则均值 μ 和方差σ^2 的无偏估计分别为

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

(1) 正态分布总体均值检验

1）总体方差 σ^2 已知的情况　对均值提出的原假设为

$$H_0:\mu=\mu_0 \quad (5.3\text{-}205)$$

给出的检验水平为 α。

如果原假设成立，则

$$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$$

因此

$$P\left(\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\leqslant u_{\alpha/2}\right)=1-\alpha$$

故若出现

$$\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|>u_{\alpha/2} \quad (5.3\text{-}206)$$

则应否定原假设 H_0。这是因为一次试验中竟出现了小概率事件，故可认为原假设 H_0 不合理。相反若出现

$$\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\leqslant u_{\alpha/2} \quad (5.3\text{-}207)$$

则不能否定原假设 H_0，一般可认为 H_0 成立。$u_{\alpha/2}$为上侧分位数。

2）总体方差 σ^2 未知的情况　原假设为

$$H_0:\mu=\mu_0$$

若原假设成立，则统计量

$$\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1) \quad (5.3\text{-}208)$$

所以，如果出现

$$\left|\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\right|\leqslant t_{\alpha/2}(n-1) \quad (5.3\text{-}209)$$

就可以接受 H_0 成立，否则应否定 H_0。

(2) 方差检验

原假设为

$$H_0:\sigma^2=\sigma_0^2$$

给出检验水平为 α。

若原假设成立时有

$$\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\sim\chi^2(n-1)$$

如果出现

$$\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}<\chi^2_{\alpha/2}(n-1) \tag{5.3-210}$$

或

$$\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}>\chi^2_{1-\alpha/2}(n-1) \tag{5.3-211}$$

则可认为原假设 H_0 不成立，否则 H_0 成立。

(3) 两个正态总体均值和方差的检验

设两个独立正态总体 $X_1\sim N(\mu_1,\sigma_1^2)$，$X_2\sim N(\mu_2,\sigma_2^2)$。现各取随机样本 X_{11}，X_{12}，…，X_{1n_1} 和 X_{21}，X_{22}，…，X_{2n_2}，其样本均值和方差分别为

$$\overline{X}_1=\frac{1}{n_1}\sum_{i=1}^{n_1}X_{1i},\ S_1^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_{1i}-\overline{X}_1)^2$$

$$\overline{X}_2=\frac{1}{n_2}\sum_{i=1}^{n_2}X_{2i},\ S_2^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(X_{2i}-\overline{X}_2)^2$$

1) 两个正态总体的方差 σ_1^2 和 σ_2^2 的检验　原假设为

$$H_0:\sigma^2=\sigma_0^2$$

则统计量

$$S_1^2/S_2^2\sim F(n_1-1,n_2-1) \tag{5.3-212}$$

若出现

$$S_1^2/S_2^2>F_{\alpha/2}(n_1-1,n_2-1) \tag{5.3-213}$$

或

$$S_1^2/S_2^2<F_{1-\alpha/2}(n_1-1,n_2-1) \tag{5.3-214}$$

则否定原假设，否则接受原假设。

2) 两个独立正态总体的均值检验

① σ_1^2 和 σ_2^2 未知，但通过 F 检验能够认为 $\sigma_1^2=\sigma_2^2$ 的情况

原假设为

$$H_0:\mu_1=\mu_2$$

则统计量

$$\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}}\sqrt{\frac{n_1n_2}{n_1+n_2}}\sim t(n_1+n_2-2) \tag{5.3-215}$$

对于给定的显著性水平 α，若出现

$$\left|\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}}\sqrt{\frac{n_1n_2}{n_1+n_2}}\right|\leqslant t_{\alpha/2}(n_1+n_2-2) \tag{5.3-216}$$

则接受原假设，否则否定原假设。

② σ_1^2 和 σ_2^2 未知，并且不能认为 $\sigma_1^2=\sigma_2^2$ 的情况

原假设

$$H_0:\mu_1=\mu_2$$

则统计量

$$\frac{\overline{X}_1-\overline{X}_2}{\sqrt{S_1^2/n_1+S_2^2/n_2}}\sim t(d)$$

对于给定的显著性水平 α，若出现

$$\frac{\overline{X}_1-\overline{X}_2}{\sqrt{S_1^2/n_1+S_2^2/n_2}}\leqslant t_{\alpha/2}(d) \tag{5.3-217}$$

则接受原假设，否则否定原假设。

式中

$$d=\left[\frac{C^2}{n_1-1}+\frac{(1-C)^2}{n_2-1}\right]^{-1},C=\left[1+\frac{S_2^2/n_2}{S_1^2/n_1}\right]^{-1}$$

当 d 是非整数时应取整。

③ σ_1^2 和 σ_2^2 为已知的情况

原假设

$$H_0:\mu_1=\mu_2$$

则统计量

$$\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1)$$

对于给定的显著性水平 α，若出现

$$\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\leqslant u_{\alpha/2} \tag{5.3-218}$$

则接受原假设，否则否定原假设。

实际上此方法也适合于两个独立的分布未知的总体的均值检验，因为 n 充分大时

$$X_1\sim N\left(\mu_1,\frac{\sigma_1^2}{n_1}\right),X_2\sim N\left(\mu_2,\frac{\sigma_2^2}{n_2}\right)$$

$$\overline{X}_1-\overline{X}_2\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$$

所以

$$\frac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1)$$

若 $\mu_1=\mu_2$，则

$$\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1) \tag{5.3-219}$$

(4) 假设检验中可能犯的两类错误

假设检验是在认为小概率事件不可能发生的基础上进行的，由于实际推断中小概率事件仍有可能发生，所以接受或拒绝假设都不是绝对无误的。这就导致了假设检验中可能出现的两类错误。

1) 原假设 H_0 为真时，作出拒绝 H_0 的错误推断。这个错误的概率很小，不超过给定的显著性水平 α，即

$$P(\text{拒绝 } H_0 \mid H_0 \text{ 为真})=\alpha \tag{5.3-220}$$

称这类错误为第一类错误。

2) 另一类错误是原假设 H_0 为假时，作出接受 H_0 的错误推断，称它为第二类错误。犯第二类错误的概率记为

$$P(\text{接受 } H_0 \mid H_0 \text{ 不真})=\beta \tag{5.3-221}$$

图 5.3-40 是双边检验时可能犯两类错误的示意图。

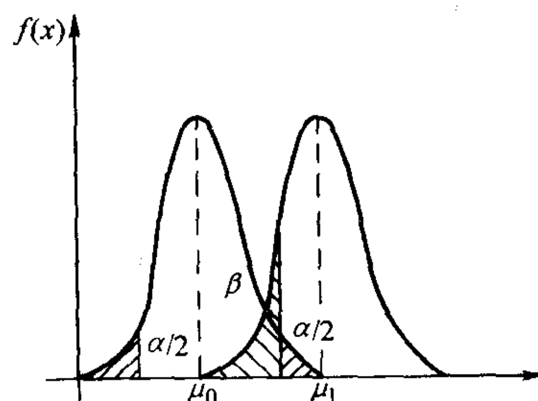


图 5.3-40　犯两类错误示意图

设正态总体双边检验问题中，原假设为

$$H_0:\mu=\mu_0,\quad H_1:\mu=\mu_1\neq\mu_0$$

且

$$|\mu_1-\mu_0|=\delta>0$$

已知

$$P(\text{拒绝 } H_0 | H_0 \text{ 为真})=\alpha$$
$$P(\text{接受 } H_0 | H_0 \text{ 为假})=\beta$$

如果 n 满足

$$\sqrt{n}\geqslant(u_{\alpha/2}+u_\beta)\frac{\sigma}{\delta} \tag{5.3-222}$$

$u_{\alpha/2}$，u_β 都是上侧分位数，则可满足犯第一类错误的概率为 α，同时犯第二类错误的概率不超过 β。

由式（5.3-222）可见，同时减少犯两类错误的概率是不可能的。如果要使 α 和 β 都达到预先制定的小水平，就必须加大样本容量。

3.1.8 总体分布的假设检验

分布的假设检验又称为分布函数的拟合优度检验。

(1) χ^2 检验

用 $F(x)$ 表示总体 X 的分布函数，需要检验的假设是

$$H_0: F(x)=F_0(x) \tag{5.3-223}$$

其中，已知的分布函数 $F_0(x)$ 可以根据样本的分布特征作出，并用极大似然估计法估计未知参数，未知参数的个数为 r。

已知实验数据是容量为 n 的样本，将总体 X 的取值范围分成 k 个区间，如果 H_0 成立，则可根据 $F_0(x)$ 计算出总体 X 在第 i 个区间（$i=1, 2, \cdots, k$）内取值的概率 p_i，而 np_i 就是总体 X 落入第 i 个区间内的理论频数。而实际 n 个观察值落入第 i 个区间内的实际频数为 f_i，则当 H_0 成立时，n 充分大（$n\to\infty$）时统计量

$$\sum_{i=1}^{k}\frac{(f_i-np_i)^2}{np_i}\sim\chi^2(k-r-1) \tag{5.3-224}$$

只要 $n>50$，就可以认为上述统计量近似服从 $\chi^2(k-r-1)$ 分布。

给定显著度，使得

$$P\Big[\sum_{i=1}^{k}\frac{(f_i-np_i)^2}{np_i}>\chi^2(k-r-1)\Big]=\alpha$$

若出现

$$\sum_{i=1}^{k}\frac{(f_i-np_i)^2}{np_i}>\chi^2(k-r-1) \tag{5.3-225}$$

则拒绝原假设 H_0；反之则接受原假设。

(2) 柯尔莫哥罗夫－斯米尔诺夫（$K-S$）检验

柯尔莫哥罗夫－斯米尔诺夫检验，简称 $K-S$ 检验。它比 χ^2 检验具有更高的灵敏度，还适合于小样本的情况。

需要检验的假设是

$$H_0: F(x)=F_0(x)$$

其中，$F_0(x)$ 是完全确定的不含任何未知参数的函数。

将样本 X_1，X_2，…，X_n 的各分量按大小重新排列 $X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$，样本的经验分布 $F_n(x)$ 为

$$F_n(x)=\begin{cases}0 & (X<X_{(1)})\\ i/n & (X_{(i)}\leqslant X\leqslant X_{(i+1)})\\ 1 & (X\geqslant X_{(n)})\end{cases} \tag{5.3-226}$$

柯氏统计量为

$$D_n=\mathrm{Sup}|F_n(x)-F_0(x)| \tag{5.3-227}$$

$|F_n(x)-F_0(x)|$ 的上确界可在 n 个点 $X_{(i)}$（$i=1, 2, \cdots, n$）处寻找。

给定显著度 α（$\alpha=0.05$ 或 $\alpha=0.1$）的柯氏检验临界值 $D_{n,\alpha}$ 可查表。

如果出现 $D_n>D_{n,\alpha}$ 时，拒绝原假设；反之则接受原假设。

3.1.9 连续分布的容忍限与容忍区间

设 $F(x)$ 是总体 X 的分布函数，X_1，X_2，…，X_n 是总体 X 的简单随机样本，要从样本 X_1，X_2，…，X_n 算出一个下限值 $I=I(X_1,X_2,\cdots,X_n)$，使得

$$F[I(X_1,X_2,\cdots,X_n)]\leqslant\beta$$

或

$$P(X\leqslant I)\leqslant\beta \tag{5.3-228}$$

如图 5.3-41 所示。但是由于样本的随机性，我们只能以一定的概率（或置信度）$1-\gamma$（通常 γ，β 为极小的正数）保证式（5.3-228）成立，要求

$$P\{F[I(X_1,X_2,\cdots,X_n)]\leqslant\beta\}\geqslant1-\gamma \tag{5.3-229}$$

如果对给定的（γ，β），某统计量 I 满足式（5.3-229），则称它是总体分布 F 的（γ，β）容忍下限。容忍下限与样本 X_1，X_2，…，X_n 和（γ，β）有关。

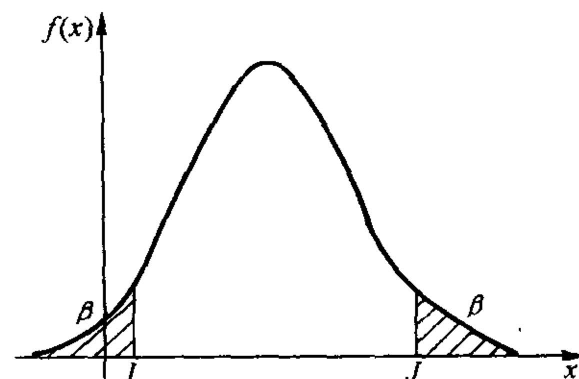


图 5.3-41 容忍限、容忍区间示意图

同样从样本 X_1，X_2，…，X_n 等计算出有关上限值 $J=J(X_1,X_2,\cdots,X_n)$ 使得

$$P\{F[J(X_1,X_2,\cdots,X_n)]\geqslant\beta\}\geqslant1-\gamma \tag{5.3-230}$$

如果对给定的（γ，β），某统计量 J 满足式（5.3-230），则称它是总体分布 F 的（γ，β）容忍上限。容忍上限 J 与样本 X_1，X_2，…，X_n 和（γ，β）有关。

如果区间（I，J）满足

$$P[F(J)-F(I)\geqslant1-\beta]\geqslant1-\gamma \tag{5.3-231}$$

则区间（I，J）称为总体分布 F 的容忍区间。

3.1.10 线性回归

(1) 一元线性回归

假设两个变量 x 与 y 之间线性相关，现由实验获得 x 和 y 的一组样本数据（x_i，y_i），记它们的线性关系如下：

$$y_i=a+bx_i+\varepsilon_i(i=1,2,\cdots,n,n>2) \tag{5.3-232}$$

式中，a、b 为待定估计量；ε_i 为独立等方差的正态随机变量，$\varepsilon_i\sim N(0,\sigma^2)$；$x_i$ 为普通自变量，如有随机性，则归于 ε_i 之中。于是，$y_i\sim N(a+bx_i,\sigma^2)$。

用线性最小二乘法确定待定系数 a 和 b，即令

$$\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}(y_i-a-bx_i)^2$$

得方程组

$$\begin{cases}1\times a+\overline{x}\times b=\overline{y}\\ \overline{x}\times a+\overline{x^2}\times b=\overline{xy}\end{cases}$$

求得 a、b 的估计值 $\hat{a}$、$\hat{b}$ 为

$$\begin{cases}\hat{b}=\dfrac{n(\overline{xy}-\overline{x}\overline{y})}{n(\overline{x^2}-\overline{x}\overline{x})}=\dfrac{l_{xy}}{l_{xx}}\\ \hat{a}=\overline{y}-\overline{x}\hat{b}\end{cases} \tag{5.3-233}$$

其中

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\overline{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$$

$$l_{xx}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2=n\overline{x^2}-n(\overline{x})^2$$

$$l_{xy}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})=n\overline{xy}-n\overline{x}\overline{y}$$

$$l_{yy}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\overline{y})^2=n\overline{y^2}-n(\overline{y})^2$$

由此得到线性回归方程为

$$\hat{y}=\hat{a}+\hat{b}x \tag{5.3-234}$$

(2) 方差的估计

根据线性回归假设 $y_i\sim N(a+bx_i,\sigma^2)$，线性回归方程为 $\hat{y}_i=\hat{a}+\hat{b}x_i$，$y_i-\hat{y}_i$ 称为 x_i 处的残差，残差平方和为

$$Q_e=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2=\sum_{i=1}^{n}(y_i-\hat{a}-\hat{b}x_i)^2$$

残差平方和服从分布

$$\frac{Q_e}{\sigma^2}\sim\chi^2(n-2)$$

于是

$$E\left(\frac{Q_e}{\sigma^2}\right)=n-2$$

因此 σ^2 的无偏估计量为

$$\hat{\sigma}^2=\frac{Q_e}{n-2}=\frac{1}{n-2}(l_{yy}-\hat{b}l_{xy}) \tag{5.3-235}$$

(3) 系数的置信区间

在线性回归假设中，我们假设 y_i 在每个 x_i 处都是正态分布，a，b 是 y_i 的线性函数，因而也是正态分布的随机变量。

由线性回归假设可证明

$$\frac{\hat{b}-b}{S_b}\sim t(n-2)$$

$$\frac{\hat{a}-a}{S_a}\sim t(n-2)$$

式中，a、b 的样本方差分别为

$$S_b^2=\frac{\hat{\sigma}^2}{l_{xx}}$$

$$S_a^2=\hat{\sigma}^2\left[\frac{1}{n}+\frac{(\overline{x})^2}{l_{xx}}\right]$$

因此系数 a、b 置信度为 $1-\alpha$ 的置信区间为

$$\left[\hat{b}\pm t_{\alpha/2}(n-2)\frac{\hat{\sigma}}{\sqrt{l_{xx}}}\right] \tag{5.3-236}$$

$$\left[\hat{a}\pm t_{\alpha/2}(n-2)\frac{\hat{\sigma}}{\sqrt{1/n+(\overline{x})^2/l_{xx}}}\right] \tag{5.3-237}$$

式中，$\hat{\sigma}=\sqrt{\hat{\sigma}^2}$。

(4) 线性回归的显著性检验

1) t 检验法　线性回归方程是否具有实用价值，一般来说，需要经过假设检验。

假设：$H_0:b=0$，$H_1:b\neq0$

若 $b=0$ 成立，则可知 $t=\dfrac{\hat{b}}{S_b}\sim t(n-2)$

给定显著性水平 α，出现

$$|t|=\left|\frac{\hat{b}}{S_b}\right|=\left|\frac{\hat{b}}{\hat{\sigma}}\sqrt{l_{xx}}\right|\geqslant t_{\alpha/2}(n-2) \tag{5.3-238}$$

则拒绝 H_0，此时 y 与 x 存在线性关系。否则 y 与 x 不存在线性关系。

2) F 检验法

假设：$H_0:b=0$，$H_1:b\neq0$

给定显著性水平 α，可证明

$$F=\frac{\hat{b}}{\hat{\sigma}^2}l_{xy}\sim F(1,\ n-2) \tag{5.3-239}$$

若出现 $F>F_\alpha(1,n-2)$，则存在线性关系。

F 检验法与 t 检验法是等价的。

(5) 回归预测区间

对于给定的点 $x=x_0$，可以以一定的置信度预测对应的 y 值的取值范围，称为预测区间。

设 y_0 是在 $x=x_0$ 处对随机变量 y 的预测结果，由线性回归假设 $y_0=a+bx_0+\varepsilon_0$，$\varepsilon_0\sim N(0,\sigma^2)$ 作为 y_0 的预测值。y_0 的预测值的均值为

$$\hat{y}_0=\hat{a}+\hat{b}x_0$$

可以证明 y_0，$\hat{y}_0$ 相互独立，且

$$\frac{y_0-\hat{y}_0}{\hat{\sigma}\sqrt{1+1/n+(x_0-\overline{x})^2/l_{xx}}}\sim t(n-2)$$

于是，对于给定置信度 $1-\alpha$，y_0 的置信度为 $1-\alpha$ 的预测区间为

$$\left|\hat{y}_0\pm t_{\alpha/2}(n-2)\hat{\sigma}\sqrt{1+1/n+(x_0-\overline{x})^2/l_{xx}}\right| \tag{5.3-240}$$

(6) 二元线性回归

1) 二元线性回归方程　对一组样本 x_{1i}，x_{2i}；y_i，设 x_{1i}，x_{2i} 与 y_i 之间线性相关。

$$y_i=b_0+b_1x_{1i}+b_2x_{2i}+\varepsilon_i\quad(i=1,\ 2,\ \cdots,\ n,\ n>3) \tag{5.3-241}$$

式中，b_0、b_1 和 b_2 为待定的估计量；ε_i 为 n 个相互独立和等方差的正态随机变量，$\varepsilon_i\sim N(0,\sigma^2)$。

按最小二乘法的原理求解得：

$$\begin{aligned}b_1&=\frac{l_{10}l_{22}-l_{20}l_{12}}{l_{11}l_{22}-l_{12}l_{21}}\\b_2&=\frac{l_{20}l_{11}-l_{10}l_{21}}{l_{11}l_{22}-l_{12}l_{21}}\\b_0&=\overline{y}-b_1\overline{x}_1-b_2\overline{x}_2\end{aligned} \tag{5.3-242}$$

式中

$$l_{11}=n(\overline{x_1^2}-(\overline{x}_1)^2)\qquad l_{12}=n(\overline{x_1x_2}-\overline{x}_1\ \overline{x}_2)$$

$$l_{10}=n(\overline{x_1y}-\overline{x}_1\ \overline{y})\qquad l_{22}=n(\overline{x_2^2}-(\overline{x}_2)^2)$$

$$l_{21}=l_{12}\qquad l_{20}=n(\overline{x_2y}-\overline{x}_2\ \overline{y})$$

因此得二元线性回归方程

$$\hat{y}=b_0+b_1x_1+b_2x_2 \tag{5.3-243}$$

2) 回归系数的显著性检验　作统计量

$$t_i=\sqrt{U_i}/S \tag{5.3-244}$$

式中

$$U_1=b_1^2\left(l_{11}-\frac{l_{12}}{l_{22}}\right)$$

$$U_2=b_1^2\left(l_{22}-\frac{l_{12}^2}{l_{11}}\right)$$

$$S=\sqrt{\frac{l_{xy}-b_1l_{10}-b_2l_{20}}{(n-3)}}$$

若 $t_i>2$，该因素显著；若 $t_i<1$，该因素可忽略，应重新进行一元线性回归。

3) 二元线性回归效果检验　与一元线性回归分析相类似，作统计量

$$F=\frac{(b_1l_{10}+b_2l_{20})/2}{(l_{yy}-b_1l_{10}-b_2l_{20})/(n-3)}\sim F(2,n-3) \tag{5.3-245}$$

给定显著度 α，若 $F>F_\alpha(2,n-3)$，则线性回归效果显著。

(7) 非线性回归分析

在工程分析中常遇到非线性函数，对于非线性函数可以通过线性变换，将它们变成线性函数进行线性回归。

例如指数函数 $y=ax^b$，两边取对数后变为

$$\ln y=\ln a+b\ln x$$

如果令 $Y=\ln y$，$A=\ln a$，$B=b$，$X=\ln x$，则变换后得

到的线性方程为

$$Y = A + BX$$

利用此方程，就可按上述线性回归法进行求解。

又如抛物线方程 $y = a + bx + cx^2$，如果令 $Y = y$，$X_1 = x$，$X_2 = x^2$，$A = a$，$B = b$，$C = c$，则可以化成包含两个自变量的线性方程 $Y = A + BX_1 + CX_2$，从而将抛物线拟合转化成二元线性回归处理。

3.2　失效分析随机过程基础

3.2.1　随机过程及其统计描述

(1) 随机过程的概念

随机过程是指这样一个过程：①是时间的变化过程；②相同条件下，每次试验结果大都不同，记为 $x_1(t), x_2(t), \cdots, x_n(t), \cdots$，并且不可能预先确定出现哪种试验结果 $x_i(t)(i=1,2,\cdots)$，每个试验结果称为一个样本函数；③在给定的每个时刻 t，随机过程 $X(t)$ 是随机变量，由所有样本函数 $x_1(t), x_2(t), \cdots, x_n(t), \cdots$ 在时刻 t 的函数值给出。

随机过程 $X(t)$ 对于 t_1，t_2，…，$t_n \in T$ 定义函数：

$$F_n(t_1, t_2, \cdots, t_n; x_1, x_2, \cdots, x_n) = P[X(t_1) < x_1, \cdots, X(t_n) < x_n]$$

称为随机过程 $X(t)$ 的 n 维分布函数。显然，随机过程可用一族分布函数 $\{F_1, F_2, \cdots, F_n, \cdots\}$ 确定。

(2) 随机过程的数字特征

随机过程的统计描述一般用数字特征。

随机过程 $\{X(t), t \in T\}$ 对给定的 $t \in T$，$X(t)$ 是一个随机变量，它的均值一般与 t 有关，记为

$$\mu_X(t) = E[X(t)] \tag{5.3-246}$$

称 $\mu_X(t)$ 为随机过程 $\{X(t), t \in T\}$ 的均值函数。

均值函数 $\mu_X(t)$ 表示了随机过程 $X(t)$ 在各个给定时刻的摆动中心。如图5.3-42所示。

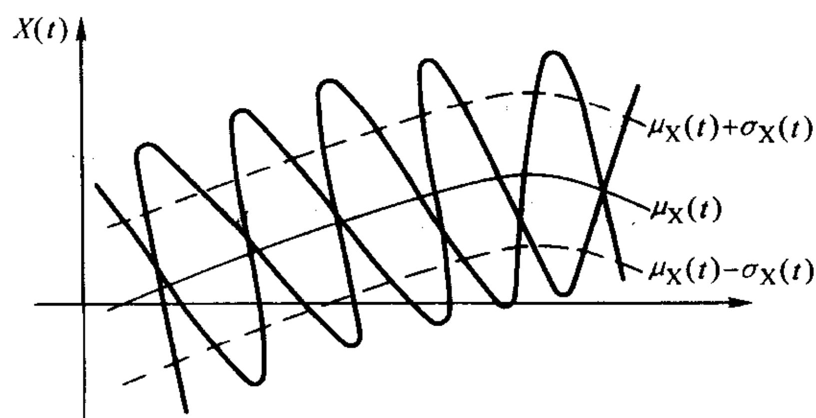


图 5.3-42　均值函数示意图

$\mu_X(t)$ 是随机过程的所有样本函数在时刻 t 的函数值平均值。通常，这种给定时刻所有样本函数值的平均称为集平均或统计平均。还有一种平均的方法是一个样本函数在整个时间轴上的平均，称为时间平均。

$X(t)$ 的二阶原点矩和二阶中心矩分别记作

$$\psi_X^2(t) = E[X^2(t)] \tag{5.3-247}$$

$$\sigma_X^2(t) = D_X(t) = E\{[X(t) - \mu_X(t)]^2\} \tag{5.3-248}$$

它们分别称为随机过程 $\{X(t), t \in T\}$ 的均方值函数和方差函数。$\sigma_X^2(t)$ 表示随机过程 $X(t)$ 在时刻 t 对于均值 $\mu_X(t)$ 的偏离程度。

对任意给定 t_1，$t_2 \in T$，我们把随机变量 $X(t_1)$ 和 $X(t_2)$ 的二阶原点混合矩记作

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] \tag{5.3-249}$$

它称作随机过程 $\{X(t), t \in T\}$ 的自协方差函数，或简称相关函数。类似地还可以写出 $X(t_1)$ 和 $X(t_2)$ 的二阶中心混合矩：

$$C_X(t_1, t_2) = \mathrm{cov}[X(t_1)X(t_2)] = E\{[X(t_1) - \mu(t_1)][X(t_2) - \mu(t_2)]\} \tag{5.3-250}$$

它称作随机过程 $\{X(t), t \in T\}$ 的自协方差系数，或简称协方差系数。

对于任意两个固定时刻 t_1，$t_2 \in T$，$X(t_1)$ 和 $X(t_2)$ 是随机变量，它们之间线性联系的密切程度可用相关系数描述，即：

$$\rho(t_1, t_2) = \frac{\mathrm{cov}[X(t_1), X(t_2)]}{\sqrt{D_X(t_1)}\sqrt{D_X(t_2)}} \quad (t_1, t_2 \in T) \tag{5.3-251}$$

实际上，随机过程的基本数字特征为均值函数和相关函数，其他数字特征都可以由它们表示：

$$\psi_X^2(t) = R_X(t, t)$$

$$C_X(t_1, t_2) = R_X(t_1, t_2) - \mu_X(t_1)\mu_X(t_2)$$

$$\sigma_X^2(t) = C_X(t, t) = R(t, t) - \mu_X^2(t)$$

$$\rho(t_1, t_2) = \frac{R_X(t_1, t_2) - \mu_X(t_1)\mu_X(t_2)}{\sqrt{R_X(t_1, t_1) - \mu_X^2(t_1)}\sqrt{R_X(t_2, t_2) - \mu_X^2(t_2)}}$$

3.2.2　平稳随机过程和各态历经性

(1) 平稳随机过程

设 $X(t)$ 为一随机过程，若对于任意 t_1，t_2，…，$t_n \in T$，以及任意满足 $t_i + \tau \in T(i = 1, 2, \cdots, n)$ 的 τ，有

$$F_n(t_1, t_2, \cdots, t_n; x_1, x_2, \cdots, x_n) = F_n(t_1 + \tau, t_2 + \tau, \cdots, t_n + \tau; x_1, x_2, \cdots, x_n) \tag{5.3-252}$$

则称该过程为平稳随机过程或严格平稳随机过程。

式(5.3-252)表示平稳过程的所有各维分布函数在时间参数 t 上作任意平移保持不变。若令 $\tau = -t_1$，则得

$$F_n(t_1, t_2, \cdots, t_n; x_1, x_2, \cdots, x_n) = F_n(0, t_2 - t_1, \cdots, t_n - t_1; x_1, x_2, \cdots, x_n)$$

即分布函数只与时间参数有关。

设随机过程，如果对任意 t，$t + \tau \in T$，满足

$$\begin{cases} E[X(t)] = \mu_X \\ E[X(t)X(t + \tau)] = R_X(\tau) \end{cases} \tag{5.3-253}$$

则称随机过程为宽平稳随机过程。

一个严平稳随机过程只要二阶矩存在，则它必定也是宽平稳随机过程。但是反过来一般是不成立的。不过一个宽平稳的正态过程必定是严平稳的。在实际应用中，平稳正态随机过程应用最多。

(2) 各态历经性

如果平稳随机过程 $X(t)$ 使得

$$\langle X(t) \rangle = E[X(t)] = \mu_X \tag{5.3-254}$$

以概率1成立，则称随机过程 $X(t)$ 的均值具有各态历经性。

式中〈·〉表示平均，“以概率1成立”是指对 $X(t)$ 的所有样本函数都成立。

如果平稳随机过程 $X(t)$ 对于任意实数 τ

$$\langle X(t)X(t + \tau) \rangle = E[X(t)X(t + \tau)] = R_X(\tau) \tag{5.3-255}$$

以概率1成立，则称随机过程 $X(t)$ 的自相关函数具有各态历经性。

如果平稳随机过程 $X(t)$ 的均值和自相关函数都具有各态历经性，则称 $X(t)$ 是各态历经过程，或者说 $X(t)$ 是各态历经的。

各态历经性的重要价值在于从理论上给出了如下保证：从一次试验所得到的样本函数 $x(t)$ 来确定该过程的均值和自相关函数。

如果试验记录 $x(t)$ 只是在时间区间 $[0, T]$ 上给出，则相应的均值和自相关函数的无偏估计式为

$$\hat{\mu}_X = \frac{1}{T}\int_0^T x(t)\mathrm{d}t \tag{5.3-256}$$

$$\hat{R}_X(\tau) = \frac{1}{T}\int_0^{T-\tau} x(t)x(t+\tau)\mathrm{d}t = \frac{1}{T}\int_{\tau}^{T} x(t)x(t-\tau)\mathrm{d}t$$
$$(0 \leqslant \tau < T) \quad (5.3\text{-}257)$$

具体数值计算方法如下：把 $[0,\ T]$ 等分为 N 个长度为 $\Delta t = T/N$ 的小区间，$t_k = k\Delta t = kT/N$，$k=1,\ 2,\ \cdots,\ N$，$\tau_r = r\Delta t = rT/N$，则

$$\hat{u}_X = \frac{1}{T}\sum_{k=1}^{N} x_k \Delta t = \frac{1}{N}\sum_{k=1}^{N} x\left(k\frac{T}{N}\right) \quad (5.3\text{-}258)$$

$$\hat{R}_X = \left(r\frac{T}{N}\right) = \frac{1}{N-r}\sum_{k=1}^{N} x\left(k\frac{T}{N}\right)x\left((k+r)\frac{T}{N}\right) \quad (5.3\text{-}259)$$

式中，$r=0,\ 1,\ 2,\ \cdots,\ m$，$m<N$，通常取 $m=N/5 \sim N/2$。

自相关函数 $\hat{R}(\tau)$ 如图 5.3-43 所示。

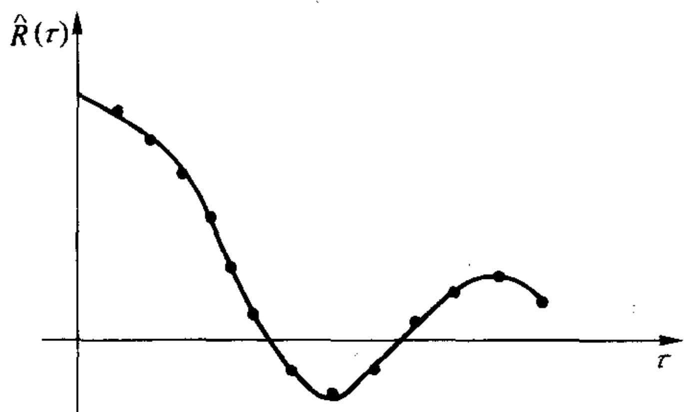


图 5.3-43 自相关函数 $\hat{R}(\tau)$ 示意图

3.2.3 平稳随机过程的自相关系数与功率谱密度

用数字特征来描述随机过程，比用分布函数或概率密度函数方便，特别是正态平稳过程，它的均值 μ_X 和自相关函数 $R_X(\tau)$ 完全刻画了该过程的统计特性，又根据各态历经性，通过对随机过程的一个样本函数的数值分析就可求得它的均值和自相关函数。为了成功使用数字特征研究随机过程，有必要进一步研究自相关函数的性质。

(1) 自相关函数的性质

平稳随机过程 $X(t)$ 的自相关函数 $R_X(\tau)$ 具有如下性质。

1) $R_X(0) = E[X^2(t)] = \psi_X^2 \geqslant 0$

式中，$R_X(0)$ 称为平稳过程的“平均功率”。

2) $R_X(-\tau) = R_X(\tau)$

即 $R_X(\tau)$ 是 τ 的偶函数。根据这个性质，在实际问题中只需要计算或测量 $R_X(\tau)$ 在 $\tau \geqslant 0$ 的值。

3) $|R_X(\tau)| \leqslant R_X(0)$ 或 $|C_X(\tau)| \leqslant C_X(0) = \sigma_X^2$（常数）

4) $R_X(\tau) = \sigma_X^2\rho + \mu_X^2$，$\rho$ 为和 $X(t+\tau)$ 的相关系数。因为 $|\rho| \leqslant 1$，故

$$-\sigma_X^2 + \mu_X^2 \leqslant R_X(\tau) \leqslant \sigma_X^2 + \mu_X^2$$

如图 5.3-44 所示。

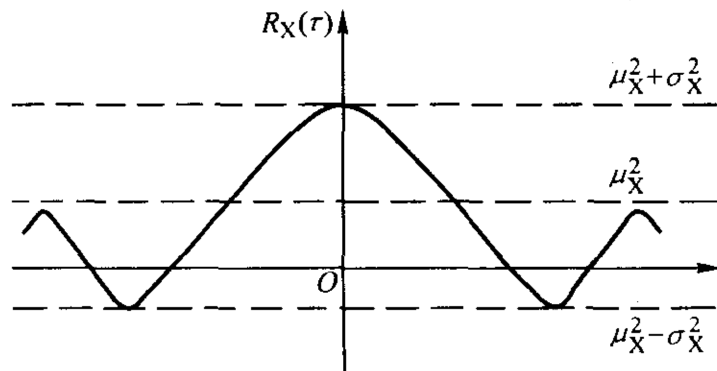


图 5.3-44 平稳过程自相关函数特性图

(2) 功率谱密度性质

设平稳过程 $X(t)$ 的相关函数 $R_X(\tau)$，$R_X(\tau)$ 的傅立叶变换为

$$S_X(\omega) = \int_{-\infty}^{+\infty} \mathrm{e}^{-i\tau\omega} R_X(\tau)\mathrm{d}\tau \quad (-\infty < \omega < +\infty)$$
$$(5.3\text{-}260)$$

式中，$S_X(\omega)$ 称为平稳随机过程在 ω 处的功率谱密度。傅立叶反变换得

$$R_X(\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{-i\tau\omega} S_X(\omega)\mathrm{d}\omega \quad (-\infty < \omega < +\infty)$$
$$(5.3\text{-}261)$$

或表示为

$$S_X(\omega) = 2\int_0^{+\infty} R_X(\tau)\cos(\omega\tau)\mathrm{d}\tau$$

$$R_X(\tau) = \frac{1}{\pi}\int_0^{+\infty} S_X(\omega)\cos(\omega\tau)\mathrm{d}\tau$$

$S_X(\omega)$ 是 ω 的实的、非负的偶函数。

由式（5.3-261）和各态历经性式（5.3-255）可知

$$R_X(0) = \lim_{T\to\infty}\frac{1}{T}\int_0^T x^2(t)\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_X(\omega)\mathrm{d}\omega$$
$$(5.3\text{-}262)$$

若将 $x(t)$ 看作电流，则 $x^2(t)$ 就是这个电流在电阻上产生的瞬时功率；而 $R_X(0)$ 就是电流的平均功率，由 $S_X(\omega)$ 与 x 轴所围成的面积表示。

以上定义的谱密度称为“双边谱密度”，对 ω 的正负值都是有意义的；而在实际应用中平稳过程 $X(t)$，定义在 $t \geqslant 0$，此时定义“单边谱密度”

$$G_X(\omega) = \begin{cases} 2S_X(\omega) & (\omega > 0) \\ 0 & (\omega < 0) \end{cases} \quad (5.3\text{-}263)$$

这相当于利用 $S_X(\omega)$ 的偶函数性质，把负频率范围内的谱密度折算到正频率范围内，用图 5.3-45 表示。

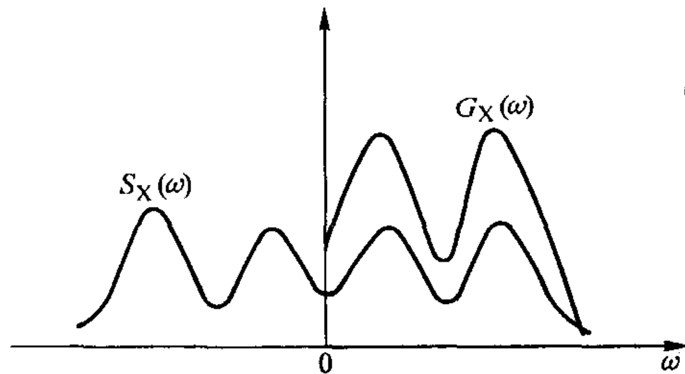


图 5.3-45 单边谱密度示意图

(3) 窄带过程和宽带随机过程

1) 窄带过程 频率的主要成分集中在很窄的一段频率范围内，称为窄带过程。此时随机过程频率变化较小，只是振幅和相位角随机变化。如图 5.3-46 所示。

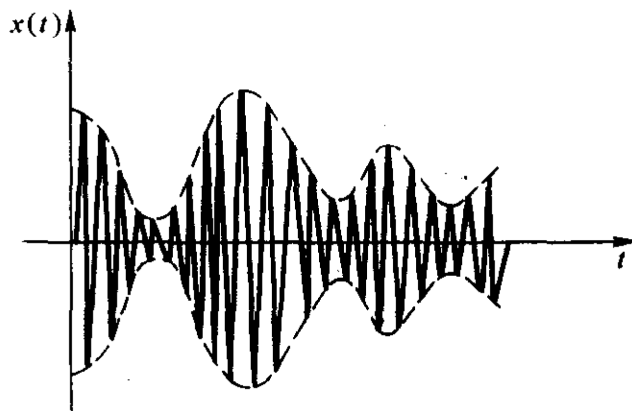


图 5.3-46 窄带平稳随机过程

2) 宽带过程 谱密度占有一个宽的频带。此时随机过程频率变化较大，是这频带内的各种频率叠加的结果。如图 5.3-47 所示。

(4) 有关随机过程的重要积分

若随机过程 $X(t)$ 在区间 $[a,\ b]$ 上均方连续，则 $X(t)$ 在 $[a,\ b]$ 上均方可积，并且

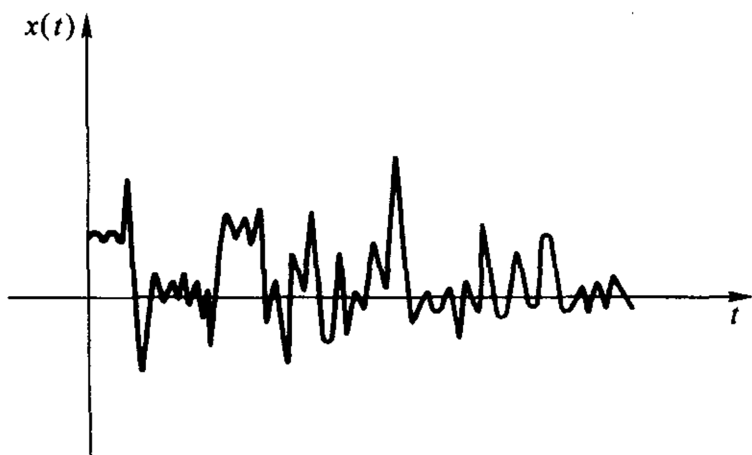


图 5.3-47 宽带平稳随机过程

$$\int_a^b \alpha X(t)\mathrm{d}t = \alpha\int_a^b X(t)\mathrm{d}t \qquad (\alpha\text{ 是常数})$$

$$\int_a^b f(t)X\mathrm{d}t = X\int_a^b f(t)\mathrm{d}t \qquad (X\text{ 是随机变量})$$

$$E\left[\int_a^b f(t)X(t)\mathrm{d}t\right] = \int_a^b f(t)E[X(t)]\mathrm{d}t$$

$$E\left|\int_a^b f(t)X(t)\mathrm{d}t\right|^2 = \int_a^b\int_a^b f(S)f(t)R_X(s,t)\mathrm{d}s\mathrm{d}t$$

3.2.4 快速傅立叶变换计算自相关系数

(1) 离散傅立叶变换

$x(t)$是随机过程$X(t)$的一个样本函数，样本序列$\{x(j\Delta t)\}$简记为$\{x(j)\}$，$\Delta t = T/N$的采样间隔，$j=1, 2, \cdots, N-1$。

样本函数$x(j)$的有限离散傅立叶变换（简记为 DFT）为

$$A(k) = \frac{1}{N}\sum_{k=0}^{N-1}x(j)W^{jk} \quad (k=0,1,\cdots,N-1) \tag{5.3-264}$$

式中，$W=\exp(-2\pi i/N)$；$i=\sqrt{-1}$。

有限离散傅立叶反变换为

$$x(j) = \sum_{k=0}^{N-1}A(k)W^{-jk} \quad (j=0,1,\cdots,N-1) \tag{5.3-265}$$

式中，序列$x(j)$和$A(k)$称为傅立叶变换对。

快速傅立叶变换（FFT）是计算有限离散傅立叶变换的有效方法，而且有现成的计算程序可供参考。

(2) FFT 算法计算自相关系数

如果平稳随机过程的均值为$\mu_X \neq 0$，则减去均值后的平稳随机过程$Y(t)=X(t)-\mu_X$，满足

$$E[Y(t)]=0$$
$$\sigma_Y^2=\sigma_X^2$$
$$C_Y(\tau)=R_Y(\tau)=C_X(\tau)-\mu_X^2$$

只要求得$Y(t)$的自相关函数，那么$X(t)$的自相关函数为$R_X(\tau)=R_Y(\tau)+\mu_X^2$。实际讨论时，为了方便，只讨论均值为零的平稳随机过程；对于非零均值的平稳随机过程，可用上述方法变化为零均值随机过程。

对于离散样本序列$\{x(j)\}$，$j=0, 1, \cdots, N-1$，$\Delta t = T/N$为采样间隔。如果$x(j)$已经减去平均值$\bar{x} = \frac{1}{N}\sum_{j=0}^{N-1}x(j)$，则随机过程$X(t)$的自相关函数为

$$R(p)\frac{1}{N}\sum_{j=0}^{N-1-p}x(j+p)x(j) \quad (p=0,1,\cdots,N-1) \tag{5.3-266}$$

显然，当N很大时，计算自相关函数需要作很多次乘法和加法运算，从而占用较多的计算时间。采用 FFT 算法可以使运算量减少。

使用 FFT 计算自相关函数$R(p)$的具体步骤如下。

1) 假设样本序列$\{x(j)\}$具有$N=2^m$个点（m是大于 1 的正整数），一般取$N=16, 32, 64, 128, \cdots$，此时$m=4, 5, 6, 7, \cdots$，并且$\{x(j)\}$已经减去平均值$\bar{x}$，这时在$\{x(j)\}$的末尾补加$N$个零；

2) 使用 FFT 变换，计算有限离散傅立叶变换$\{A(k)\}$，其中，$k=0, 1, \cdots, 2N-1$；

3) 使用 FFT 变换，计算$|A(k)|^2$的有限离散傅立叶变换得$\{R^c(p)\}$，其中，$p=0, 1, \cdots, 2N-1$；

4) 取$\{R^c(p)\}$的前N个点得$R(p)$，即$R(p)=R^c(p)$。

3.2.5 快速傅立叶变换计算功率谱密度

直接从样本序列的傅立叶变换（DFT）来估计功率谱密度是不能令人满意的，这是因为它的方差不随记录长度的增加而减少，必须作一些修正。

(1) 功率谱密度的样本谱

在有限时间区间$0\leqslant t\leqslant T$内，考虑平稳随机过程$X(t)$，如果它的数学期望为零，即$E[X(t)]=0$，则自相关函数为

$$R_X(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T x(t+\tau)x(t)\mathrm{d}t$$

式中，$x(t)$为随机过程$X(t)$的一个样本函数。

根据纳－辛钦公式，平稳随机过程的自相关函数$R_X(\tau)$与功率谱密度$S_X(\omega)$互为傅立叶变换：

$$S_X(\omega) = \int_{-\infty}^{+\infty}R_X(\tau)e^{-2\pi i\omega\tau}\mathrm{d}\tau \tag{5.3-267}$$

$$R_X(\tau) = \int_{-\infty}^{+\infty}S_X(\omega)e^{2\pi i\omega\tau}\mathrm{d}\omega \tag{5.3-268}$$

或者

$$S_X(\omega) = \int_{-\infty}^{+\infty}R_X(\tau)e^{-i\omega\tau}\mathrm{d}\tau$$

$$R_X(\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}S_X(\omega)e^{i\omega\tau}\mathrm{d}\omega$$

由式（5.3-267）和式（5.3-268）可知，随机过程$X(t)$的方差为

$$Var[X(t)] = R(0) = \lim_{T\to\infty}\frac{1}{T}\int_0^T x^2(t)\mathrm{d}t = \int_{-\infty}^{+\infty}S_X\mathrm{d}\omega$$

假设样本函数的有限离散傅立叶变换$X(\omega)$存在，则有下面的等式。

$$X(\omega) = \frac{1}{T}\int_0^T x(t)e^{-2\pi i\omega t}\mathrm{d}t$$

$$x(t) = \int_0^T X(\omega)e^{2\pi i\omega t}\mathrm{d}\omega$$

$T|X(\omega)|^2$称为样本谱，在$T\to\infty$时，可以证明样本谱的数学期望就是功率谱密度。

$$\lim_{T\to\infty}E[T|X(\omega)|] = S_X(\omega)$$

考虑有限时间区间$[0, T]$内的离散样本序列$\{x_j\}$，$x_j = x(j\Delta t) = x(t_j)$，$j=0, 1, \cdots, N-1$，$\Delta t = T/N$，$\{x_j\}$的有限离散傅立叶变换（DFT）是

$$X_k = \frac{1}{N}\sum_{j=0}^{N-1}x_j e^{-2\pi ikj/N}$$

式中，$X_k = X(k\Delta f) = X(f_k)$，$\Delta f = 1/T$，$f_k = k/T$，并且有

$$-\frac{N}{2}\leqslant k\leqslant\frac{N}{2}$$

即

$$-\frac{1}{2\Delta t}\leqslant f_k\leqslant\frac{1}{2\Delta t}$$

式中，$1/(2\Delta t)$叫做 Nyquist 频率。根据采样定理，它是采样间隔为Δt的序列所具有的最高频率。

这时样本频率为

$$T|X_k|^2 = T|X(f_k)|^2 = N\Delta t\left|\frac{1}{N}\sum_{j=0}^{N-1}x_j e^{-2\pi ikj/N}\right|^2 =$$

$$\frac{\Delta t}{N}\left[\left(\sum_{j=0}^{N-1}x_j\cos\frac{2\pi jk}{N}\right)^2+\left(\sum_{j=0}^{N-1}x_j\sin\frac{2\pi jk}{N}\right)\right] \tag{5.3-269}$$

样本谱 $T|X_k|^2$ 是功率谱密度的无偏估计：但是样本谱的方差大，而且不随样本长度的增加而减少，意味着其方差不能依靠增加样本序列长度来改善。

(2) 数据窗

设 $x_T(t)$ 是区间 $-T/2\leqslant t\leqslant T/2$ 上观测到的信号，那么对 $x_T(t)$ 进行傅立叶变换将会引起截断误差。为了看出这个影响，考虑函数：

$$w(t)=\begin{cases}1 & (-T/2\leqslant t\leqslant T/2)\\ 0 & (\text{其他})\end{cases} \tag{5.3-270}$$

上面的函数通常称为矩形数据窗。如果 $x(t)$ 是在无穷区间上存在的信号，则

$$x_T(t)=x(t)w(t)$$

于是有限长度的信号 $x_T(t)$ 和 $x(t)$ 矩形数据窗 $w(t)$ 的乘积。

假设 $x(t)$ 和 $w(t)$ 的傅立叶变换分别为 $X(\omega)$ 和 $W(\omega)$，由卷积定理可以得到 $x_T(t)$ 的傅立叶变换为

$$X_T(\omega)=X(\omega)*W(\omega)=\int_{-\infty}^{+\infty}X(g)W(\omega-g)\mathrm{d}g \quad (-\infty<\omega<\infty) \tag{5.3-271}$$

对于矩形数据窗有

$$W(\omega)=T\frac{\sin(\pi\omega T)}{\pi\omega T}\quad(-\infty<\omega<\infty) \tag{5.3-272}$$

$W(\omega)$ 称为频率窗，其图形如图 5.3-48 所示。

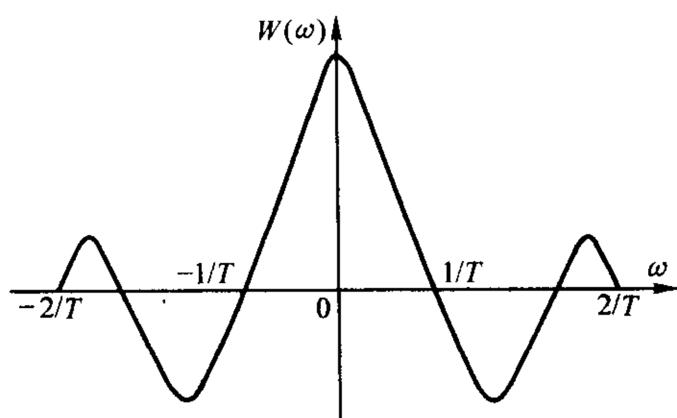


图 5.3-48　矩形数据窗的频率窗

频率窗有如下的性质：

1) $\int_{-\infty}^{+\infty}W(\omega)\mathrm{d}\omega=1$

2) $W(\omega)=W(-\omega)$

3) $W(\omega)$ 的基本宽度是 $2/T$。基本宽度是指 $W(\omega)$ 主峰两边第一个零点之间的距离。

从式 (5.3-271) 和式 (5.3-252) 可以看出，对于较小的 T 而言，由于频率窗的基本宽度较大，离某一频率 ω 较远的频率 g 上的 $X(g)$ 值也将影响 $X_T(\omega)$，从而降低了对 $X(\omega)$ 的分辨能力，$X_T(\omega)$ 中出现了相邻频率分量的干扰。当 T 增大时，基本宽度减小，使得当 $\omega_1\neq\omega_2$ 时，对 $X(\omega_1)$ 和 $X(\omega_2)$ 的分辨能力得到提高。

另一个问题是由于边瓣的影响。在某一频率 ω 处卷积的结果必然也会使其他频率成分不同程度地“泄漏”过来，这种泄漏现象的严重程度主要取决于谱窗的边瓣。

所以，为了提高分辨能力和削弱泄漏现象，要求主瓣宽度要足够窄和减小边瓣的幅度，即一个合适的数据窗的主瓣应集中在 $\omega=0$ 附近，而且边瓣将随 $|\omega|$ 的增大而衰减。

这里介绍两种常用数据窗：

$$C^{(1)}(t)=\begin{cases}1 & (-4T/2\leqslant t\leqslant 4T/10)\\ \left(1+\cos\frac{10\pi t}{T}\right)/2 & (-T/2\leqslant t<-4T/10,\ 4T/10<t\leqslant T/2)\\ 0 & (\text{其他})\end{cases}$$

$$C^{(2)}(t)=\begin{cases}\left(1+\cos\frac{10\pi t}{T}\right)/2 & (-T/2\leqslant t\leqslant T/2)\\ 0 & (\text{其他})\end{cases}$$

如图 5.3-49 所示。

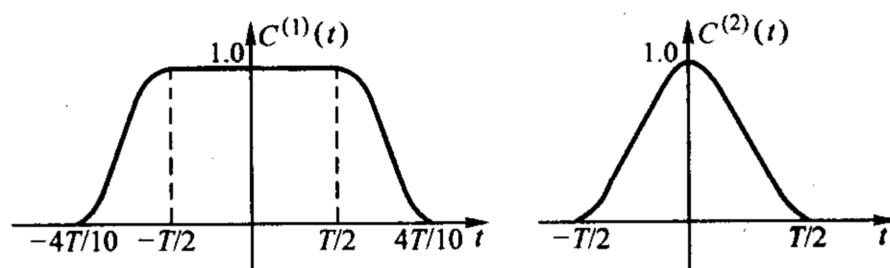


图 5.3-49　两种数据窗示意图

令

$$C=\int_{-\infty}^{+\infty}C^2(t)\mathrm{d}t\Big/\int_{-\infty}^{+\infty}w^2(t)\mathrm{d}t \tag{5.3-273}$$

式中，$w(t)$ 是矩形数据窗，$C(t)$ 是实际要用的数据窗。若 $C(t)=C^{(1)}(t)$，则 $C=0.875$；若 $C(t)=C^{(2)}(t)$，则 $C=0.375$。

(3) FFT 计算功率谱密度

设 $\{x_j\}$ 是点数为 N 的样本序列，由式 (5.3-269) 的“双边谱密度”得到“单边谱密度”估计

$$S=S(f_k)=2N\Delta t\left|\frac{1}{N}\sum_{j=0}^{N-1}x_j\mathrm{e}^{-2\pi ijk/N}\right|^2 \tag{5.3-274}$$

式中，$\Delta t=T/N, f_k=k/T=k/(N\Delta t), 0\leqslant f_k\leqslant 1/(2\Delta t), k=0,1,\cdots,N/2$。

这是谱估计的基础。FFT 程序通常要求 $N=2^m$（m 为大于 1 的正整数），如果给出的初始样本序列长度不符合这一要求时，应通过截断或补加零的办法使之成为 $N=2^m$ 的形式。

具体过程如下。

1) 选取数据窗（比如 $C^{(1)}(t)$）或 $C^{(2)}(t)$ 乘样本序列得到一个新序列 $\{x_j\}$。

2) 用 FFT 算法计算 $\{x_j\}$ 的有限离散傅立叶变换为

$$A(k)=\frac{1}{N}\sum_{j=0}^{N-1}x_j\mathrm{e}^{-2\pi ijk/N}\qquad k=0,1,\cdots,N/2$$

3) 由式 (5.3-274) 计算单边功率谱密度估计为

$$S_k=S(f_k)=2T|A(k)|^2\qquad k=0,1,\cdots,N/2$$

4) 如果数据窗是 $C^{(1)}(t)$，则令 $S_k=S_k/0.875$；如果数据窗是 $C^{(2)}(t)$，那么令 $S_k=S_k/0.375$；如果是其他数据窗，则 $S_k=S_k/C$，C 由式 (5.3-274) 确定；

5) 上面得到的 S_k 只是功率谱密度的粗糙估计，为了减少谱估计的方差，还需进一步处理。这种处理分为初始样本序列的分段平滑和频率平滑。现分述如下。

初始样本序列的分段平滑法：如果样本序列比较长，共分 L 段，则信号总长为 $T=LT_S$，总点数 $N=LN_S$，每段满足 $N_S=2^m$。假设第 l 段粗糙功率谱估计为

$$S_k^{(l)}, l=1,2,\cdots,L; k=0,1,\cdots,N_S/2$$

则平滑估计为

$$\overline{S}_k=\frac{1}{L}\sum_{l=0}^{L}S_k^{(l)}, k=0,1,\cdots,N_S/2$$

式中，$\overline{S}_k$ 是功率谱密度的估计。

频率平滑法：对 S_k 的 L 个相邻频率分量平均得平滑功率谱密度估计

$$S_{k+L/2} = \frac{1}{L}\sum_{l=0}^{L} S_{k+l}, k = 0, L, 2L, \cdots, [N/(2l)]L$$

3.3 失效分析中随机变量和随机过程的模拟方法

随机变量和随机过程的模拟是 Monte - Carlo 方法的重要组成部分。这里给出几种常见分布的随机数的产生方法和随机过程的模拟方法。

3.3.1 （0，1）随机数及其检验

具有$F(x)$分布的随机变量的一列独立样本值称为$F(x)$分布随机数，因为最基本的是（0，1）区间上分布的随机数，所以它又简称为随机数。（0，1）区间上均匀分布的随机数是产生其他分布随机数以及随机过程的基础，为了保证随机变量和随机过程的模拟质量，要对均匀分布随机数统计性质进行检验。

（1）伪随机数及其产生方法

随机数可以通过在计算机上附加一些物理设备来产生，这种设备称为随机数发生器。在进行随机性模拟的时候，人们通常采用在计算机上产生伪随机数的方法。

所谓伪随机数是在数字计算机上采用某种完全确定的规则，通过递推运算而产生的一列数。这一列数不是由真实的随机现象所产生的，因而不是真正的随机数。但由于这种数列具有类似于随机数的统计性质，因此，可以把它当作随机数来运算，故这种数列就称为伪随机数。

一般来说伪随机数应满足如下要求：

1）具有较好的随机性与均匀性；

2）产生伪随机数的速度要快；

3）算法程序应尽量少占用内存单元；

4）一批随机数的周期尽可能长。

（2）（0，1）区间上均匀分布的随机数

产生均匀随机数的方法较多，下面介绍字长为32位的计算机上易于实现的一种方法。

$$x_i = 314\,159\,269 x_{i-1} + 453\,806\,245 (\mathrm{mod} 2^{31}) \quad (5.3\text{-}275)$$

$$r_i = x_i/2^{31} (i = 1, 2, \cdots) \quad (5.3\text{-}276)$$

由于$2^{31}-1$是32位字长的计算机所能表示的最大正整数，所以随机数列的周期约为$2^{31}-1\approx 10^9$，即所产生的随机数个数约10^9个。

（3）（0，1）均匀分布随机数的检验

在计算机上用迭代方法产生的随机序列$\{r_i\}$作为均匀分布随机数使用之前，需要检验其随机性和均匀性。

1）均匀分布的假设检验

① χ^2 检验　将［0，1］区间等分成 m 个子区间，随机序列 r_1，r_2，…，r_N 种落入第 i 个子区间的数值个数为n_i $(i=1,2,\cdots,m)$，r_i 值落入各个子区间的理论概率 $P_i=1/m$，$(i=1, 2, \cdots, m)$。如果$\{r_i\}$是均匀分布的，则统计量

$$\frac{m}{N}\sum_{i=1}^{n}\left(n_i - \frac{N}{m}\right)^2 \sim \chi^2(m-1) \quad (5.3\text{-}277)$$

按式（5.3-147）可进行总体分布的假设检验。

② 柯尔莫哥罗夫－斯米尔诺夫检验　将被检验的 N 个随机数从小到大排列

$$r_1, r_2, \cdots, r_i, \cdots, r_N$$

统计量

$$D = \max_{1\leqslant i\leqslant N}\left\{\left|\frac{i}{N} - r_i\right|, \left|\frac{i-1}{N} - r_i\right|\right\} \quad (5.3\text{-}278)$$

当 $N\to\infty$时，D_N 的极限分布为

$$\lim_{N\to\infty} P(\sqrt{N}D_N \leqslant \lambda) = \sum_{k=-\infty}^{\infty}(-1)^k \cdot e^{-2k^2\lambda^2} = Q(\lambda)$$

当给定显著性水平 α，则查表求得 λ_α，使得

$$P(\sqrt{N}D_N > \lambda_\alpha) = \alpha$$

如果$\sqrt{N}D_N > \lambda_\alpha$，拒绝均匀性假设；如果$\sqrt{N}D_N < \lambda_\alpha$，接受均匀性假设。

2）参数的显著性检验　（0，1）区间上均匀分布随机变量的数学期望和方差分别为

$$E(r) = \frac{1}{2}, \sigma^2(r) = \frac{1}{12}$$

对于假设

$$H_0: E(r) = 1/2$$

统计量

$$u = \frac{\bar{r} - 1/2}{\sigma(r)/\sqrt{N}} = \sqrt{12N}(\bar{r} - 1/2) \quad (5.3\text{-}279)$$

渐近服从标准正态分布，式中

$$\bar{r} = \frac{1}{N}\sum_{i=1}^{N} r_i$$

在 H_0 成立的条件下，若取显著性水平 $\alpha = 0.05$，则检验的拒绝域为$|u|\geqslant 1.96$。

3）独立性检验　一组随机数 r_1，r_2，…，r_N，若前后相距为 j 的随机数是相互独立的，则前后为 j 的相互系数为

$$\hat{\rho} = \left[\frac{1}{N-j}\sum_{i=1}^{N-j} r_i r_{j+1} - (\bar{r})^2\right] \Big/ S, (j = 1, 2, \cdots) \quad (5.3\text{-}280)$$

对于充分大的 N（如 $N-j>50$），相关系数取零，假设 $H_0:\rho=0$，则统计量

$$u = \hat{\rho}_j\sqrt{N-j} \quad (5.3\text{-}281)$$

渐近服从标准正态分布。式中

$$S^2 = \frac{1}{N}\sum_{i=1}^{N}(r_i - 1/2)^2 \quad (5.3\text{-}282)$$

在 H_0 成立的条件下，取显著性水平 α，可进行假设检验。

3.3.2 常见连续型随机变量的模拟

（1）逆变换法

定理：如果随机变量 X 的分布函数$F(x)$连续，则$R=F(x)$是区间（0，1）上均匀分布的随机变量。

1）区间［a，b］上均匀分布的随机数　设 X 为［a，b］上均匀分布的随机变量，其分布函数为

$$F(x) = \frac{x-a}{b-a}$$

如果 R 为（0，1）上均匀分布的随机变量，则

$$R = \frac{x-a}{b-a} \quad (5.3\text{-}283)$$

$$x = a + (b-a)R$$

式（5.3-254）就是［a，b］上均匀分布随机变量 X 的抽样公式。

2）指数分布的随机数　若随机变量 X 的概率密度函数为

$$f(x) = \lambda e^{-\lambda x} (x\geqslant 0, \lambda > 0)$$

其分布函数为

$$F(x) = \int_0^x f(x)\mathrm{d}x = 1 - e^{-\lambda x}$$

由于

$$R = 1 - e^{-\lambda x}, \ 1 - R = e^{-\lambda x}$$

$$\ln(1-R) = -\lambda x$$

$$x = -\frac{1}{\lambda}\ln(1-R)$$

又在区间（0，1）上 R 与$1-R$同分布，故得指数分布随机变量的抽样公式为

$$x = -\frac{1}{\lambda}\ln R \quad (5.3\text{-}284)$$

3) 威布尔分布的随机数 若威布尔分布的概率密度函数为

$$f(x)=\frac{\beta}{\alpha}(x-\gamma)^{\beta-1}\exp\left[-\frac{(x-\gamma)}{\alpha}\right](x\geqslant\gamma)$$

分布函数为

$$F(x)=\int_{\gamma}^{x}f(x)\mathrm{d}x=1-\exp\left[-\frac{(x-\gamma)^{\beta}}{\alpha}\right](x\geqslant\gamma)$$

因此，解方程 $R=1-\exp\left[-\frac{(x-\gamma)^{\beta}}{\alpha}\right]$，并且在区间 [0, 1] 上 R 与 $1-R$ 同分布，得威布尔分布的抽样公式为

$$x=(-\alpha\ln R)^{1/\beta}+\gamma \tag{5.3-285}$$

逆变换法是一种较普遍的方法，只要分布函数的反函数存在，就可以采用逆变换法产生随机数。

(2) 近似法

设 r_1，r_2，…，r_n 相互独立，并且在区间 (0, 1) 均匀分布随机变量 $E(r_i)=1/2$，$D(r_i)=1/12$，故有

$$x_n=\frac{\sum_{i=1}^{n}r_i-n/2}{\sqrt{n/12}}$$

渐近服从标准正态分布，特别是 $n=12$ 时，有

$$x_{12}=\sum_{i=1}^{12}r_i-6 \tag{5.3-286}$$

式中，$r_i(i=1,2,\cdots,12)$是区间 (0, 1) 上均匀分布的随机数；x_{12}是标准正态分布随机数。近似法效率较差，但精度较高。当均匀随机数偏离一点均匀分布时，也能得到较好的正态分布随机数。

对于服从$N(\mu,\sigma^2)$的正态随机变量的抽样公式为

$$x_i=\mu+\sigma\left(\sum_{i=1}^{12}r_i-6\right) \tag{5.3-287}$$

如果对数正态分布的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma x}\exp\left[-\frac{(\ln x-\mu)^2}{2\sigma^2}\right]\quad x>0$$

并且$W=\ln x\sim N(\mu,\sigma^2)$，则对数正态分布随机数的抽样公式为

$$x=\exp(w) \tag{5.3-288}$$

式中，x 为对数正态分布随机数；w 为正态分布随机数。

(3) 基于正态分位数的近似采样

正态分布函数的逆函数的有理逼近，是产生正态分布随机数最好的一种方法。由变换法可知

$$\Phi(x)=R$$
$$x=\Phi^{-1}(R)$$

所以问题归结于求标准正态分布 $\Phi(x)$ 的反函数 $\Phi^{-1}(x)$。Hasting 有理逼近公式的计算简单，精度较高。

首先

$$x=\begin{cases}\sqrt{-2\ln r_i} & (0\leqslant r_i\leqslant0.5)\\ \sqrt{-2\ln(1-r_i)} & (0.5\leqslant r_i\leqslant1)\end{cases} \tag{5.3-289}$$

令

$$\xi_i=\begin{cases}x_i-\dfrac{a_0+a_1x_i+a_2x_i^2}{1+b_1x_i+b_2x_i^2+b_3x_i^3} & (0\leqslant r_i\leqslant0.5)\\ \dfrac{a_0+a_1x_i+a_2x_i^2}{1+b_1x_i+b_2x_i^2+b_3x_i^3}-x_i & (0.5\leqslant r_i\leqslant1)\end{cases} \tag{5.3-290}$$

则 ξ_i 是服从标准正态分布的随机数。式中，r_i 是均匀分布随机数，且

$a_0=2.515\ 17$ $b_1=1.432\ 788$

$a_1=0.802\ 853$ $b_2=0.189\ 269$

$a_2=0.010\ 328$ $b_3=0.001\ 308$

如果用 ξ_i 表示标准正态分布的随机数，则普通正态分布$X\sim N(\mu,\sigma^2)$的随机数 x_i 为

$$x_i=\mu+\xi_i \tag{5.3-291}$$

对于对数正态分布的随机数 y_i 有

$$y_i=\exp(\mu+\sigma\xi_i) \tag{5.3-292}$$

3.3.3 随机过程的模拟

(1) n 维正态向量的模拟

基本思想是将一个 n 维向量的模拟问题转换为n 个一元随机变量的模拟问题。

设 n 维向量$\xi=(\xi_1,\xi_2,\cdots,\xi_n)$具有均值$\mu=(\mu_1,\mu_2,\cdots,\mu_n)$，协方差矩阵$B_\xi=(b_{ij})_{m\times n}$，$B_\xi$ 为 n 阶正定对称矩阵。其中：

$$b_{ij}=E[(\xi_i-\mu_i)(\xi_j-\mu_j)] \tag{5.3-293}$$

取 n 个相互独立的一元标准正态随机变量 y_1，y_2，…，y_n，作三角变换

$$\begin{cases}\xi_1=\mu_1+C_{11}y_1\\ \xi_2=\mu_2+C_{21}y_1+C_{22}\\ \vdots\\ \xi_n=\mu_n+C_{n1}y_1+C_{n2}y_2+\cdots+C_{nn}y_n\end{cases} \tag{5.3-294}$$

若设$\xi=(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$

$$y=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$$
$$\gamma=(\mu_1,\mu_2,\cdots,\mu_n)^{\mathrm{T}}$$

$$C=\begin{bmatrix}C_{11}&0&0&\cdots&0\\ C_{21}&C_{22}&0&\cdots&0\\ \cdots&\cdots&\cdots&\cdots&\cdots\\ C_{n1}&C_{n2}&C_{n3}&\cdots&C_{nn}\end{bmatrix}$$

则由式 (5.3-294) 得

$$\xi=\mu+Cy \tag{5.3-295}$$

考虑到 B_ξ 为正定矩阵，选取下三角阵 C，使得

$$C\cdot C^T=B\xi \tag{5.3-296}$$

对于 $j=1$，2，…，n 有 (Cholesky 分解)

$$\begin{cases}C_{jj}=\left(b_{jj}-\sum_{k=1}^{j-1}C_{jk}^2\right)^{\frac{1}{2}}\\ C_{ij}=\left(b_{ij}-\sum_{k=1}^{j-1}C_{ik}C_{jk}\right)/C_{jj}(i=j+1,\cdots,n)\\ |B_\xi|=\left(\prod_{k=1}^{n}C_{kk}\right)^2\end{cases} \tag{5.3-297}$$

可以证明，ξ 就是具有概率密度函数为

$$f(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{n/2}\cdot|B_\xi|^{1/2}}\exp\left[-\frac{1}{2}(x-\mu)^{\mathrm{T}}B_\xi^{-1}(x-\mu)\right]$$

的 n 维正态向量。用随机向量方法模拟随机过程，当 n 取值很大时，这种方法困难很大。

(2) n 维对数正态随机向量的模拟

对数正态随机向量$\eta=(\eta_1,\eta_2,\cdots,\eta_n)$，其中 ξ_i 具有概率密度

$$f_i(x)=\frac{1}{\sqrt{2\pi}\sigma_i x}\exp\left[-\frac{(\ln x-\mu_i)^2}{2\sigma_i^2}\right] \tag{5.3-298}$$

且 η_i 和 η_j 的相关系数为

$$\rho_{ij}=\frac{E(\eta_i\eta_j)-E(\eta_i)E(\eta_j)}{\sqrt{D(\eta_j)}\sqrt{D(\eta_j)}} \tag{5.3-299}$$

则可借助于 n 维正态随机向量$\xi_i\sim N(\mu,\sigma^2)$，模拟 n 维对数正态随机向量

$$\eta_i=\exp\xi_i \tag{5.3-300}$$

其中，n 维正态随机向量$\xi=(\xi_1,\xi_2,\cdots,\xi_n)$的协方差矩

阵$B_\xi=(b_{ij})_{n\times n}$的元素为

$$b_{ij}=\ln\{\sqrt{[\exp(\sigma_i^2)-1][\exp(\sigma_i^2)-1]}\rho_{ij}+1\}\tag{5.3-301}$$

(3) 宽平稳随机过程的模拟

在许多实际问题中，常常要求模拟具有给定的数学期望、相关函数$R(\tau)$或功率谱密度$S(\overline{\omega})$的平稳随机过程$\eta(t)$，而不再考虑它们更细致的统计性质作为实际物理随机过程的模拟。无疑，这时随机过程$\eta(t)$并不惟一，存在许多不同的模拟方法。下面给出几个简单的例子。

如果宽平稳随机过程$\eta(t)$，均值为零，功率谱密度已知，令

$$a^2=\frac{1}{\pi}\int_{-\infty}^{+\infty}S(\omega)\mathrm{d}\omega\tag{5.3-302}$$

函数

$$f(\overline{\omega})=\frac{S(\omega)}{\pi a^2}$$

是一个非负规范化的实函数，可以作为随机变量$\overline{\omega}$的概率密度函数。而θ是在$[-\pi,\pi]$上均匀分布的随机变量，则

$$\eta(t)=a\cos(\omega t+\theta)\tag{5.3-303}$$

就是功率谱密度为$S(\overline{\omega})$的所求宽平稳随机过程。

如果模拟数学期望$E[\eta(t)]=0$，相关函数

$$R(\tau)=\sigma^2\exp(-\beta|\tau|)\tag{5.3-304}$$

的正态平稳随机过程$\eta(t)$，则产生$N(0,1)$分布的简单子样μ_1，μ_2，…：

$$\begin{cases}\eta(t_1)=\sigma\mu_1\\ \eta(t_i)=\exp(-\beta|t_i-t_{i-1}|)\cdot\eta(t_i-1)+\sigma\mu_i\\ \qquad\sqrt{1-\exp(2-\beta|t_i-t_{i-1}|)}\end{cases}\tag{5.3-305}$$

实际模拟时，为了减少模拟偏离，可以把u_1取为随机过程$\eta(t)$的数学期望。如果模拟数学期望$E[\eta(t)]=0$，相关系数

$$R(\tau)=\sigma^2\cos(\overline{\omega}\tau)\exp(-\beta|\tau|)\tag{5.3-306}$$

的正态平稳随机过程$\eta(t)$的模拟公式为

$$\eta(t)=\eta_1(t)\cos(\omega t)+\eta_2(t)\sin(\omega t)\tag{5.3-307}$$

式中，$\eta_1(t)$、$\eta_2(t)$是两个相互独立，以$\sigma^2\exp(-\beta|\tau|)$为相关函数的正态平稳过程，可由式（5.3-305）求得。

3.4 断裂失效概率预测

断裂力学是固体力学近代发展的一个新分支，它是研究带裂纹材料和构件的强度和断裂规律的一门科学。近三十年来，断裂力学在理论、实验和工程应用方面都取得了重大的进展，无论在设计、选材、估计构件的疲劳寿命以及制定验收标准方面都引入了断裂力学的理论和方法。可以说，现在对断裂问题的分析是离不开断裂力学这一工具的。

断裂力学的发展，也碰到不少有待解决的问题。例如对于缺陷和裂纹的定性定量，目前广泛使用的超声波探伤方法，还不能提供很准确的信息。而这对于断裂力学赖以计算的裂纹模型的选取和裂纹尺寸的确定都是很有影响的，有时是举足轻重的。又如关于工作应力问题，外加载荷引起的应力比较好估计一些，当然，这个由外载引起的应力也带有不同程度的随机变化成分在内，但它毕竟好估计；最难于处理的是残余应力的问题，残余应力的产生、分布、消除机制不清，理论和测试技术均处于萌芽时期，因此，由残余应力造成的影响，现在无更好的估计方法，这就给断裂力学计算带来了威胁，往往会使判断失误。另外，断裂韧度的测试也十分困难，这是因为很多的内外因素都在裂纹尖端这个关键区域起作用。

断裂力学在许多方面的偏差和不确定性叠加起来，降低了断裂力学的工程分析方法的可靠性。应用概率统计的方法，可以估计这些偏差和不确定因素，以弥补断裂力学应用的不足之处。

在断裂力学分析中引入概率理论，国外也称为概率断裂力学（Probability Fracture Mechanics，简记为FFM），是断裂力学的一个分支，它是用概率统计的理论和方法来分析研究和预测带裂纹的材料或结构的强度和断裂的一种工程方法。具体地说，它是研究在应力、强度、环境等作为随机参数时，裂纹结构发生断裂的概率大小的。

目前来说，确定发生断裂的概率的方法有两种。第一种方法是建立在过去的失效资料和经验的基础上的，它不顾结构的很多明晰的物理细节，而直接运用过去大量的资料，不需要使用分析模型或试验。这种方法的局限性是不能预言主要参数变化时对结构性能的影响。第二种是发展一种工程模型，是建立在对失效方式和控制参数的统计分布充分了解的基础上的。这种工程模型有明确的物理背景，是对断裂力学分析、疲劳寿命分析、应力分析和材料特性分析进行综合的结果。而作这种综合所运用的是概率统计方法，它可以给出可靠性的概念。事实上，这两种方法是有联系的，现在概率断裂力学的发展正是两者结合的结果。

3.4.1 概率断裂力学的统计基础

(1) 正态分布和对数正态分布下失效概率的计算方法

1) 正态分布下的失效概率的计算 假设裂纹尖端的应力强度因子$K_{\rm I}$和材料的断裂韧度$K_{\rm IC}$，均符合正态分布（下面的计算对裂纹尖端的COD和材料的$\delta_{\rm C}$，$J_{\rm I}$和材料的$J_{\rm Ic}$同样适用，只要验证了它们符合正态分布即可），则定义一个函数Y：

$$Y=K_{\rm IC}-K_{\rm I}\tag{5.3-308}$$

因为$K_{\rm IC}=(\mu_{K_{\rm IC}},\sigma_{K_{\rm IC}})$；$K_{\rm I}=(\mu_{K_{\rm I}},\sigma_{K_{\rm I}})$

所以$Y=(\mu_Y,\sigma_Y)=[(\mu_{K_{\rm IC}}-\mu_{K_{\rm IC}}),\sqrt{\sigma^2_{K_{\rm IC}}-\sigma^2_{K_{\rm I}}}]$

这时，我们可以直接应用联结方程：

$$u=-\frac{\mu_{K_{\rm Ic}}-\mu_{K_{\rm I}}}{\sqrt{\sigma^2_{K_{\rm IC}}+\sigma^2_{K_{\rm I}}}}\tag{5.3-309}$$

查u值表即可得失效概率和可靠度。

2) 对数正态分布下失效概率的计算 假设裂纹尖端的应力强度因子$K_{\rm I}$和材料的断裂韧性$K_{\rm IC}$均符合对数正态分布（对于COD和J积分方法下述计算同样适用，只要验证它们符合对数正态分布即可），则定义一个函数Z：

$$Z=\ln\frac{K_{\rm IC}}{K_{\rm I}}=\ln K_{\rm IC}-\ln K_{\rm I}$$

因为 $\ln K_{\rm IC}=(\mu_{K_{\rm IC}},\sigma_{K_{\rm IC}})$；$\ln K=(\mu_{K_{\rm I}},\sigma_{K_{\rm I}})$

则

$$\mu=\mu_{K_{\rm IC}}-\mu_{K_{\rm I}}$$

$$\sigma_Z=\sqrt{\sigma^2_{K_{\rm IC}}+\sigma^2_{K_{\rm I}}}$$

应用联结方程：

$$u=-\frac{\mu_{\rm KIc}-\mu_{\rm KI}}{\sqrt{\sigma^2_{K_{\rm IC}}+\sigma^2_{K_{\rm I}}}}\tag{5.3-310}$$

查u值表即可得失效概率和可靠度。

(2) $K_{\rm I}$和$K_{\rm IC}$为任意分布时失效概率的计算方法

我们仍用K判据来讨论在任意分布时构件的失效概率计算问题。设$K_{\rm I}$的概率密度函数为$\varphi(K_{\rm I})$，断裂韧度$K_{\rm IC}$的概率密度函数为$f(K_{\rm IC})$。利用$K_{\rm I}=K_{\rm IC}$判据，可知当$K_{\rm IC}>K_{\rm I}$时构件将是安全的，因此$K_{\rm IC}>K_{\rm I}$的概率就是可靠度，即$P(K_{\rm IC}>K_{\rm I})=R$；而$K_{\rm IC}<K_{\rm I}$，构件将产生失稳扩展，所以$K_{\rm IC}<K_{\rm I}$的概率就是失效概率$P_{\rm f}$，记为$P(K_{\rm IC}<K_{\rm I})=P_{\rm f}$。

可靠度的计算公式如下：

$$R=\int_{-\infty}^{+\infty}\varphi(K_{\rm I})\left[\int_{K_{\rm I}}^{\infty}f(K_{\rm IC})\right]{\rm d}K_{\rm I}\qquad(5.3\text{-}311)$$

或

$$R=\int_{-\infty}^{+\infty}f(K_{\rm IC})\left[\int_{-\infty}^{K_{\rm IC}}\varphi(K_{\rm IC})\right]{\rm d}K_{\rm IC}\qquad(5.3\text{-}312)$$

当给出 $K_{\rm I}$ 和 $K_{\rm IC}$的不同分布类型后，可以得出相应的可靠度计算式。

(3) 断裂力学判据中概率方法的引入

假定我们研究的构件的断裂适合使用判据 $K_{\rm I}=K_{\rm IC}$。

设 $K_{\rm I}=\alpha\sigma\sqrt{\pi a}=K_{\rm IC}$，则可写为

$$f(K_{\rm IC},\alpha,\sigma,a)=0\qquad(5.3\text{-}313)$$

代入不同参量进行求解：

$$K_{\rm IC}=\varphi_1(\sigma,\alpha,a)\qquad(5.3\text{-}314)$$

$$\sigma=\varphi_1(K_{\rm IC},\alpha,a)\qquad(5.3\text{-}315)$$

$$a=\varphi_3(K_{\rm IC},\alpha,a)\qquad(5.3\text{-}316)$$

上述三个式子就代表了断裂力学工程应用的三个方面。式（5.3-314）表示在一个结构或构件中如已知工作应力 σ 和裂纹尺寸（假设形状系数 α 已知），则可确定为了安全工作所需要的材料断裂韧度值，这就是断裂力学设计对材料提出的要求，也可以说式（5.3-314）是按断裂力学观点进行选材或制定热处理规范的依据。

式（5.3-315）表示在已知构件内的缺陷（有时在设计阶段均可假设在危险部位有一个一定尺寸的缺陷，这被称为设计缺陷）和材料的断裂韧性 $K_{\rm IC}$的情况下，要计算构件能承受的最大工作应力，即作强度校核。它是结构或构件进行断裂力学设计的依据。

式（5.3-316）则是在构件工作应力一定及材料断裂韧性 $K_{\rm IC}$已知的情况下，来确定构件能承受的最大裂纹尺寸。因此，在配合寿命要求后，可以估计它在验收时所允许的初始裂纹尺寸，也限制了验收标准（或产品的质量标准）。所以它是讨论裂纹容限的依据。

除了上述三方面外，再附加一个$\frac{{\rm d}a}{{\rm d}N}=C\Delta K^n$，即裂纹扩展速率公式。从它积分可得到构件寿命的估计式，它是从断裂力学角度对带裂纹的作寿命计算的依据。由此得到的信息，可用来安排检修期或调换构件。

上述四方面，就是全部断裂力学的工程应用。下面以断裂力学设计为例介绍如何引入概率统计分析。

以式（5.3-315）为例，考虑无限板中心贯穿裂纹受拉伸的简单情况。

$$\sigma_{\rm c}=\frac{K_{\rm IC}}{\sqrt{\pi a}}\qquad(5.3\text{-}317)$$

式中，$\sigma_{\rm c}$ 为临界工作应力。

两边取对数：

$$\lg\sigma_{\rm c}=\lg K_{\rm IC}-\frac{1}{2}\lg\pi-\frac{1}{2}\lg a\qquad(5.3\text{-}318)$$

现假定材料的断裂韧性 $K_{\rm IC}$，裂纹尺寸 a 均符合对数正态分布。

$$\lg K_{\rm IC}=N(\mu_{K_{\rm IC}},\sigma_{K_{\rm IC}})$$

$$\lg a=N(\mu_{\rm a},\sigma_{\rm a})$$

则根据式（5.3-318）可知，$\lg\sigma_{\rm c}$ 也符合对数正态分布。

$$\lg\sigma_{\rm c}=N(\mu_{\sigma_{\rm c}},\sigma_{\sigma_{\rm c}})$$

按式（5.3-318）的关系，有（设 $\lg K_{\rm IC}$、$\lg a$ 均为独立随机变量）：

$$\begin{cases}\mu_{\sigma_{\rm c}}=\mu_{K_{\rm IC}}-\frac{1}{2}\lg\pi-\frac{1}{2}\mu_{\rm a}\\ \sigma_{\sigma_{\rm c}}=\sqrt{\sigma_{K_{\rm IC}}^2+\frac{1}{4}\sigma_{\rm a}^2}\end{cases}$$

代入 $\lg K_{\rm IC}$和 $\lg a$ 的具体分布参数即可对 $\sigma_{\rm c}$ 和失效概率进行计算。

(4) 概率断裂力学的一般统计基础

概率断裂力学（PFM）是一种预测并提高结构可靠性的工具，它用概率统计方法来估计裂纹体的剩余强度和寿命，在一定假设的基础上，对结构的可靠性进行评定；现在，先介绍一下 PFM 的统计基础。

假定当一种与失效方式相联系的应力 s 超过该失效方式所控制的强度 r 时，我们认为失效事件 E 将发生。

$$P(E)=P(Y<0)\qquad(5.3\text{-}319)$$

式中，Y 是强度差。

$$Y=r-s=G(X_i)\qquad(5.3\text{-}320)$$

式中，X_i 为独立的参量，$i=1,2,3,\cdots,n$。当然，我们把 X_i 作为随机变量来处理，因此，Y 也是随机变量，当 $Y<0$ 时表示结构将失效。

函数 G 是由以前的经验，在作了一定的假设之后提炼出模型而得到的，如果 G 的形式不能确定，则可根据已有数据用最小二乘法拟合得到。除此之外，在 G 函数中也可引进不确定性的估计，这种不确定性是由工程判断得来的。

每一个独立随机变量 X_i 的累积概率分布函数是

$$P(X_i<x_i)=F(x_i)\qquad(5.3\text{-}321)$$

在实际应用中，需要首先找出基本的随机变量 X_i 的分布$F(x_i)$，然后代入式（5.3-320）便可得到 Y，最后把 Y 代入式（5.3-319）就可得到失效概率分布。在简单的情况下，可以得到解析解，但大多数情况下，只能得到数值解。

3.4.2　带裂纹构件强度的概率断裂力学预测

概率断裂力学的出发点之一是，考虑应力和强度的统计变化，调整各个参数，以使某种失效方式的失效概率达到一般的极小值，从而进行结构设计和材料选择。

假设：

工作应力为 σ，它的分布符合正态分布$N(\mu_\sigma,\sigma_\sigma)$；

屈服强度为 $\sigma_{\rm s}$，它的分布符合正态分布$N(\mu_{\sigma_{\rm s}},\sigma_{\sigma_{\rm s}})$；

屈服强度为 $\sigma_{\rm f}$，它的分布符合正态分布$N(\mu_{\sigma_{\rm f}},\sigma_{\sigma_{\rm f}})$。

定义：

$Y_{\rm s}=\sigma_{\rm s}-\sigma$ 为屈服强度差；

$Y_{\rm f}=\sigma_{\rm f}-\sigma$ 为脆断强度差；

则当：

$Y_{\rm s}<0$ 时，发生屈服失效事件；

$Y_{\rm f}<0$ 时，发生脆断失效事件；

安全系数定义为：

$n_{\rm s}=\frac{\mu_{\sigma_{\rm s}}}{\mu_\sigma}$ 为屈服安全系数；

$n_{\rm f}=\frac{\mu_{\sigma_{\rm f}}}{\mu_\sigma}$ 为脆断安全系数；

强度和应力的分散性对于构件的可靠度有直接的影响，在安全系数不变的情况下，缩小强度和应力的分散性，会较大地提高构件的可靠性。

分布的参数可用变差系数 c 表示：

$c_\sigma=\sigma_\sigma/\mu_\sigma$ 为工作应力变差系数；

$c_{\rm s}=\sigma_{\sigma_{\rm s}}/\mu_{\sigma_{\rm s}}$ 为屈服强度的变差系数；

$c_{\rm f}=\sigma_{\sigma_{\rm f}}/\mu_{\sigma_{\rm f}}$ 为脆断强度的变差系数；

如果 $c_{\rm s}=c_{\rm f}$，且 $\mu_{\sigma_{\rm s}}=\mu_{\sigma_{\rm f}}$，此时有 $n_{\rm s}=n_{\rm f}$ 和 $P_{\rm s}=P_{\rm f}$。但实际上 $c_{\rm s}\neq c_{\rm f}$，这是因为 $\sigma_{\rm f}$ 是由下式给出的：

$$\sigma_f = \frac{K_{IC}}{\sqrt{\pi a}} \qquad (5.3\text{-}322)$$

在上式中，由于断裂韧度 K_{IC} 的分散性比屈服点 σ_s 的分散性大，此外，裂纹尺寸 a 的分散性和 K_{IC} 的分散性都提高了 σ_f 的分散性，所以 c_f 比 c_s 大。从物理意义上讲，脆断是一种不稳定的过程，而屈服断裂是一种稳定过程。

因此，在设计中，应使脆断的可能性尽量减小，即提高防止脆断的安全性。在实际中通常使用诺谟图（Nomogram）选择安全系数从而调整两种失效方式的破坏概率，如图5.3-50所示。

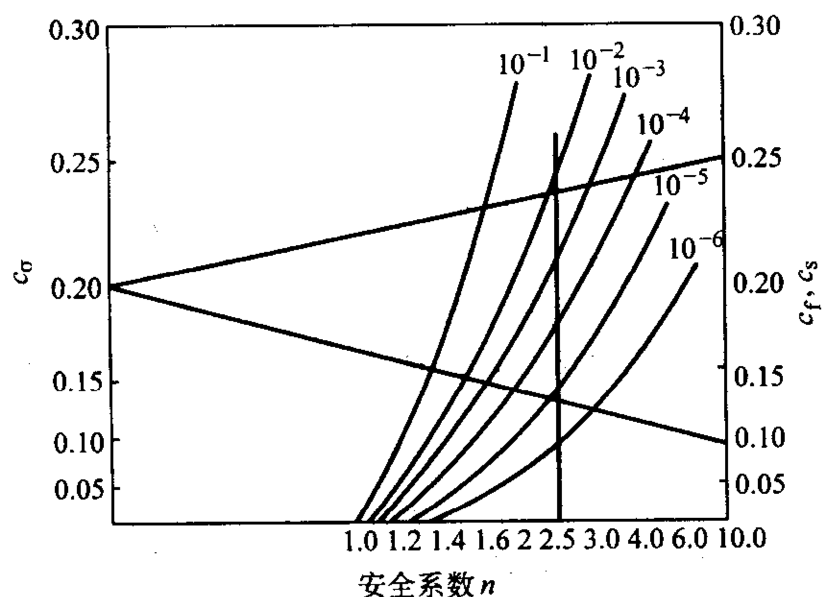


图5.3-50 安全系数诺谟图

图5.3-50左边纵坐标代表 c_σ，右边纵坐标代表 c_s 和 c_f，横坐标为安全系数。为了求得失效 $P(Y<0)$，可在 c_σ 和 c_s 以及 c_σ 和 c_f 之间引两条直线，它们分别与 $P(Y<0)$ 相交，交点对应于不同的安全系数。例如，$c_\sigma=0.20$，$c_s=0.10$，且 $c_f=0.25$，在给定的某一设计中，其安全系数为2.5（$n_s=n_f=2.5$），于是可从图5.3-50查得，对于两种失效方式的失效概率为

对于屈服破坏 $P_s=2\times10^{-6}$；

对于脆断破坏 $P_f=1.3\times10^{-2}$。

结果表明，该设计不能被接受，因为脆断破坏的概率远大于屈服破坏概率。同时，这个例子也说明，常规设计的安全系数是以 σ_s 或 σ_b 与工作应力 σ 相比来确定的，这就存在一种潜在的危险，由于脆断强度 σ_f（与 σ_b 不同，它是考虑裂纹以后的强度）的 c_f 很大，（从图5.3-50可见，由于 σ_{σ_f} 大，μ_{σ_f} 小，所以 c_f 大）所以在诺谟图中它的破坏概率很大，这一点只有在概率断裂力学的计算中才能揭示出来，并用来指导设计，一直调整到 $P_f<P_s$ 为止。这就是从可靠性意义来说的最佳设计。

3.5 疲劳失效概率预测

3.5.1 载荷的统计

载荷（此处主要指动载荷）的内容很广泛，这里主要讨论引起构件疲劳失效的一类动载荷，亦常被称做疲劳载荷或交变载荷。

（1）疲劳载荷的形式

疲劳载荷的形式很多，图5.3-51给出了典型的疲劳载荷的形式。疲劳载荷一般可分为确定的和随机的两种。确定的疲劳载荷是一种按一定规律变化的，可以重复的载荷，可用数学公式表达。随机疲劳载荷是一种无规则的不能重复的载荷，只能进行统计描述。随机疲劳载荷又可进一步分为平稳的和非平稳的。

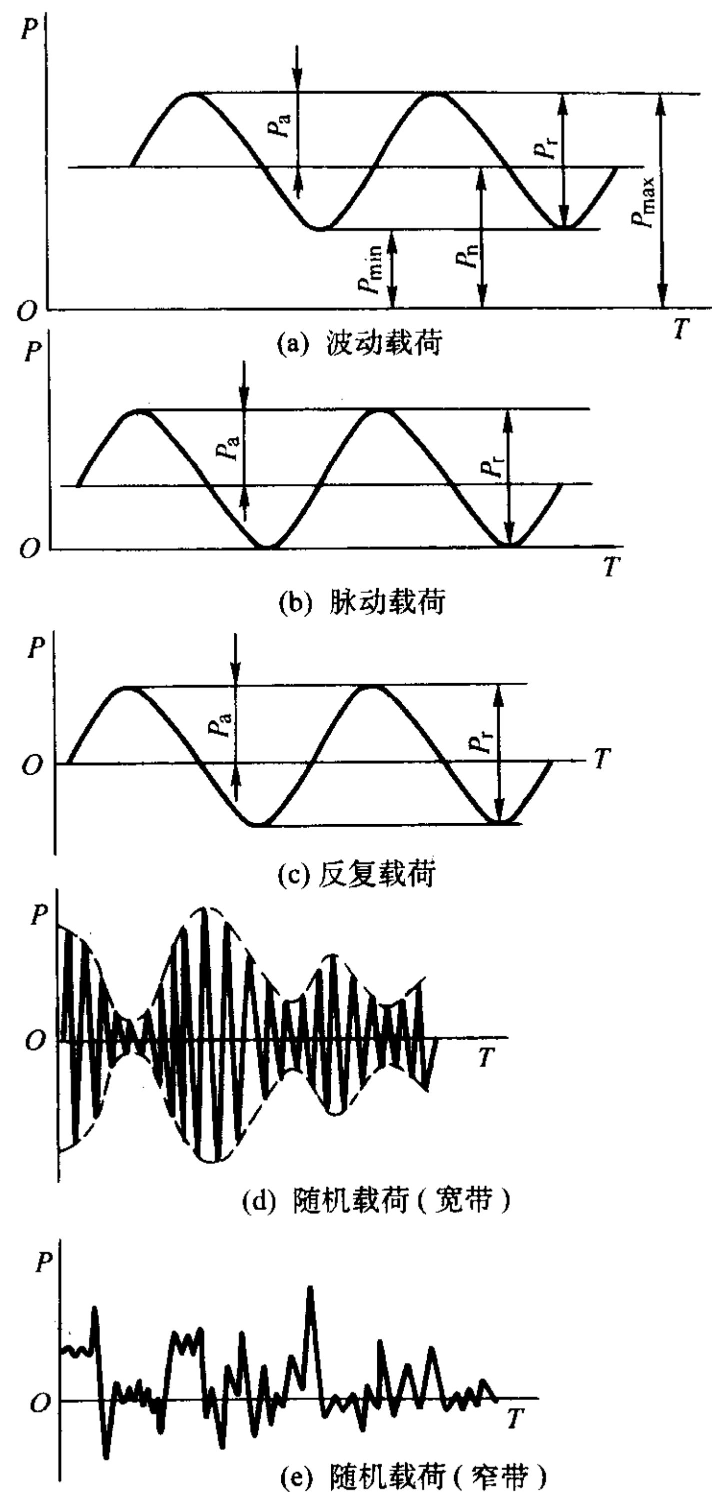
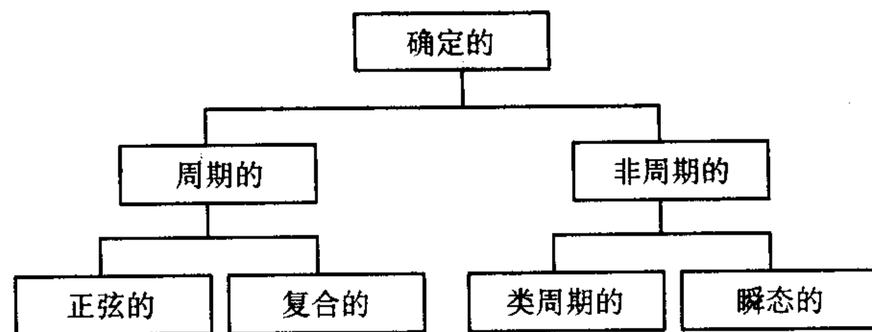


图5.3-51 典型的疲劳载荷

确定的疲劳载荷又可以进一步分类如下：



在日常的工程实践中，完全反复正弦波具有零平均值的加载情况，这时载荷可用下式表示。

$$f(t)=p\sin(wt)$$

式中，p 为最大的载荷幅值。

对于这种疲劳载荷，在可靠性研究中较多地采用下述两种模型，因为使用这两种模型可以得到问题的封闭解。这两种模型如下。

1）完全反复载荷的瞬时值服从正态分布，而峰值载荷 P_T 服从瑞里（Rayleigh）分布，其概率密度函数为

$$f(P_T)=\frac{m}{h^2}e^{-P_T^2/h^2}$$

式中，$h^2=\dfrac{P_{T_1}^2+P_{T_2}^2+\cdots+P_{T_n}^2}{2n}$；$m=(h^2\pi)^{1/2}\sqrt{2}$；$n=$ 峰值载荷的个数。

2）峰值载荷 P_T 可以认为是正态或近似正态分布，其概率密度函数为

$$f(P_T)=\frac{1}{\sqrt{2\pi}\sigma_{P_T}}\exp\left\{-\frac{1}{2}\left(\frac{P_T-\mu_{P_T}}{\sigma_{P_T}}\right)\right\}$$

式中，$\mu_{P_T}=\frac{1}{n}\sum_{i=1}^{n}P_{T_i}$；$\sigma_{P_T}=\left[\frac{1}{n}\sum(P_{T_i}-\mu_{P_T})^2\right]^{1/2}$

（2）随机载荷的统计分析

对随机载荷只能进行统计描述。有两种统计分析方法用来描述随机载荷，即功率谱法和计数法。

1）功率谱法　任何随机过程的基本特征主要用四种统计函数来描述，分别是：均方值 σ^2，概率密度函数 $P_{(x)}$——幅值域，自相关函数 $R_{(\tau)}$——时域，功率谱密度函数 $C_x(f)$——频域。

如果过程是一个平稳高斯过程，则可用功率谱密度函数来完全确定该过程。功率谱法给出载荷幅值的均方值随概率的分布，即载荷的功率谱密度函数。功率谱法是一种较精确、严密的统计方法，但只有在载荷过程是一种较精确、严密的统计方法，但是只有在载荷过程是一种平稳高斯随机过程的条件下，才有可能用功率谱密度函数把该过程完全确定下来。因此，在采用功率谱法的时候，应对测定的载荷工况有一定的条件限制。

2）计数法　计数法是一种建立在概率论和数理统计方法基础上的统计分析方法，可以分为三种类型。

① 简单的单参数计数法。例如以随机载荷的峰值、振幅或穿过某一载荷级的频次为计数对象进行统计分析，分别称为峰值计数法、振幅计数法和穿级计数法。

这种统计分析方法比较简单、直观，但不够严密、精确，把本来连续的载荷过程离散化了，既得不到载荷频率变化的资料，也没有载荷发生次序的资料，只是对随机载荷过程的一种近似的描述。

② 以材料的应力－应变相应为基础的计数法，如振幅对计数法和雨流计数法。它们是用载荷（或应变）循环，而不是用振幅（即半个循环）来分析载荷历史。振幅对计数法是对载荷历史中量值相等、符号相反的振幅成对的计数。雨流法也是对载荷历史中量值相等、符号相反的振幅对计数，但这些振幅对在应力－应变滞后回线中必须是从一顶点出发又回到同一顶点形成一个闭环时的一对振幅。这种计数方法如同雨水顺着屋顶流下的情况，故名雨流法。振幅对法和雨流法的计数结果常常是大致相同的，但后者比较适用于数据处理的自动化。这类计数法把载荷统计分析的过程和材料的循环应力－应变概念来预估构件疲劳寿命的方法。这种方法已经使构件疲劳寿命的预估有了重大的改进，并可采用计算机分析的技术，为估算构件疲劳寿命提供了有效的分析方法。

③ 双参数计数法。这种计数法是为了克服峰值法和振幅法各自的缺点而提出的一种计数法，它以载荷历史中的两个参数为对象，同时进行统计计数，能保留载荷历史中的幅值－均值关系。

（3）载荷谱

1）载荷谱的定义及测定　计数法给出了载荷的概率密度函数或概率分布函数。由此，我们可以确定载荷的大小与其出现频次（或频率）的关系。我们把表示载荷的大小与其出现频次（或频率）关系的图形、表格、数字、矩阵等称为载荷谱。单参数计数法的结果可用图形来表示，如图5.3-52所示。

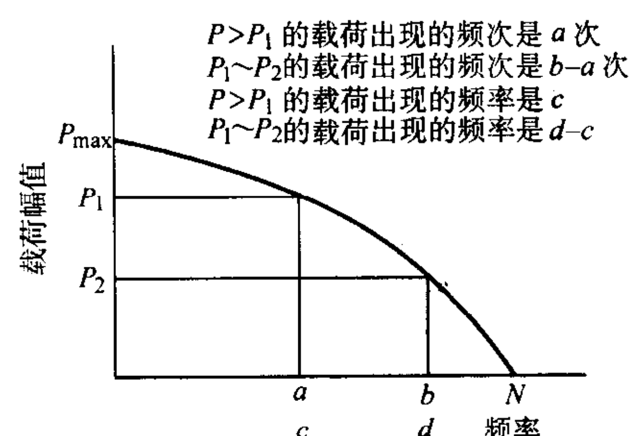


图 5.3-52　载荷谱

2）载荷谱的典型分布形式　根据实际测定可以认为，常见的载荷谱主要有两种典型的分布形式。一种是正态分布，一种是直线（或指数）分布。如图5.3-53所示。第二类载荷谱如图5.3-54所示，其特点是载荷的总作用次数是有限的，其实际工作应力高于疲劳极限。对于这一类构件，主要考虑其静强度外，还要考虑其疲劳强度。第三类载荷谱是最一般的情况，即大载荷高于疲劳极限，但出现次数少，而大量出现的载荷是低于疲劳极限的小载荷。如图5.3-55所示。

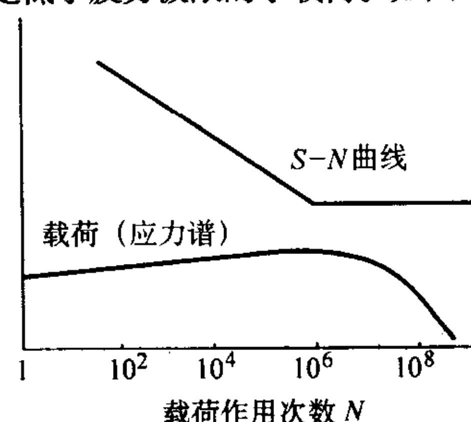


图 5.3-53　第一类载荷谱

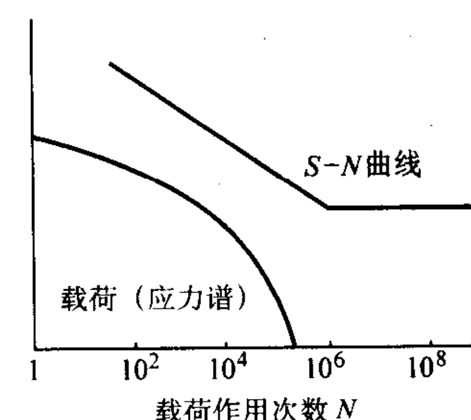


图 5.3-54　第二类载荷谱

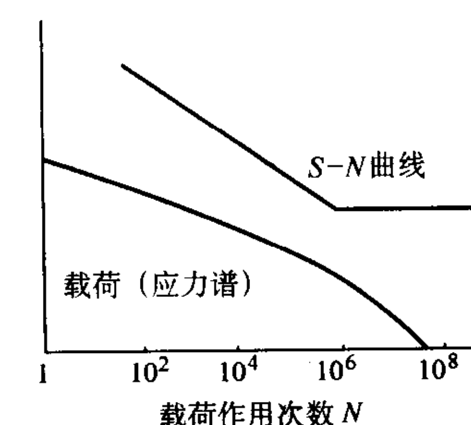


图 5.3-55　第三类载荷谱

3）载荷谱的类型　载荷谱主要有三种类型。第一类载荷谱如图5.3-53所示，其特点是最大载荷的出现频次极多，因此必须把该载荷作为持久极限来考虑；第二类载荷谱如图5.3-54所示，特点是载荷的总作用次数是有限的，其实际工作应力高于疲劳极限。这类载荷，除了必须考虑构件的静强度，还需要考虑其疲劳强度；第三类载荷谱是最一般的情况，即大载荷高于疲劳极限，但出现的次数少，而大量出现的载荷是低于疲劳极限的小载荷。

3.5.2　给定寿命下的疲劳强度分布

在疲劳强度的可靠性设计中，需要知道在某一给定寿命下材料的疲劳强度的分布，这可以通过两种途径来得到。

一是从疲劳的 $P-S-N$ 图中得到。应该指出，到目前为止，疲劳的 $P-S-N$ 曲线均为在恒定应力 S 下，通过试验求得 $P-N$ 曲线，即得到恒应力下的寿命分布，再把 $P-N$ 曲线和 $S-N$ 曲线相结合，从而得到 $P-S-N$ 曲线；

二是直接从实验中，在固定 N 的情况下求 $P-S$ 曲线，即得到固定寿命下的应力分布，这个应力分布，就代表了给定寿命下材料的疲劳强度分布。

(1) $P-S-N$ 曲线

对称应力循环条件下平均应力为零，选择4～5个应力水平，每个应力水平至少15个试件进行成组法试验。在较长寿命区（应力接近疲劳强度）采用升降法得给定寿命的疲劳强度的近似分布。如图5.3-56所示，得到 $S-N$ 曲线。

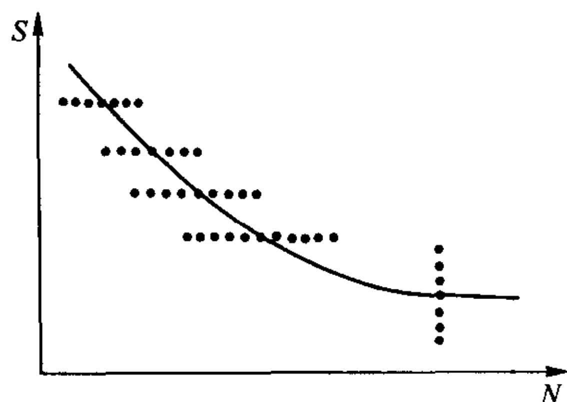


图5.3-56 $S-N$ 曲线实验测定

在恒定应力水平下，疲劳寿命分布一般服从对数正态分布，并且由恒定应力下的寿命样本，可以计算寿命分布的均值和方差。

各应力水平下，把等失效概率的点连接起来，就可以得到许多失效概率曲线，称为 $P-S-N$ 曲线。如果失效概率为 P，则不失效保证构件正常工作的可靠度为 $R=1-P$。例如5%失效概率曲线也可以反过来称为95%可靠度曲线。而 $S-N$ 曲线通常是指平均曲线，即失效概率 $P=50\%$ 的曲线，如图5.3-57所示。

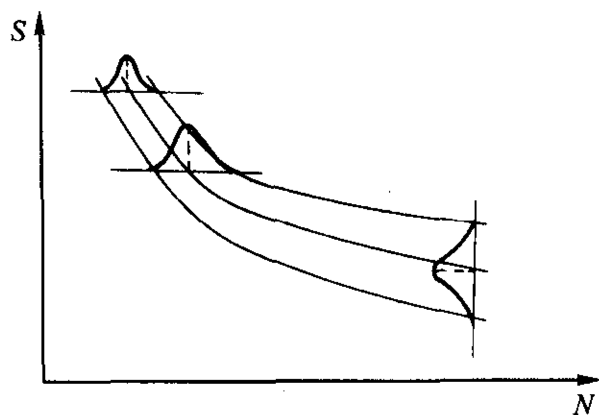


图5.3-57 $P-S-N$ 曲线示意图

经受无限循环而不失效的应力，称为疲劳极限。一般非铁合金的长寿命区没有平台，所以用规定寿命的应力（例如对应 10^7 或 10^8 周次应力）称为条件疲劳极限和疲劳强度。

(2) 影响 $P-S-N$ 曲线的因素

材料成分、晶粒尺寸大小和晶粒方向、热处理状态、焊接等都会影响疲劳强度，在上述条件基本相同时，构件与光滑小试件主要差异在于构件外形的差异、尺寸大小差异和表面质量的差异。综合考虑上述三种因素，构件在对称循环下的疲劳强度

$$\sigma_{-1}^{0}=\frac{\varepsilon_\sigma\beta}{K_\sigma}\sigma_{-1}$$

式中，σ_{-1} 为光滑小试件的疲劳极限；ε_σ 为对称循环下的尺寸系数。由于构件尺寸大，形成疲劳裂纹机会也更多，所以 $\varepsilon_\sigma<1$；β 为对称循环下的表面质量系数。当构件表面质量低于小试件时，$\beta<1$；而表面经强化后，$\beta>1$；K_σ 为对称循环下的有效应力集中系数。一般 $K_\sigma>1$。

实验证明：构件应力集中、尺寸大小和表面质量等因素，只对属于动应力的应力幅度 σ_a（或用 S_a 表示）有影响，而对属于静应力的平均应力 σ_m（或用 S_m 表示）并没有影响。

3.5.3 材料疲劳强度的统计数据

金属材料疲劳极限的分布数据，是要通过很多疲劳试验数据得来的。由于疲劳试验所得数据的分散性很大，要得到精确度很高的 $P-S-N$ 曲线，必须取足够多的试样，是一项投入很大的工作。表6.5-2～表6.5-5是根据有关文献提供的疲劳数据，用上述方法将这些数据进行处理后，换算得到在给定寿命 N 下的疲劳极限分布的参数值。由于原始实验数据的数量有限，分散性大，需要继续做补充实验使数据精确化。

3.5.4 疲劳极限线图

在常规疲劳强度设计中，所用的疲劳极限线图是由各种不对称系数 r 下的均值画出的，是一条直线。而在可靠性设计中疲劳极限线图是一条曲线分布带，见图5.3-58。设不对称系数 r 的直线与疲劳极限的均值图线相交于 A 点，向量 OB 表示平均应力 σ_m，其标准差为 S_m；向量 BA 表示应力幅 σ_a，其标准差为 S_a；向量 OA 是合成向量 σ_r，其标准差为 S_r。

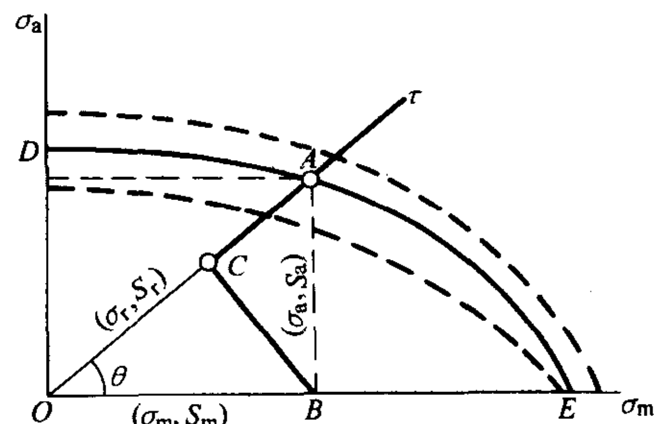


图5.3-58 疲劳极限线图

由图5.3-58可以看出，均值 σ_r 为

$$\sigma_r=\sqrt{\sigma_a^2+\sigma_m^2} \tag{5.3-323}$$

求标准差 S_r 时，可作辅助线 $BC\perp OA$，则 OC 为 σ_m 在 σ_r 上的投影，CA 为 σ_a 在 σ_r 上的投影，有

$$OA=OC+CA$$

由计算可得，两正态分布函数之和的标准差 S_r 为

$$S_r=\left[\frac{\sigma_a^2S_a^2+\sigma_a^2S_m^2}{\sigma_a^2+\sigma_m^2}\right] \tag{5.3-324}$$

因为应力幅 σ_a 和平均应力 σ_m 为

$$\sigma_a=(\sigma_{max}-\sigma_{min})/2,\sigma_m=(\sigma_{max}+\sigma_{min})/2$$

不对称系数 $r=\sigma_{min}/\sigma_{max}$，代入有

$$\sigma_a=(1-r)\sigma_{max}/2,\sigma_m=(1+r)\sigma_{max}/2$$

当 r 为常数时，应力幅 σ_a 的标准差 S_a 为

$$S_a=\left(\frac{1-r}{2}\right)S_\sigma \tag{5.3-325}$$

平均应力 σ_m 的标准差 S_m 为

$$S_m=\left(\frac{1-r}{2}\right)S_\sigma \tag{5.3-326}$$

这里 S_a 为最大应力 σ_{max} 的标准差。在疲劳试验中，σ_{max} 就是疲劳极限，其标准差在表6.5-3～表6.5-5中给出。当 S_σ 已知，则 S_a 和 S_m 可以计算得到，应力幅 σ_a 和 σ_m 由式(5.3-324)得到，这样，由式(5.3-324)可以求得图5.3-58中在给定不对称系数 r 的标准差 S_r，代入式 $x_{0.999}=\bar{x}-3S$ 中的 S，即可求得相应于可靠度 $R=0.999$ 的疲劳极限线图。

3.5.5 疲劳剩余强度的确定方法

评价一个结构的疲劳可靠性取决于对它的设计要求，如果结构需要安全地运行一段时间而不进行检查，那么在整个服役期间必须保持一个低的可接受的失效概率。如民用工程结构就属于这一类。然而，由于质量和经济的要求，如某些航天结构，在它们在被检查修理或替换以前，允许有一个疲劳损伤的累积过程，这时就要求结构的剩余强度保持高于某一极限值。

在第一种情况下的可靠性分析，仅仅考虑疲劳过程的最终结局——断裂，在给定循环次数 N 下的失效概率为

$$P_f(N) = P(N_f \leqslant N) \tag{5.3-327}$$

式中，N_f 为失效循环数，是一个随机变量。

在第二种情况下的可靠性分析，是把结构的强度 R（也称为剩余强度）作为循环数 N 的函数，并假定当作用应力等于剩余强度时就发生断裂。于是失效概率就有

$$P_f(N) = P[R(N) \leqslant S] \tag{5.3-328}$$

其中 S 为作用的交变应力最大值。

第二种情况是更普遍、更有意义的一种分析方法，它被用来进行疲劳过程理论和机理的研究，同时也应用于工程分析中。

许多研究者对第二种情况进行了分析，但他们都基于同一个假定：在形成一个宏观的初始疲劳裂纹之前，材料（或构件）的强度保持不变。但在随后的研究中，经过试验资料的统计分析发现，即使在裂纹的形成阶段，强度也随时在降低。其强度降低的原因还不是十分清楚，但初步认为这是在应变不相容的中心（如空穴、晶界等）形成显微裂纹所造成的。忽略了这一部分的强度的降低，如果裂纹形成阶段占整个疲劳寿命的很大一部分，会造成错误的可靠性估计。因此有必要对疲劳裂纹的形成和扩展阶段的强度变化提出一个可靠性分析的程序。研究者分析指出：①对于初始强度 R_0 服从威布尔分布的材料（或构件），其剩余强度也符合威布尔分布；②进一步分析表明，分布的形状参数在疲劳的两个阶段是不变的，但位置参数和尺度参数随着循环次数在降低；③两个阶段比较指出，初始强度 R_0 是一元分布，而剩余强度则是二元分布，高剩余强度部分对应于疲劳裂纹的形成阶段，低剩余强度部分相应于疲劳裂纹的扩展阶段。

（1）疲劳过程的剩余强度的考虑

为了估算式（5.3-328）给出的疲劳失效概率，我们将分别考虑疲劳裂纹形成阶段和扩展阶段的剩余强度的概率分布。

1）裂纹的形成阶段　疲劳断口的显微研究揭示了这个阶段的一些现象和特征，研究者提出了一些现象和特征，得到了许多裂纹的形成机制。但这些机制的定量描述仍在发展中。

引入一个损伤参数 D，在循环开始时其值为 0，随后单调增长，当一个初始的宏观裂纹形成时，它的值为 1。故有

$$D = \frac{R_0 - R}{R_0 - R_c} \tag{5.3-329}$$

式中，R_0 为初始拉伸强度；R_c 为刚好形成宏观裂纹时的强度；R 为某一瞬时的强度。R_c 标志着裂纹形成阶段的结束，它与宏观裂纹尺寸 a_c 有关。可以用下述 Griffith－Irwin 关系来计算。

$$R_c = \alpha K_C a_c^{-\frac{1}{2}} \tag{5.3-330}$$

式中，K_C 为材料的断裂韧性；α 为形状因子。

进一步假定，疲劳损伤率是由下式来表示。

$$\frac{dD}{dN} = n\left(\frac{S}{1-D}\right)^m \tag{5.3-331}$$

式中，n、m 为材料的参数。

对式（5.3-331）进行积分，并利用式（5.3-329），可得

$$R = R_c + (R_0 - R_c)(1 - N/N_c)^{m'} \tag{5.3-332}$$

式中，$m' = \frac{1}{1+m}$；$N_c = \frac{1}{n\ (1+m)\ S^m}$。

从式（5.3-332）可知，由于 R_c，R_0，N_c 和 m 是随机变数，因此 R 也是随机变数。于是，可由这些随机变数的分布来估计 R 的分布（或其参数）。假定 R_0 和 N_c 均符合三参数的威布尔分布。对于给定的式样，试验结果表明，R_c 仅仅有较小的离散度。而对于 m（和 m'）规定一个分布还缺乏物理基础。

因此，在式（5.3-332）中，如果仅仅假定 R 和 R_0 还有 N_c 是随机变量的话，那么假如 R_0 符合威布尔分布，则 R 也必然符合威布尔分布，有

$$P(R_0 \leqslant r_0) = 1 - \exp\left[-\frac{r_0 - a}{b}\right]^c \tag{5.3-333}$$

和

$$P(R \leqslant r) = 1 - \exp\left[-\frac{r - a'}{b'}\right]^c \tag{5.3-334}$$

式中，$a' = R_c + (a - R_c)(1 - N/N_c)^{m'}$；$b' = b(1 - N/N_c)^{m'}$。

2）裂纹的扩展阶段　一旦宏观裂纹形成，它将在交变应力作用下直到失效。在任意给定时刻，裂纹扩展阶段的剩余强度与裂纹长度 a 可同样用 Griffith－Irwin 关系：

$$R = \alpha K_c a^{-\frac{1}{2}} \quad a \geqslant a_c \tag{5.3-335}$$

裂纹扩展速率可用 Paris－Erdogen 关系：

$$\frac{da}{dN} = C(\Delta K)^n$$

式中，ΔK 为应力强度因子幅度；c、n 为实验常数。

上式可改为

$$\frac{da}{dN} = \beta a^{n/2} \tag{5.3-336}$$

式中，β 为应力幅度的函数。

对式（5.3-335）两边对 N 求导，并将式（5.3-336）代入：

$$\frac{dR}{dN} = \alpha K_c\left(-\frac{1}{2}\right)a^{-3/2}\frac{da}{dN}$$

$$\frac{dR}{dN} = \alpha K_c\left(-\frac{1}{2}\right)a^{-3/2}\beta a^{n/2} = -\gamma\left(\frac{1}{R}\right)^{n-3} \tag{5.3-337}$$

其中 $\gamma = \frac{1}{2}\beta(K_c\alpha)^{n-2}$。

一般材料 $n \geqslant 2$，特别对低断裂韧性的材料，其 $n > 2$。所以我们分 $n = 2$ 和 $n > 2$ 来讨论剩余强度的分布。

① $n = 2$　这时式（5.3-337）成为

$$\frac{dR}{dN} = -\gamma R = -\frac{\beta}{2}R$$

积分上式得

$$R = R_c \exp\left[-\frac{\beta}{2}(N - N_c)\right] \tag{5.3-338}$$

类似可得到

$$R = S\exp\left[-\frac{\beta}{2}(N_f - N)\right] \tag{5.3-339}$$

假设 β 为常数，N_f 服从威布尔分布：

$$P\ (N_f \leqslant N)\ = 1 - \exp\left[-\frac{N - a_n}{b_n}\right]^{C_n} \tag{5.3-340}$$

改写式（5.3-337）得

$$N_f = \frac{2}{\beta}\ln R - \ln S + N$$

则可以得到

$$P(\ln R \leqslant \ln r) = P\left[N_f \leqslant \frac{2}{\beta}\ln r - \frac{2}{\beta}\ln S + N\right]$$

$$= 1 - \exp\left[-\frac{r^* - a^*}{b^*}\right]^{c_n} \quad (5.3\text{-}341)$$

式中，$r^* = \ln r$，$a^* = \ln S + \frac{\beta}{2}(N - a_n)$，$b^* = \frac{\beta}{2}b_n$。

② $n>2$ 对式（5.3-336）从 N 到 N_f 进行积分，可以得到

$$R^{n-2} = S^{n-2} + \gamma(n-2)(N_f - N) \quad (5.3\text{-}342)$$

同样，假定 N_f 符合威布尔分布，则可以得到

$$P(R^{n-2} \leqslant r^{n-2}) = 1 - \exp\left[-\frac{r'' - a''_n}{b''_n}\right]^{c_n} \quad (5.3\text{-}343)$$

式中，$r'' = r^{n-2}$；$a''_n = S^{n-2} + \gamma(n-2)(a_n - N)$；$b''_n = \gamma(n-2)b_n$。

对于 $n=3$ 的特殊情况，从式（5.3-336）可知，剩余强度降低速率是一个常数，即剩余强度将随着循环数 N 线性地降低，在这种情况下

$$P(R \leqslant r) = 1 - \exp\left[-\frac{r'' - a''_n}{b''_n}\right]^{c''_n} \quad (5.3\text{-}344)$$

式中，$a''_n = S + \frac{1}{2}\alpha\beta K_c(a_n - N)$；$b''_n = \frac{1}{2}\alpha\beta K_c b_n$；$c''_n = c_n$。

（2）$P-R-S-N$ 曲线的作法

在前面的 $P-S-N$ 曲线的作法中，我们已经总结指出，它是由 $S-N$ 曲线（由实验得来的函数关系，在 $P=50\%$ 的情况下）和 $P-N$（在一定的交变应力下测定一组试样的断裂寿命 N_f）曲线组合而得到。结合的方法是在不同的应力水平 S_i 下，做几组试验，然后确定 $S-N$ 中的待定参数。类似地，$P-R-N$ 曲线的作法是：先得到疲劳裂纹不同阶段的 $R-N$ 关系，然后在 N 固定的情况下做试验，以测定 $P-R$ 关系（概率分布），这类似于前边的 $P-N$ 分布。最后将两者组合，便可得到 $P-R-S-N$ 曲线。

1）$R-N$ 曲线　裂纹形成阶段的 $R-N$ 关系为

$$R = R_c + (R_0 - R_c)(1 - N/N_c)^{m'} \quad (5.3\text{-}345)$$

上式在固定的应力 S 下确定，其中 R_c、R_0 分别由实验测定，待定系数为 m' 和 N_c。

裂纹扩展阶段的 $R-N$ 关系为

在 $n=2$ 的情况下

$$R = R_c\exp\left[-\frac{\beta}{2}(N - N_c)\right] \quad (5.3\text{-}346)$$

或

$$R = S\exp\left[-\frac{\beta}{2}(N_f - N)\right] \quad (5.3\text{-}347)$$

式中，R_c，N_c 及 N_f 由实验确定，惟一的参数是 β，同样是在应力 S 固定的情况下取得上式的。

在 $n>2$ 的情况下

$$R^{n-2} = S^{n-2} + \gamma(n-2)(N_f - N) \quad (5.3\text{-}348)$$

式中，N_f，n 由试验确定，惟一的参数是 γ。上式也是在固定的 S 下得到的。

2）$P-R$ 曲线　对于初始强度 R_0，其分布为

$$P(R_0 \leqslant \gamma_0) = 1 - \exp\left[-\frac{r_0 - a}{b}\right]^{c} \quad (5.3\text{-}349)$$

对于裂纹形成阶段

$$P(R \leqslant \gamma) = 1 - \exp\left[-\frac{r_0 - a'}{b'}\right]^{c} \quad (5.3\text{-}350)$$

式中，$a' = R_c + (R_0 - R_c)(1 - N/N_c)m'$；$b' = b(1 - N/N_c)m'$。

对于裂纹扩展阶段

$$n=2, P(\ln R \leqslant \ln r) = P\left[N_f \leqslant \frac{2}{\beta}\ln r - \frac{2}{\beta}\ln S + N\right]$$

$$= 1 - \exp\left[-\frac{r^* - a^*}{b^*}\right]^{c_n} \quad (5.3\text{-}351)$$

式中，$r^* = \ln r$；$a^* = \ln S - \frac{\beta}{2}(N - a_n)$；$b^* = \frac{\beta}{2}b_n$。

$$n>2, P(R^{n-2} \leqslant r^{n-2}) = 1 - \exp\left[-\frac{r'' - a''_n}{b''_n}\right]^{c_n} \quad (5.3\text{-}352)$$

式中，$r'' = r^{n-2}$；$a''_n = S^{n-2} + \gamma(n-2)(a_n - N)$；

$b''_n = \gamma(n-2)b_n$

a_n、b_n 和 c_n 都是 N_f 的威布尔分布的参数，即

$$P(N_f \leqslant N) = 1 - \exp\left(\frac{N - a_n}{b_n}\right)^{c_n} \quad (5.3\text{-}353)$$

3）$P-R-S-N$ 曲线　为了求得 $P-R-N$ 曲线，首先需要从试验中求得 $P-R$ 曲线，然后再与已知的 $R-N$ 曲线结合。而求 $P-R$ 曲线，则关键需要求初始强度 R_0 的分布式（5.3-333）和最终寿命 N_f 的分布式（5.3-340），其他的三个分布［式（5.3-334）、式（5.3-341）、式（5.3-343）］的分布参数便可依此而求得。

3.6 安全系数的统计分析

3.6.1 经典意义下的安全系数

在常规设计中，安全系数被定义为材料的强度（如屈服极限、强度极限或疲劳极限等）除以构件中的实际应力。这个实际应力是指构件最薄弱环节（截面）上的最大应力。但是材料强度具有离散性，即同一种材料在相同的热处理制度和试验条件下，其强度数值呈现了不同程度的波动；而且，在薄弱环节上的最大应力在不同的条件下也在变动，即使是静载荷，这种变动也是存在的。因此，上述安全系数的定义就具有某些不确定性，在具体应用（设计）的场合，还必须进一步明确“强度”和“实际应力”的取值方法。比较常见的有：取强度和实际应力的平均值，这时便可得到平均安全系数 $\bar{n}$。

$$\bar{n} = \frac{\bar{r}}{\bar{S}} \quad (5.3\text{-}354)$$

式中，$\bar{r}$ 为材料强度的平均值；$\bar{S}$ 为薄弱环节上最大应力在不同时间的平均值。

也有取强度的极小值 r_{min} 和应力 S_{max} 来定义安全系数的，即

$$n = \frac{r_{min}}{S_{max}} \quad (5.3\text{-}355)$$

有时亦采用强度的平均值 $\bar{r}$ 和应力的极大值 S_{max} 之比来确定安全系数。

上述经典意义下的安全系数，其特点是直观、简便并具有一定的实践依据，所以一直沿用。但它们也具有很大的经验性和盲目性；而且 r_{min} 和 S_{max} 无明确的定量概念；更为重要的是，经典的安全系数没有和构件的破坏概率相联系。因此，对于关键构件，其安全性要求是很高的，上述安全系数的应用就受到了限制。其实，只有当强度和工作应力的离散性很小时，式（5.3-355）定义的安全系数才有意义；而当强度和工作应力的离散性很大时，式（5.3-355）就无法定义，此时，式（5.3-354）定义的平均安全系数也失去了意义。

3.6.2 可靠性意义下的安全系数

对经典意义下的安全系数引入可靠性，便成为可靠性意义下的安全系数。它已成为可靠性方法来校核常规设计的构件的强度安全性的一个指标。

从可靠性理论来看，材料强度和构件的工作应力都服从某一概率分布的随机变量，经典安全系数定义中的强度和工

作应力的不同含义，可以用在一定概率（或可靠度，或存活率）下随机变量的取值来明确规定。如强度极小值 r_{min} 可规定为可靠度 $R=95\%$ 时的下限值，而工作应力的极大值 S_{max} 可规定为可靠度 $R=99\%$ 的上限值，于是

$$n_R=\frac{r_{min}}{S_{max}}=\frac{r_{95}}{S_{99}}=\frac{(1-1.65C_r)\bar{r}}{(1+2.33C_s)\bar{S}} \quad (5.3\text{-}356)$$

式中，n_R 表示可靠度意义下的安全系数；r_{95} 表示可靠度 $R=95\%$ 时强度的下限值，$r_{95}=(1-1.65C_r)\bar{r}$，其中 $\bar{r}$ 为平均强度，$c_r=\frac{\sigma_r}{\bar{r}}$ 称为强度的变差系数，σ_r 为强度的标准差；S_{99} 表示可靠度 $R=99\%$ 时应力的下限值，$S_{99}=(1+2.33C_S)\bar{r}$，其中 $\bar{S}$ 为平均应力，$c_S=\frac{\sigma_S}{\bar{S}}$ 称为应力的变差系数，σ_S 为应力的标准差。

因此可知，式（5.3-354）定义的平均安全系数是取可靠度 $R=50\%$ 时的强度和应力之比。有了可靠性的安全系数，就可以进一步研究它和构件的可靠度之间的关系。

假设强度和应力的分布均符合正态分布，则可得到联系强度、应力和构件可靠度三者之间的关系。

$$u=\frac{\bar{r}-\bar{S}}{\sqrt{\sigma_r^2+\sigma_S^2}} \quad (5.3\text{-}357)$$

式中，u 为可靠性指数，它与构件的可靠度 R 对应，可查标准分布表。其余符号的意义同前。

由上式可得到：

$$\frac{\bar{r}}{\bar{S}}=\frac{1+u\sqrt{c_r^2+c_S^2-u^2c_r^2c_S^2}}{1-u^2c_r^2} \quad (5.3\text{-}358)$$

这就是联系强度变差系数 c_r，应力变差系数 c_S，构件平均安全系数 $\frac{\bar{r}}{\bar{S}}$ 和可靠度 R（即 u）的方程式。代入式（5.3-356）可得到：

$$n_R=\frac{(1-1.65C_r)}{(1+2.33C_s)}\frac{1+u\sqrt{c_r^2+c_S^2-u^2c_r^2c_S^2}}{1-u^2c_r^2} \quad (5.3\text{-}359)$$

上式为可靠性意义下的安全系数与强度和应力的变差系数以及构件的可靠度 R（即与之相应的 u）之间的关系，若已知 c_r，c_S 和 R 就可求得 n_R 值。当应力或强度的离散性很小时，则上式可简化为

$$\left.\begin{aligned}&\text{若 } c_S\approx0,\text{则 } n_R=\frac{1-1.65c_r}{1-uc_r}\\&\text{若 } c_r\approx0,\text{则 } n_R=\frac{1+uc_S}{1+2.33c_S}\end{aligned}\right\} \quad (5.3\text{-}360)$$

为了与常规设计中的平均安全系数相比较，国外早期的资料中常引用下式：

$$n_R=n_{50}=\frac{\bar{r}}{\bar{S}}=\frac{1+u\sqrt{c_r^2+c_S^2-u^2c_r^2c_S^2}}{1-u^2c_r^2} \quad (5.3\text{-}361)$$

有的文献则推荐用下述可靠性安全系数的表达式：

$$\begin{aligned}n'_R&=\frac{\bar{r}}{S_{99}}=\frac{\bar{r}}{(1+2.33c_S)\bar{S}}\\&=\frac{1}{(1+2.33c_S)}\cdot\frac{1+u\sqrt{c_r^2+c_S^2-u^2c_r^2c_S^2}}{1-u^2c_r^2}\end{aligned} \quad (5.3\text{-}362)$$

由于式（5.3-359）所表达的安全系数既与常规的安全系数接近［即与式（5.3-355）接近］，又能与试验检验相联系，因此它也被广泛地使用。

3.6.3 安全系数与可靠度的关系

关于安全系数与可靠度的关系，有以下两点必须指出。

1）从运行函数表达的联结方程来看，

$$\because\quad u=\frac{\mu_z}{\sigma_z}=\frac{\ln\mu_R-\ln\mu_S}{\sqrt{\sigma_R^2+\sigma_S^2}}$$

$$\therefore\quad \ln\mu_R-\ln\mu_S=u\sqrt{\sigma_R^2+\sigma_S^2}$$

$$\ln\frac{\mu_R}{\mu_S}=u\sqrt{\sigma_R^2+\sigma_S^2}$$

$$\therefore\quad n_{50}=\frac{\mu_R}{\mu_S}=\exp[u\sqrt{\sigma_R^2+\sigma_S^2}] \quad (5.3\text{-}363)$$

① 当强度和应力的标准差不变时，提高平均安全系数就会提高可靠性；

② 当强度和应力的平均值的比值不变时，缩小它们的离散性，即减小 σ_R 和 σ_S，就可提高构件的可靠度。

2）要注意输入误差对可靠度的影响。如果要得到一个较好的可靠度估计值，必须严格控制输入量 μ_Y 和 σ_Y 的精度，因为它们对可靠度的影响是很灵敏的。

3.6.4 安全系数的统计分析方法

前面介绍的是在经典的安全系数中引入可靠度，也称为可靠性安全系数方法，它是可靠性设计和常规设计之间联系的桥梁。现在，要将安全系数 n 当作随机变量，而且把它看作是强度 r 和应力 S 这两个随机变量的函数，即定义

$$n=\frac{r}{S} \quad (5.3\text{-}364)$$

存在下列概率不等式：

$$P(|n-a|\leqslant\varepsilon)\geqslant1-\frac{E[(n-a)^2]}{\varepsilon^2} \quad (5.3\text{-}365)$$

式中，a 和 ε 均为大于零的任意常数。

对式（5.3-365）的证明如下：

先考察下式

$$E\{(n-a)^2\}=\left\{\int_{-\infty}^{\infty}(n-a)^2f(n)\mathrm{d}n\geqslant\int_{n^*}(n-a)^2f(n)\mathrm{d}n\geqslant\int_{n^*}(n-a)^2f(n)\mathrm{d}n\right. \quad (5.3\text{-}366)$$

其中积分限 n^* 表示积分仅对 n 的这些值进行，即满足 $|n-a|>\varepsilon$ 的 n 值。由于

$$\int_{n^*}(n-a)^2f(n)\mathrm{d}n>\varepsilon^2\int_{n^*}f(n)\mathrm{d}n=\varepsilon^2P(|n-a|)>\varepsilon \quad (5.3\text{-}367)$$

结合式（5.3-366）和式（5.3-367）可得

$$\frac{1}{\varepsilon^2}=E\{(n-a)^2\}>P(|n-a|>\varepsilon)$$

$$\therefore\quad \frac{1}{\varepsilon^2}E\{(n-a)^2\}>1-P(|n-a|<\varepsilon)$$

即

$$P(|n-a|<\varepsilon)\geqslant1-\frac{1}{\varepsilon^2}E\{(n-a)^2\} \quad (5.3\text{-}368)$$

令 $a=k\bar{n}$，因此有

$$\begin{aligned}E\{(n-a)^2\}&=E\{(n-k\bar{n})^2\}=E\{n^2-2kn\bar{n}+k^2\bar{n}^2\}\\&=\bar{n}^2\left[\frac{\sigma_n^2}{\bar{n}^2}+(1-k)^2\right]=\bar{n}^2[c_n^2+(1-k)^2]\end{aligned} \quad (5.3\text{-}369)$$

其中 $c_n=\frac{\sigma_n}{\bar{n}}$ 是安全系数的变差系数。

把不等式（5.3-368）改写成

$$P(a-\varepsilon\leqslant n\leqslant a+\varepsilon)\geqslant1-\frac{1}{\varepsilon^2}E\{(n-a)^2\}$$

令 $a-\varepsilon=1$，则上式可写为

$$P(1\leqslant n\leqslant2k\bar{n}-1)\geqslant1-\frac{\bar{n}^2[c_n^2+(1-k)^2]}{(k\bar{n}-1)^2} \quad (5.3\text{-}370)$$

按照可靠性定义，可得

$$R=P(n\geqslant 1)$$

结合上式和式（5.3-370）可得

$$R\geqslant 1-\frac{\bar{n}^2[c_n^2+(1-k)^2]}{(k\bar{n}-1)^2} \tag{5.3-371}$$

不等式右边代表 R 的下限，为了得到最高的下限，求下式的极小值：

$$W=\frac{\bar{n}^2[c_n^2+(1-k)^2]}{(k\bar{n}-1)^2}$$

可得

$$k^*=\frac{\bar{n}(c_n^2+1)-1}{\bar{n}-1}$$

同时可验证 $\frac{\partial^2 W}{\partial^2 k}<0$。把上述 k^* 代入式（5.3-370）可得

$$R\geqslant 1-\frac{\bar{n}^2c_n^2}{\bar{n}^2c_n^2+(n-1)^2} \tag{5.3-372}$$

或

$$\bar{n}\geqslant\frac{1}{1-c_n\sqrt{\frac{R}{1-R}}} \tag{5.3-373}$$

上面两式表示了平均安全系数 $\bar{n}$，安全系数的变差系数 c_n 和可靠度 R 的关系。式（5.3-373）表示平均安全系数 $\bar{n}$ 的下限值，可由此求得安全系数的范围：

$$1\leqslant n\leqslant 2k^*\bar{n}-1 \tag{5.3-374}$$

其可靠度为 R。

4 以可靠性工程为基础的失效预测技术和方法

4.1 可靠性的基本概念和数学基础

失效分析时，结构的危险程度用失效概率描述，结构的安全程度用可靠度描述。有了失效概率和可靠度的概念，就可以定量描述危险程度和安全程度。

4.1.1 可靠性的基本概念

可靠性是产品的一个质量指标，这个指标具有如下的特点：

1）产品的可靠性指标是与时间有关的一个参量；

2）这个指标是与失效方式、失效定义有关的综合指标；

3）产品的可靠性可分为“固有可靠性”和“使用可靠性”两种，前者决定于设计和生产，后者则与使用条件有关。

产品在规定的条件下，在规定的时间内，完成规定功能的能力，就叫做产品的可靠性。

1）产品的规定条件　如温度，气压，震动，冲击，介质，载荷等；

2）产品的规定时间　对保持产品的质量和性能要有一定的时间要求，对于不同的零件，其时间要求也不一样；

3）产品的规定功能　何种功能，用何种指标来衡量，什么是失效，对于每一个产品都要有明确的定义；

4）在规定的条件、规定的时间和规定的功能下，产品可能完成任务，也可能不能完成任务，我们称之为随机事件，而随机事件可以用概率来定量地描述，因此，使用了概率来度量产品能否完成规定的功能。于是，产品在规定的条件下，规定的时间内，完成规定的功能的概率就叫做产品的可靠度，即可靠性的概率度量。

4.1.2 可靠性指标的定义

设产品总数为 N_0，累计失效数为 N_f，仍正常工作数为 N_s，则定义

$$\bar{R}=\frac{N_s(t)}{N_0} \tag{5.3-375}$$

为存活频率。当 $N_0\to\infty$ 时，$\lim\limits_{N_0\to\infty}\bar{R}=R(t)$，则 $R(t)$ 为产品至时刻 t 的可靠度。

下面介绍其他有关可靠性的统计指标。

（1）失效（故障）密度

设 N_0 是试验产品总数，ΔN 是时刻 t 到 $t+\Delta t$ 时间间隔内产生的失效产品数，比值

$$\bar{f}=\frac{\Delta N(t)}{N_0\Delta t} \tag{5.3-376}$$

称为 t 到 $t+\Delta t$ 时间间隔内的平均失效（故障）密度，它表示这段时间间隔内平均单位时间的故障频率。若 $N_0\to\infty$，$\Delta t_i\to 0$，则频率元变为概率元。式（5.3-376）变为瞬时失效（故障）密度：

$$f(t)=\frac{1}{N_0}\frac{\mathrm{d}N}{\mathrm{d}t} \tag{5.3-377}$$

（2）失效率（故障率或故障强度）

设 N_0 是试验产品总数，$N_f(t)$ 是 t 时刻的累积故障数，$\Delta N(t)$ 是时间间隔 $t\sim t+\Delta t$ 内出现的故障数，则定义

$$\bar{\lambda}(t)=\frac{\Delta N(t)}{\Delta t}\cdot\frac{1}{N_0-N_f(t)} \tag{5.3-378}$$

为时间间隔 $t\sim t+\Delta t$ 内的平均故障强度或平均失效率。当 $\Delta t\to 0$，$N_0\to\infty$ 时，它就是时刻 t 的失效率 $\lambda(t)$，即

$$\lambda(t)=\lim_{\substack{\Delta t\to 0\\ N_0\to 0}}\bar{\lambda}(t)=\frac{\mathrm{d}N(t)}{\mathrm{d}t}\cdot\frac{1}{N_0-N_f(t)} \tag{5.3-379}$$

（3）平均寿命

有时更为关心的是产品从投入运行到产生失效（故障）的平均工作时间——平均寿命有多大。定义

$$T_0=\int_0^{\infty}tf(t)\mathrm{d}t \tag{5.3-380}$$

为产品的平均寿命。

（4）寿命方差

寿命方差代表了产品寿命的离散程度。定义

$$V(t)=\int_{-\infty}^{+\infty}(t-T)^2f(t)\mathrm{d}t$$

为寿命方差，它是关于子样均值 T 的二次中心矩，是度量各寿命观测值对平均值 T 的离散程度的一个量。它和 T 结合起来，能综合评定产品的可靠度。

（5）可靠寿命、中位寿命和特征寿命

用产品的寿命指标来描述其可靠性的还有可靠寿命、中位寿命和特征寿命，它们的定义分别如下：

1）可靠寿命　在某些场合需要知道在给定的可靠度 $R(t)=a$ 时产品的寿命，这时只要利用其反函数反解出 t，得

$$t_a=R^{-1}(a)$$

式中，t_a 为可靠度 $R=a$ 时的寿命。

2）中位寿命　若令可靠度 $R(t)=0.5$ 即可得到中位寿命 $t_{0.5}$。

3）特征寿命　若令可靠度 $R(t)=e^{-1}$ 即可得到中位寿命 $t_{e^{-1}}$。

4.1.3 可靠度（或失效概率）的计算方法

1）可靠度的一般表达式　一般用 s 表示应力，用 r 表示强度，并用 $f_s(s)$ 和 $f_r(r)$ 分别表示应力和强度的概率密度函数，见图 5.3-59。

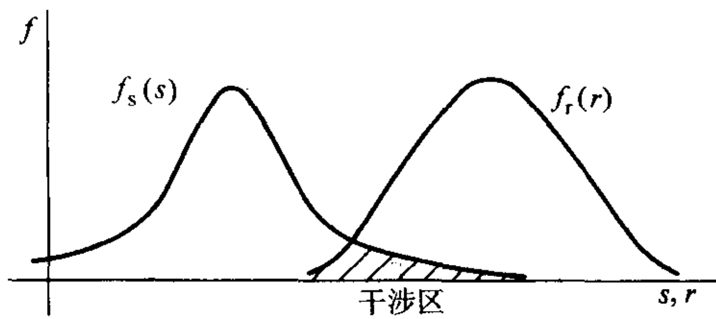


图 5.3-59 强度－应力干涉图

定义

$$R = P(r > s) = P[(r - s) > 0] \quad (5.3\text{-}381)$$

图 5.3-59 中阴影部分表示干涉面积，它表示了失效概率的大小。现考虑应力落在 ds 间隔内的概率，见图 5.3-60。可计算得

$$P\left(s_0 - \frac{ds}{2} \leqslant s \leqslant s_0 + \frac{ds}{2}\right) = f_s(s_0)ds$$

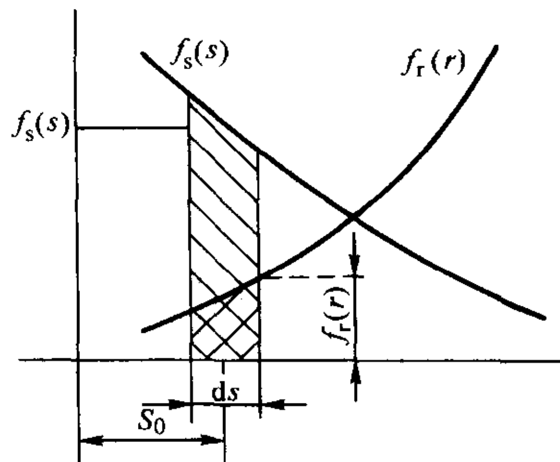


图 5.3-60 可靠度计算－干涉部分放大图

而强度 r 大于 s_0 的概率为

$$P(r > s_0) = \int_{s_0}^{\infty} f_r(r)dr$$

假定这是两个独立的随机事件，则它们同时发生的概率为

$$f_s(s_0)ds\int_{s_0}^{\infty} f_r(r)dr$$

上式对 s_0 取任意值时均成立，所以

$$R = \int_{-\infty}^{\infty} f_s(s)\left[\int_{s}^{\infty} f_r(r)dr\right]ds \quad (5.3\text{-}382)$$

同理可以得到

$$R = \int_{-\infty}^{\infty} f_r(r)\left[\int_{-\infty}^{r} f_s(s)ds\right]dr \quad (5.3\text{-}383)$$

式（5.3-382）和式（5.3-383）提供了在已知强度和应力的分布密度函数后，计算可靠度的一般表达式。

2）两种特殊情况

① 强度和应力均为正态分布的情况　如果强度 r 和应力 s 都服从正态分布，且相互独立，设

$$r \sim N(\mu_r, \sigma_r^2)$$
$$s \sim N(\mu_s, \sigma_s^2)$$

则有 $r - s \sim N(\mu_r - \mu_s, \sigma_r^2 + \sigma_s^2)$

则可计算得

$$\begin{aligned} R &= P(r - s > 0) \\ &= P\left[\frac{(r-s)-(\mu_r-\mu_s)}{\sqrt{\sigma_r^2+\sigma_s^2}} > -\frac{\mu_r-\mu_s}{\sqrt{\sigma_r^2+\sigma_s^2}}\right] \\ &= P\left[\frac{(r-s)-(\mu_r-\mu_s)}{\sqrt{\sigma_r^2+\sigma_s^2}} \leqslant \frac{\mu_r-\mu_s}{\sqrt{\sigma_r^2+\sigma_s^2}}\right] \\ &= \Phi\left(\frac{\mu_r-\mu_s}{\sqrt{\sigma_r^2+\sigma_s^2}}\right) \end{aligned} \quad (5.3\text{-}384)$$

式中，$\Phi(\cdot)$为标准正态分布函数。

② 强度和应力都服从对数正态分布　如果强度 r 和应力 s 都服从对数正态分布，且相互独立，设

$$\ln r \sim N(\mu_{\ln r}, \sigma_{\ln r}^2)$$
$$\ln s \sim N(\mu_{\ln s}, \sigma_{\ln s}^2)$$

可知

$$\ln r - \ln s \sim \ln(r/s) \sim N(\mu_{\ln r} - \mu_{\ln s}, \sigma_{\ln r}^2 + \sigma_{\ln s}^2)$$

则可计算得

$$R = P(r/s) > 1 = P(\ln r - \ln s) > 0 = \Phi\left(\frac{\mu_{\ln r} - \mu_{\ln s}}{\sqrt{\sigma_{\ln r}^2 + \sigma_{\ln s}^2}}\right) \quad (5.3\text{-}385)$$

4.2 静强度可靠性的失效预测技术和方法

常规设计中的静强度设计，指的是结构（或构件）所受的载荷和结构的抗力都不随时间而变化的设计问题。显然，这种处理方法是理想化的。事实上，结构所受到的载荷不论在出现的频次、作用时间的长短、幅值的大小等方面均随着时间在变化；而结构的抗力不仅随着时间在变化，而且还随着材料的冶炼工艺、热处理规范、加工条件等在变化。按照事物的本质，载荷和抗力都是随机过程，应按随机过程的理论和分析方法来处理，然后进行设计。但是这样处理的复杂程度大为增加，而且对于那些载荷和抗力随时间变化不大的结构来说，常规的静强度设计已抓住了主要矛盾，实践证明它是行之有效的。

然而，随着生产建设的发展，一部分原来按静强度设计的结构（或构件）在使用中就暴露出一些问题，或是强度不足而断裂，或是寿命不足而失效。对这些问题的分析说明，问题的关键在于对结构的载荷和抗力的变化考虑不够，估计不足。人们现在已认识到，对于那些主要承受静载荷的结构（如房屋）和机械（如机床），还必须考虑它们受到的动载荷（如风载、雪载等），如何综合地计及不同类型的载荷，使设计的结构（或构件）达到经济、安全和可靠，这就需运用以概率为基础的可靠性设计。

金属材料的静载力学性能（静强度）和结构的静载抗力，相对来说是比较稳定的。但是由于原材料、冶炼、热处理、加工等条件的不同，各地生产的（或同一厂不同时间生产的）同一牌号的金属材料，其静态力学性能也不尽相同，存在着差异。而对于一些新型材料，如复合材料、陶瓷材料，其静强度波动就很大，必须经过统计得到其概率分布，以便为可靠性设计作好准备。

4.2.1 载荷的统计分析

（1）静载荷与动载荷

静载和动载的划分要根据它对机械零件和系统的作用来进行，通常用应变速率这个参量来区分载荷的动、静程度，如表 5.3-13 所示。

表 5.3-13 载荷划分范围

载荷	蠕变	静载荷	动载荷	冲击	超高速冲击
应变速率 /s^{-1}	$0 \sim 10^{-5}$	$10^{-5} \sim 10^{-1}$	$10^{-1} \sim 5$	$5 \sim 10^4$	$> 10^4$
常用的加载方法	恒载荷或恒应力机械	液压机械	快作用液压机械或气动机械	机械冲击或爆炸性冲击	充气枪或炸弹
试验中考虑的动力因素	记录的应变与时间或蠕变率	恒应变率试验	试样和机械中的机械共振	弹塑性波传播	冲击波或炸弹
说明	惯性力可忽略不计		惯性力显著		

很多建筑结构，如房屋、水坝、桥梁以及很多机械构件或结构如车床、机床、转轴、容器等，常规设计时均按静强度处理。

大量统计证明，静载荷可用正态或对数正态分布来描述，见图5.3-61a，而一般的动载荷可用正态或对数正态来描述，见图5.3-61b。

（2）载荷与载荷效应

结构受到载荷 P 的作用后，在各构件的计算截面上就产生内力 Q（可以是弯矩 M 或轴向力 N），这种内力就称为载荷效应。最简单的情况是，载荷效应与载荷成线性关系：

$$Q = kP \tag{5.3-386}$$

式中，k 为常数，它与结构的几何和载荷类型有关。

由上式可知，由载荷的统计特性参数可以求得载荷效应的统计特性参数，其平均值和变差系数为

$$\mu_Q = k\mu_P \tag{5.3-387a}$$

$$c_Q = c_P \tag{5.3-387b}$$

当结构存在几何非线性（如大变形）或物理非线性（如材料的弹塑性）时，式（5.3-386）的线性关系便不存在。

有文献认为，载荷效应 Q 与载荷 P 的关系可取为

$$Q = k'BP \tag{5.3-388}$$

式中，k' 为影响系数，是等效静力载荷转换成载荷效应的系数；B 为模型化参数，表示实际变动载荷转换成等效静载荷的系数。假设 k'，B，P 之间相互独立，则

$$\mu_Q = \mu_{k'}\mu_B\mu_P \tag{5.3-389a}$$

$$c_Q = \sqrt{c_{k'}^2 + c_B^2 + c_P^2} \tag{5.3-389b}$$

这里，把参数 k'，B 均作为随机变量，而且从式（5.3-388）可知，载荷效应的统计规律不能简单地用载荷的统计规律来表示，说明式（5.3-388）比式（5.3-386）更合理。

（3）载荷的统计方法

从载荷-时间历程来看，较典型的有如图5.3-61所示的三种。图中（a）表示的是永久载荷，如结构的自重等。它随时间的变化很小，因而其概率密度函数曲线较陡。图中（b）表示的是持续载荷，如房屋的楼面活载荷和机械的移动载荷等。它的特点是载荷幅值在一定的时间内基本恒定，也可以是零（即不出现），且各个时段长短不一。图中（c）是短期瞬时载荷，如最大的风载和地震力。这种载荷的作用时间很短，瞬时的幅值较大。

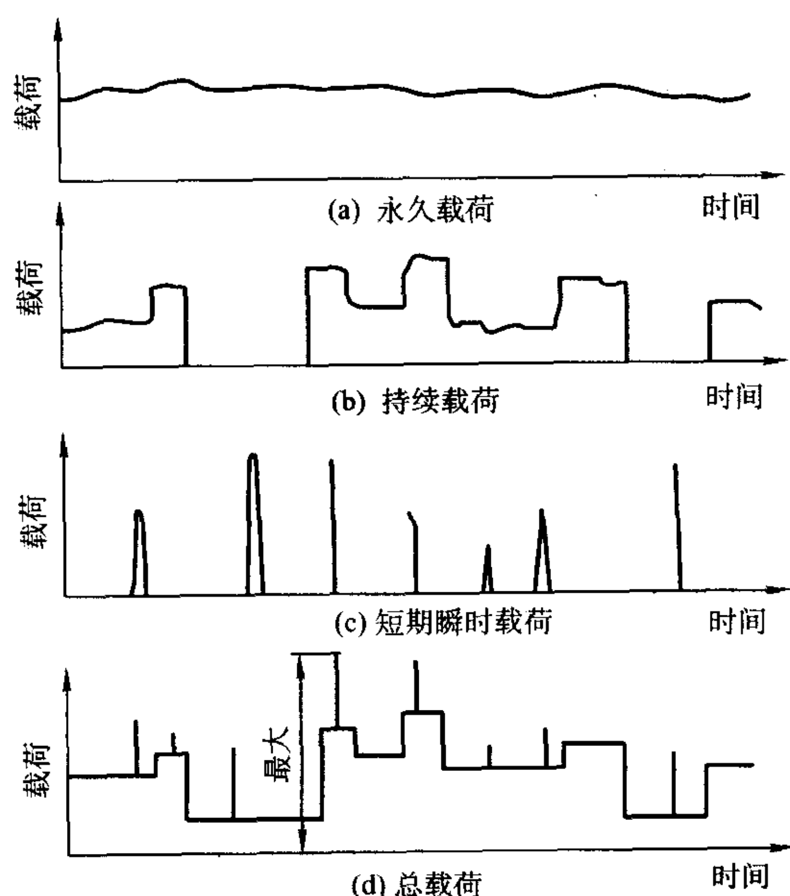


图5.3-61 三种典型的载荷-时间模型

应该指出，每个结构在设计时都有一个预期的寿命，叫做设计基准期 T，例如，我国的“建筑结构设计统一标准”中建议，对于一般的工业与民用建筑，取 $T=50$ 年。载荷统计的目的就是要求出在 T 期间的最大载荷值 Q_T，它被当成一个随机变量，因而进一步要求得 Q_T 的概率分布。下面介绍两种求 Q_T 分布的方法。

1）平稳二项随机过程模型（即等时段矩形波模型）

该模型的基本假设为：①按变动载荷一次作用在结构上的时间长短，将设计基准期 T 分为 r 个相等的时段 τ，即 $\tau=\frac{T}{r}$；②在时段 τ 内，载荷 $Q(t)$ 出现的概率为 p，不出现的概率为 $q=1-p$（p，q 均为常数）；③设 $Q(t)$ 为非负的随机变量，它在每个时段 τ 上的概率分布是相同的，记为 $F_i(x)=P[Q(t)\leqslant x, t\in\tau](x>0)$；④不同时段 τ 上 $Q(t)$ 的幅值是相互独立的，并且与在时段 τ 上是否出现载荷，也是相互独立的。

根据上述假设，我们可将图5.3-62所示随机过程 $Q(t)$ 转换为 T 年内的载荷最大值 Q_T，其方法如下。

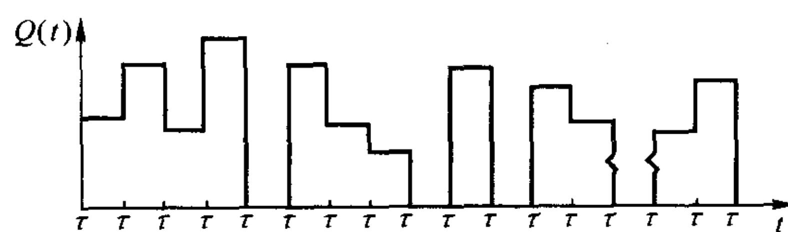


图5.3-62 等时段矩形波

① 在任一时段 τ，载荷的概率分布 $F_\tau(x)$ 可表达为

$$\begin{aligned}
&F_\tau(x)=P[Q(t)\leqslant x, t\in\tau]\\
&=P[Q(t)\neq 0]P[Q(t)\leqslant x, t\in\tau \mid Q(t)\neq 0]+P[Q(t)=0]\\
&\quad P[Q(t)\leqslant x, t\in\tau \mid Q(t)=0]\\
&=pF_i(x)+q\times 1=1-p[1-F_i(x)](x\geqslant 0)
\end{aligned}$$

② 有了每一时段 τ 概率分布 $F_\tau(x)$，则 T 年内载荷最大值 Q_T 的概率分布 $F_T(x)$ 为

$$\begin{aligned}
&F_\tau(x)=P(Q(t)\leqslant x)\\
&=P[\max_{t\in[0,T]}Q(t)\leqslant x]=pF_i(x)+q\times 1=1-p[1-F_i(x)](x\geqslant 0)\\
&=\prod_{j=1}^{r}P[Q(t)\leqslant x, t\in\tau_j]\\
&=\{1-p[1-F_i(x)]\}^r(x\geqslant 0)
\end{aligned} \tag{5.3-390}$$

③ 设可变载荷在 T 年内平均出现 m 次，$m=pr$，当 $p=1$ 时，$m=r$，则由式（5.3-390）可得

$$F_T(x)=\{F_i(x)\}^m \tag{5.3-391}$$

当 $p<1$ 时，如式（5.3-390）中的 $p[1-F_i(x)]$ 充分小，由 $e^{-x}=1-\frac{x}{1!}+\frac{x}{2!}-\cdots$ 的等式，从式（5.3-390）可得

$$\begin{aligned}
F_T(x)&=\{1-p[1-F_i(x)]\}^r\approx\{e^{-p[1-F_i(x)]}\}^r\\
&=\{e^{-[1-F_i(x)]}\}^{pr}\approx\{1-[1-F_i(x)]\}^{pr}\\
&\therefore F_T(x)\approx[F_i(x)]^m(x\geqslant 0)
\end{aligned} \tag{5.3-392}$$

在应用上式时，需要由实测资料得出 τ 及 $F_i(x)$。

有其他文献指出，上述方法有如下缺点：①它把基准期 T 年内载荷随即过程转化成等时段 τ 内的矩形波载荷，即在各个 τ 时载荷峰值保持不变。这一假设对持续性载荷比较适合，而对最大风压、临时性活载荷及短期瞬时载荷，这个假设与实际不符；②在推导式（5.3-391）和式（5.3-392）时，涉及到时段数 r 和概率 p，而 r 和 p 一般很难从载荷的统计资料中得到，只能靠主观做出判断。

2）极值统计方法 假设：①将设计基准期 T 年分为 N

个时段，即 $\tau=\frac{T}{N}$，τ 时段内的最大载荷 Q_i 的分布为 $F_i(x)$；②各时段的 Q_i 相互独立，且具有相同的分布 $F_i(x)$。根据极值分布理论，T 年内的最大载荷 Q_T 的分布 $F_T(x)$ 为

$$F_T(x)=P(Q_T\leqslant x)=P(\max_{1\leqslant i\leqslant N}Q_i\leqslant x)$$
$$=P(Q_1\leqslant x)P(Q_2\leqslant x)\cdots P(Q_N\leqslant x)$$
$$=\prod_{i=1}^{N}P(Q_i\leqslant x)=[F_i(x)]^N \qquad (5.3\text{-}393)$$

上式与式（5.3-391）和式（5.3-392）在形式上是一样的，但在统计意义上是有区别的。式（5.3-393）实现过程是：将设计基准期 T 等分为 N 个时段 τ，注意 τ 的时间不要太短，以免影响相邻时段内最大载荷的相关性；统计 N 个时段内的最大载荷 Q_i，即对 Q_1，Q_2，…，Q_N 作出频率直方图，判断其概率分布 $F_i(x)$ 的类型；根据 $F_i(x)$ 的类型估计其分布参数；按式（5.3-393）求得 Q_T 的概率分布 $F_T(x)$，于是可由 $F_i(x)$ 的统计参数 μ_{X_i}，σ_{X_1} 求得 $F_T(x)$ 的统计参数 μ_{X_T}，σ_{X_T}。

在一些特殊的分布情况下，估计 $F_T(x)$ 参数的方法大为简化，可用下列近似式：

① $F_i(x)$ 为正态分布　当 $F_i(x)$ 为正态分布时，则 $F_T(x)$ 也近似服从正态分布，它们分布参数间的关系为

$$\mu_{X_T}\approx\mu_{X_i}+3.5(1-1/\sqrt[4]{N})\sigma_{X_i} \qquad (5.3\text{-}394a)$$

$$\sigma_{X_T}\approx\sigma_{X_1}/\sqrt[4]{N} \qquad (5.3\text{-}394b)$$

② $F_i(x)$ 为极值Ⅰ型分布　这时 $F_T(x)$ 也服从极值Ⅰ型分布，其统计参数之间的关系为

$$\mu_{X_T}=\mu_{X_i}+(\ln N)/\alpha \qquad (5.3\text{-}395a)$$

$$\sigma_{X_T}=\sigma_{X_i} \qquad (5.3\text{-}395b)$$

③ $F_i(x)$ 为极值Ⅱ型分布　这时 $F_T(x)$ 也服从极值Ⅱ型分布，二者统计参数之间的关系为

$$\mu_{X_T}=\mu_{X_1}N^{\frac{1}{k}} \qquad (5.3\text{-}396a)$$

$$\sigma_{X_T}=\sigma_{X_1}N^{\frac{1}{k}} \qquad (5.3\text{-}396b)$$

变差系数的关系是

$$C_{X_T}=C_{X_1} \qquad (5.3\text{-}396c)$$

上面式（5.3-395）中的 α 和式（5.3-396）中的 k，分别是极值Ⅰ型分布和Ⅱ型分布中的参数。

4.2.2 抗力的统计分析

对于材料和构件的抗力，在工程上一直采用统计的方法。随着实践和认识的发展，已从早期的系数发展为着重于材料强度概率分布的寻求，这项工作只有通过长期的数据积累才能奏效。

(1) 单一材料构件抗力 R 的统计分析

所谓单一材料是指牌号及热处理完全相同的钢材、铸铁、球墨铸铁或铝合金等，由单一材料制造的构件，由于尺寸不同和条件不同，其性能是有差别的，为了使结果能进行比较，必须把它们转换成属于同一个母体的参数，然后才能进行统计分析。

为此引入

$$R=\frac{R_Z}{R_K} \qquad (5.3\text{-}397)$$

式中，R_Z 为构件强度的实测值；R_K 为构件强度的设计标准值。

以工字钢为例，上式可进一步改写为：

$$R=\frac{R_Z}{R_K}=\frac{R_Z}{R_P}\frac{R_P}{\sigma_{SZ}W_{SZ}}\frac{\sigma_{SZ}}{\sigma_{SK}}\frac{W_{SZ}}{W_{SK}}=ZPMA \qquad (5.3\text{-}398)$$

式中

$$Z=\frac{R_Z}{R_P}$$

为实际构件抗力 R_Z 与实验室构件抗力 R_P 之比，R_Z 与 R_P 均为随机变量；

$$P=\frac{R_P}{\sigma_{SZ}W_{SZ}}=\frac{R_P}{R_C}$$

为实验室构件抗力 R_P 与计算抗力 R_C 之比，R_C 也是随机变量，因为计算时屈服极限 σ_{SZ} 和断面塑性抵抗矩 W_{SZ} 均取实测值；

$$M=\frac{\sigma_{SZ}}{\sigma_{SK}}$$

为实际材料屈服极限 σ_{SZ} 与标准屈服极限 σ_{SK} 之比，其中 σ_{SZ} 是随机变量，σ_{SK} 为确定的量；

$$A=\frac{W_{SZ}}{W_{SK}}$$

表示实际的与标准的断面塑性抵抗矩之比，其中 W_{SZ} 为随机变量，W_{SK} 为确定的量。

假定 Z、P、M 和 A 相互独立，则由式（5.3-398）可得 R 的平均值和变差系数。

$$\left.\begin{aligned}\overline{R}&=\mu_R=\mu_Z\mu_P\mu_M\mu_A=\overline{Z}\overline{P}\overline{M}\overline{A}\\ c_R&=\sqrt{c_Z^2+c_P^2+c_M^2+c_A^2}\end{aligned}\right\} \qquad (5.3\text{-}399)$$

式中，μ 代表某个量的平均值；c 代表某个量的变差系数。

对于上式中四个参数的统计值，其处理方法为：

1) Z 的统计值 μ_Z 和 c_Z 很难得到，一般均用经验估计值，如对于钢结构，其材质稳定，在施工安装条件及使用环境良好时可近似取 $\mu_Z\approx1.0$，$c_Z\approx0$；

2) P 的统计值 μ_P 和 c_P 可通过实验室实测的 R_P 与计算的 μ_Z 和 R_C 得到，其中 R_C 是取实际试件的材料强度和尺寸，按计算标准值 R_K 的公式进行计算；

3) M 的统计值 μ_M 和 c_M 可由 σ_{SZ} 的统计值求得，因为 σ_{SK} 是标准值，是确定的量；

4) A 的统计值 μ_A 和 c_A 可由 W_{SZ} 的统计值求得，因为 W_{SK} 是确定的常量。而 W_{SZ} 则由构件尺寸的统计值，根据 W 的表达式按统计运算方法求得。

(2) 材料静强度的概率分布

制造构件所用的材料，在化学成分、冶炼、铸造、锻造、焊接、热处理、机械加工和装配等方面存在差异，以及取样部位、试样的形式尺寸和公差、试验方法和试验环境等不尽相同，从而造成材料力学性能的离散性。所有这些因素使得材料的静强度（即静态力学性能）具有统计的特性。大量测试数据的积累和分析表明，材料的静拉伸屈服极限、强度极限、延伸率和硬度均可用正态分布来描述。下面收集的金属材料的屈服极限、强度极限、延伸率和硬度的均值和标准差数据分别列于表5.3-14、表5.3-15、表5.3-16和表5.3-17中，可供可靠性设计和分析时参考。

表5.3-14　材料屈服极限 σ_s 的标准差 S_{σ_s}

序号	材料名称	强度极限均值 σ_b /MPa	屈服极限均值 σ_s /MPa	屈服极限的标准差 S_{σ_s}		附注
				MPa	屈服极限的百分数/%	
1	碳素钢	443	667	27.5	6.21*	0.21%～0.35%C 1.28%～1.71%Mn

续表 5.3-14

序号	材料名称		强度极限均值 σ_b /MPa	屈服极限均值 σ_s /MPa	屈服极限的标准差 S_{σ_s}		附注
					MPa	屈服极限的百分数/%	
2	锰钢		418*	614	20.9	5.00*	1.19% ~ 1.37%C 1.5% ~ 12.0%Mn 0.5% ~ 2.1%Si
3	钼钢（正火）		830	935	14.06	1.69*	0.37% ~ 0.39%C 0.76% ~ 0.85%Mn 1.95% ~ 2.07%Ni 0.79% ~ 0.84%Cr 0.24% ~ 0.33%Mo
4	钼钢		1 392	1 729	90.2	6.48*	0.225% ~ 0.430%C 0.61% ~ 1.23%Mn 1.05% ~ 2.48%Ni 0.40% ~ 0.99%Cr 0.23% ~ 0.79%Mo
5	低合金钢（右为回火温度）	370℃	1 276*	1 406	57.4	4.50*	0.23% ~ 0.34%C 0.37% ~ 0.87%Mn 0.46% ~ 0.65%Cr 0.15% ~ 0.24%Mo 0.41% ~ 0.65%Ni
		454℃	1 153*	1 215	44.5	3.86*	
		538℃	1 023*	1 075	45.7	4.47*	
		620℃	907*	995	49.2	5.42*	
6	铬钼钒钢		1 444	1 749	70.3	4.87*	0.21% ~ 0.34%C 0.37% ~ 0.78%Mn 4.62% ~ 5.97%Cr 1.20% ~ 1.77%Mo 0.29% ~ 0.59%V
7	高强度合金钢		1 691	1 805	102	6.04*	0.003% ~ 0.05%C 15.0% ~ 17.8%Ni 4.28% ~ 5.10%Mo 6.40% ~ 14.10%Co 0.13% ~ 0.83%Ti 0.01% ~ 0.62%Nb
8	球墨铸铁	L4512，M4512，04512 合计	383	565	33.5	8.75	3.24% ~ 3.95%C 1.45% ~ 3.45%Si 0.17% ~ 0.67%Mn
		04128	398	601	45.0	11.30	3.37% ~ 3.84%C 2.32% ~ 3.15%Si 0.24% ~ 0.81%Mn
		L5506 M5506 合计	536.2	751.2	36.4	6.79	3.10% ~ 3.98%C 2.49% ~ 3.74%Si 0.21% ~ 0.95%Mn 0.32% ~ 0.76%Cu
		M6003 M7003 合计	531.1	847.2	50.1	9.43	3.36% ~ 3.79%C 2.50% ~ 3.12%Si 0.44% ~ 0.55%Mn 0.29% ~ 0.50%Mo
		65 - 45 - 12 系列	395.2	545.2	34.6	8.75	北美洛克维尔公司数据
		80 - 45 - 06 系列	498.5	745.0	52.6	10.55	
9	铜合金	铝青铜 A	329.4	771.6	34.6	10.50*	9.37% ~ 11.80%Al 2.74% ~ 6.05%Fe 3.07% ~ 6.26%Ni 0.31% ~ 0.63%Mn
		铝青铜 B	488.4	856	64.6	12.82*	9.37% ~ 11.80%Al 2.74% ~ 5.69%Fe 3.09% ~ 6.26%Ni 0.49% ~ 0.63%Mn

注：1. 均值上有 * 号数值，由 $\bar{x}=\frac{1}{2}(x_{max}+x_{min})$ 求得，无 * 号数值由统计得出。

2. 标准差有 * 号数值，是假设：公差 = 3 × 标准差得到的，无 * 号数值由统计得出。

表 5.3-15 材料强度极限 σ_b 的标准差 S_{σ_b}

序号	材料名称		强度极限均值 σ_b /MPa	屈服极限均值 σ_s /MPa	强度极限的标准差 S_{σ_b}		附注
					MPa	强度极限的百分数/%	
1	碳素钢		443	667	25.3	3.79*	0.21%~0.35%C 1.28%~1.71%Mn
2	锰钢		418	614*	45.8	7.46*	1.19%~1.37%C 1.5%~12.0%Mn 0.5%~2.1%Si
3	钼钢（正火）		830	935	18.75	2.00*	0.37%~0.39%C 0.76%~0.85%Mn 1.95%~2.07%Ni 0.79%~0.84%Cr 0.24%~0.33%Mo
4	钼钢		1 392	1 729	169.7	9.81*	0.225%~0.43%C 0.61%~1.23%Mn 1.05%~2.48%Ni 0.46%~0.99%Cr 0.23%~0.79%Mo
5	低合金钢（右为回火温度）	370℃	1 276	1 406*	53.9	3.83*	0.23%~0.34%C 0.37%~0.87%Mn 0.46%~0.65%Cr 0.15%~0.24%Mo 0.41%~0.65%Ni
		454℃	1 153	1 216*	42.2	3.47*	
		538℃	1 023	1 076*	42.2	3.92*	
		620℃	907	995*	50.4	5.06*	
6	铬钼钒钢		1 444	1 749	84.9	4.86*	0.21%~0.34%C 0.37%~0.78%Mn 4.62%~5.97%Cr 1.20%~1.77%Mo 0.29%~0.59%V
7	高强度合金钢		1 691	1 805	99.9	5.54*	0.003%~0.05%C 15.0%~17.8%Ni 4.28%~5.10%Mo 6.40%~14.10%Co 0.13%~0.83%Ti 0.01%~0.62%Nb
8	灰铸铁	30 级	—	230	30.5	13.26	3.19%~3.65%C 1.63%~2.79%Si 0.44%~0.91%Mn
		40 级	—	287	39.7	13.84	3.02%~3.65%C 1.51%~2.49%Si 0.31%~0.99%Mn
		30 级与 40 级合计	—	254	44.7	17.59	3.02%~3.65%C 1.51%~2.49%Si 0.31%~0.99%Mn
		高中低合计	—	173	16.7	9.46*	3.47%~3.69%C 2.17%~2.72%Si 0.61%~0.84%Mn
9	球墨铸铁	L4512, M4512, 04512 合计	383	565	66.5	11.77	3.24%~3.95%C 1.45%~3.45%Si 0.17%~0.67%Mn
		04125	398	601	67.7	11.26	3.37%~3.84%C 2.32%~3.15%Si 0.24%~0.81%Mn
		L5506, M5506 合计	536.2	751.2	59.9	7.97	3.10%~3.98%C 2.49%~3.74%Si 0.21%~0.96%Mn 0.32%~0.76%Cu

续表 5.3-15

序号	材料名称		强度极限均值 σ_b /MPa	屈服极限均值 σ_s /MPa	强度极限的标准差 S_{σ_b}		附注
					MPa	强度极限的百分数/%	
9	球墨铸铁	M6003，M7003 合计	531.1	847.2	80.5	9.50	3.36%～3.79%C 2.50%～3.12%Si 0.44%～0.85%Mn 0.29%～0.50%Mo
		65－45－15 系列	395.3	545.2	39.3	7.22	北美洛克维尔公司数据
		80－55－06 系列	498.5	745.0	117.0	15.7	
10	铜合金	锰青铜（铸件）	—	567.5	15.7	2.76*	0.28%～1.37%Sn 27.41%～41.20%Zn 0.44%～2.76%Fe 0.50%～3.45%Al 0.23%～2.45%Mn
		铝青铜 A	329.4	771.6	33.3	4.31*	9.37%～11.80%Al 2.74%～6.05%Fe 3.07%～6.26%Ni 0.31%～0.63%Mn
		铝青铜 B	488.4	856.0	41.1	4.80*	9.37%～11.80%Al 2.74%～5.69%Fe 3.09%～6.26%Ni 0.49%～0.63%Mn
11	钛合金	无缺口	—	1 141*	73.4	6.14*	0.017%～0.100%C 0.0055%～0.011%H 0.10%～0.27%O 0.014%～0.079%N 0.12%～0.42%Fe 5.60%～6.20%Al 3.00%～13.50%V
		有缺口	—	1 396*	107.9	7.73*	

注：同表 5.3-14 注 1 和 2。

表 5.3-16　材料延伸率 δ 的标准差 S_δ

序号	材料名称		强度极限均值 σ_b /MPa	屈服极限均值 σ_s /MPa	延伸率均值 $\bar{\delta}$/%	延伸率的标准差 S_δ		附注
						%	延伸率的百分数/%	
1	碳素钢		443	667	25.41	2.92	11.48*	0.21%～0.35%C 1.28%～1.71%Mn
2	锰钢		418	614	14.25*	3.92	27.49*	1.19%～1.37%C 1.5%～12.0%Mn 0.5%～2.1%Si
3	钼钢		1 392	1 729	6.95	1.68	24.22*	0.225%～0.430%C 0.61%～1.23%Mn 1.05%～2.48%Ni 0.46%～0.99%Cr 0.23%～0.79%Mo
4	低合金钢（回火温度如下）	370℃	1 276	1 406	13.75*	4.25	30.91*	0.23%～0.34%C 0.37%～0.87%Mn 0.46%～0.65%Cr 0.15%～0.24%Mo 0.41%～0.65%Ni
		454℃	1 153	1 215	7.5*	2.17	28.89*	
		538℃	1 023	1 075	7.75*	2.08	26.88*	
		620℃	907	995	12.5*	2.50	20.00*	
5	铬钼钒钢		1 444	1 749	9.33	2.27	24.29*	0.21%～0.34%C 0.37%～0.78%Mn 4.62%～5.97%Cr 1.20%～1.77%Mo 0.29%～0.59%V

续表 5.3-16

序号	材料名称		强度极限均值 σ_b /MPa	屈服极限均值 σ_s /MPa	延伸率均值 $\bar{\delta}$/%	延伸率的标准差 S_δ		附注
						%	延伸率的百分数/%	
6	高强度合金钢		1 691	1 805	5.592	1.75	31.29*	0.003% ~ 0.05% C 15.0% ~ 17.8% Ni 4.28% ~ 5.10% Mo 6.40% ~ 14.10% Co 0.13% ~ 0.83% Ti 0.01% ~ 0.62% Nb
7	球墨铸铁	L4512，M4512，04512 合计	383	565	16.31	3.99	24.46	3.24% ~ 3.95% C 1.45% ~ 3.45% Si 0.17% ~ 0.67% Mn
		04128	398	601	13.38	3.86	28.85	3.37% ~ 3.84% C 2.32% ~ 3.15% Si 0.24% ~ 0.81% Mn
		L5506 M5506 合计	536.2	751.2	5.86	2.195	37.46	3.10% ~ 3.98% C 2.49% ~ 3.74% Si 0.21% ~ 0.95% Mn 0.32% ~ 0.76% Cu
		M6003 M7003 合计	531.1	847.2	6.92	2.18	31.50	3.36% ~ 3.79% C 2.50% ~ 3.12% Si 0.44% ~ 0.55% Mn 0.29% ~ 0.50% Mo
		65 - 45 - 12 系列	395.2	545.2	15.63	3.96	25.34	北美洛克维尔公司数据
		80 - 45 - 06 系列	498.5	745.0	10.32	3.80	36.82	
8	铜合金	锰青铜（铸件）	—	567.5	30.00	5.5	18.33*	0.28% ~ 1.37% Sn 27.41% ~ 41.20% Zn 0.44% ~ 2.76% Fe 0.50% ~ 3.45% Al 0.23% ~ 2.45% Mn
		铝青铜 A	329.4	771.6	12.36	3.5	28.32*	9.37% ~ 11.80% Al 2.74% ~ 6.05% Fe 3.07% ~ 6.26% Ni 0.31% ~ 0.63% Mn
		铝青铜 B	488.4	856.0	10.40	3.33	32.05*	9.37% ~ 11.80% Al 2.74% ~ 5.69% Fe 3.09% ~ 6.26% Ni 0.49% ~ 0.63% Mn
9	钛合金		—	1 141	9.00*	2.00	22.22*	0.017% ~ 0.100% C 0.0055% ~ 0.011% H 0.10% ~ 0.27% O 0.014% ~ 0.079% N 0.12% ~ 0.42% Fe 5.60% ~ 6.20% Al 3.00% ~ 13.50% V

注：同表 5.3-14 注 1 和 2。

表 5.3-17 材料硬度的标准差 S_H

序号	材料名称	强度极限均值 σ_b /MPa	屈服极限均值 σ_s /MPa	硬度名称	硬度均值	硬度的标准差 S_H		附注
						硬度	硬度的百分数/%	
1	锰钢	418	614	HB	221*	11.33	5.13*	1.19% ~ 1.37% C 1.5% ~ 12.0% Mn 0.5% ~ 2.1% Si
				RC	46.5*	0.833	1.79*	

续表 5.3-17

序号	材料名称		强度极限均值 σ_b /MPa	屈服极限均值 σ_s /MPa	硬度名称	硬度均值	硬度的标准差 S_H		附注
							硬度	硬度的百分数/%	
2	低合金钢（右为回火温度）	370℃	1 276	1 406	RC	37.0*	1.333	3.60*	0.23%~0.34%C 0.37%~0.87%Mn 0.46%~0.65%Cr 0.15%~0.24%Mo 0.41%~0.65%Ni
		454℃	1 153	1 216		29.5*	1.833	6.21*	
		538℃	1 023	1 076		26.5*	1.5	5.66*	
		620℃	907	995	RB	62.0*	1.0	1.61*	
3	高强度合金钢		1 691	1 805	RC	49.99	2.0	4.00*	0.003%~0.05%C 15.0%~17.8%Ni 4.28%~5.10%Mo 6.40%~14.10%Co 0.13%~0.83%Ti 0.01%~0.62%Nb
4	灰铸铁		—	173	HB	186.8	11.17	5.98*	3.47%~3.69%C 2.17%~2.72%Si 0.61%~0.84%Mn
5	球墨铸铁	65-45-15 系列	395.3	545.2	HB	166.85	12.29	7.37	北美洛克维尔公司数据
		80-55-06 系列	498.5	745.0		216.82	26.07	12.02	
6	铝青铜	A	329.4	771.6	HB	178.2	15.33	8.60*	9.37%~11.80%Al 2.74%~6.05%Fe 3.07%~6.26%Ni 0.31%~0.63%Mn
		B	488.4	856.0		217.6	15.33	7.05*	9.37%~11.80%Al 2.74%~5.69%Fe 3.07%~6.26%Ni 0.49%~0.63%Mn

注：同表 5.3-14 注 1 和 2。

4.3 疲劳寿命可靠性的失效预测技术和方法

4.3.1 疲劳寿命预测的重要性

承受交变载荷的机械构件和建筑结构，其失效形式往往是疲劳失效。工程设计的主要任务就是要预测在交变载荷作用下，机械构件和建筑结构的疲劳寿命。

所谓预测就是一种预报（预估或估计）方法，它是由实验室（或现场）收集的数据（包括载荷、环境条件、强度试验、寿命试验和可靠性试验）来预测构件（或系统）在现场实际使用的性能（包括强度、寿命和可靠度）的一种方法。

疲劳寿命预测的重要性可在下列几个方面得到体现。

（1）工程预测方面

在疲劳的工程设计中，寿命预测是一项主要内容。这一点见图 5.3-63 所示的设计活动和信息流程。从图中可以知道，在预测产品寿命的整个研究过程中工程设计和试验分析可以按照三种不同的路线进行。

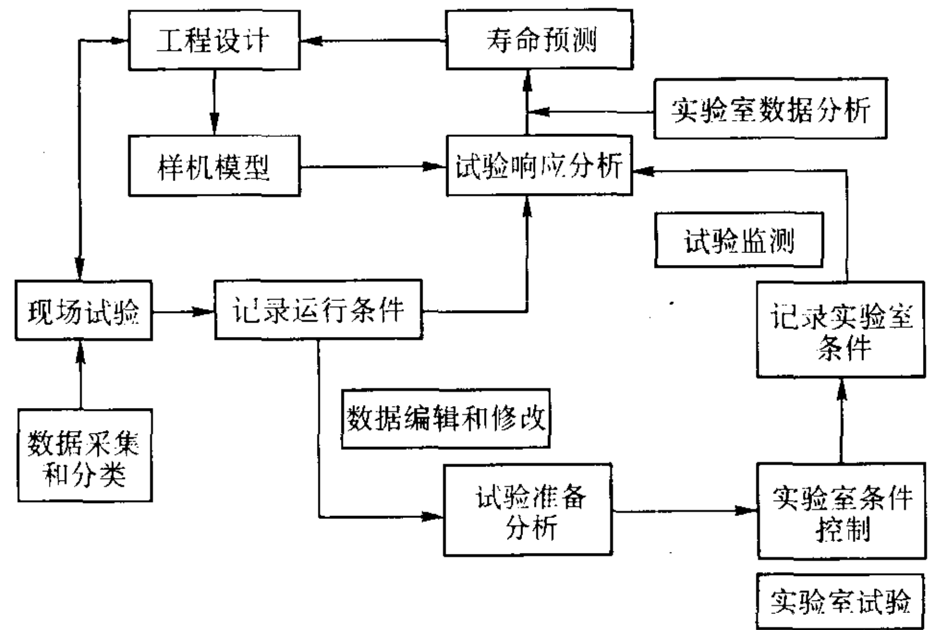


图 5.3-63 疲劳的工程预测和试验分析流程

1）样机模型的试验预测 第一条路线如图 5.3-64①所示，称为样机模型的试验预测。新产品的样机由于种种原因（或时间太长，或成本太高，或规模太大等）一时不能得到，则可使用样机模型来进行产品的设计和试验预测。采用样机模型的方法可以确定构件的应变集中区、构件的寿命和 $A-B$ 方案的比较，由此可以很快地对材料或结构型式做出改进的评价，以供不同设计方案的论证。执行这条设计路线需要有一个大容量的载荷数据库。

2）样机现场的试验预测 第二条设计路线如图 5.3-64②所示，称为样机现场的试验预测。这就是在现场试验一台样机并测量其实际响应来求得其寿命。样机现场试验，产品受到的载荷比较真实，环境条件逼真，因此其记录的运行历

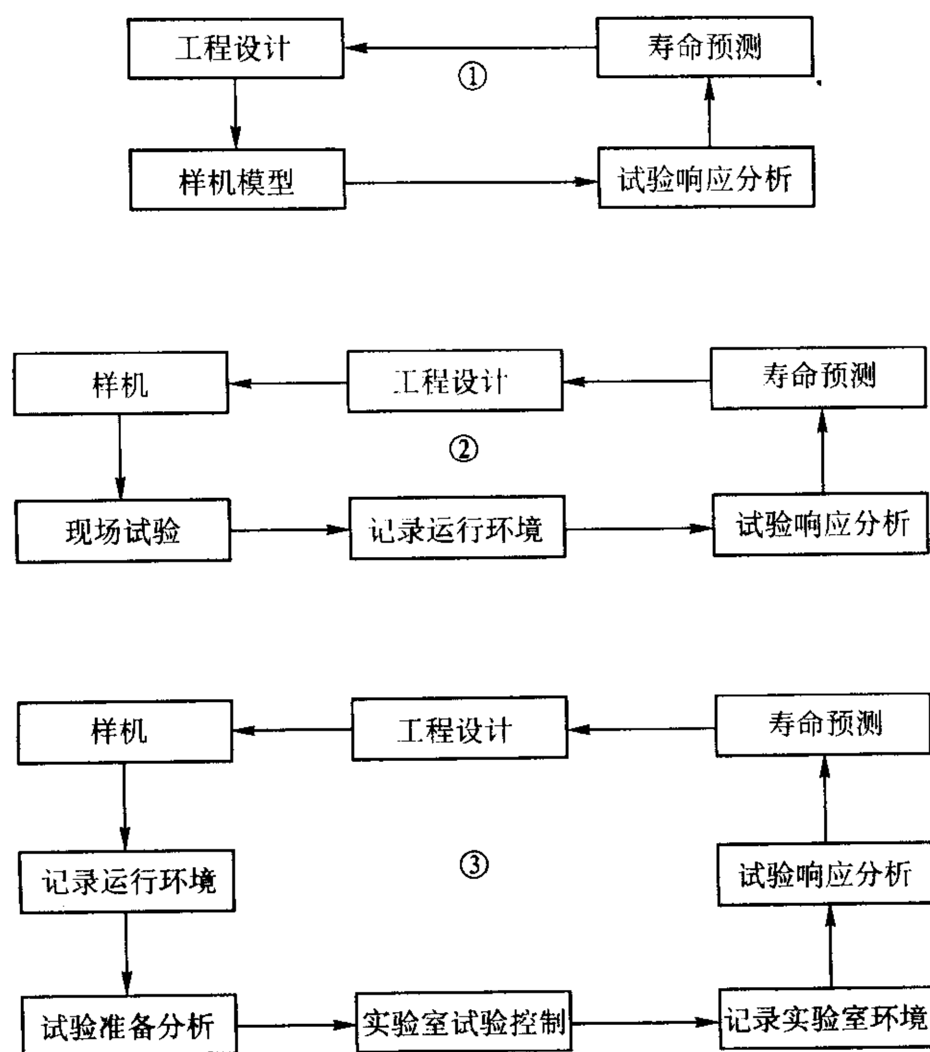


图 5.3-64　三条预测路线

程和求得的产品寿命与实际比较一致，是一种很好的寿命预测方法。特别是在载荷数据的数据库不足的情况下，因而不能进行第一条路线的设计时，应用第二条设计路线比较好。必须指出，样机现场试验的不足之处是载荷的重复性差，且不能进行加速试验（或称强化试验）。

第三条设计路线如图 5.3-64③所示，称为样机的实验室预测。在测得了典型的现场运行历程后，便可在实验室内对样机进行试验，以求得产品的寿命。这种设计路线的特点是载荷历程和环境条件均能重复，而且还能进行加速试验以缩短试验周期。它是一种较理想和完整的工程设计和寿命预测方法。

从图 5.3-63 还可看出，整个确定产品寿命的流程中要用到五种典型的工具（方法），分别是：①数据采集和分类，这主要是采集载荷数据，用它来进行疲劳试验和寿命估计；②数据编辑和修改，它把采集到的载荷数据中的干扰部分取消，以便用来加速试验；③实验室试验，它可以产生构件承受的典型载荷条件，且能测得构件的响应；④试验数据监测，其任务是确定在某一段时间内样品对载荷条件的反应，以判断试验的进行是否正常；⑤实验室数据预测，这一工具主要用来进行数据处理，即对现场或实验室试验的结果进行预测，以确定产品的寿命。

所有上述三条设计路线和五种典型工具都是围绕如何进行疲劳寿命预测。产品或构件的寿命要求在确定之后，就要从设计上满足，而方案能否达到规定的寿命指标，则必须进行寿命预测。

其次，构件和系统的设计，还要考虑材料、强度、重量、成本、加工方法等。对于受交变载荷下的构件或系统，重要的是要使其寿命和可靠度达到规定的指标，这就必须进行寿命预测和可靠性预测。

再者，在对比各种试验方案时，需要了解一个系统中各个子系统之间、一个子系统中各个元件（零件）之间关联的情况和影响，以便加以改进，要达到这一目的，也必须进行寿命和可靠性的预测。

（2）制订检修计划

构件的有限寿命设计（又称全寿命设计），要求构件存在一段安全工作的时间，这段时间又可称为无裂纹寿命。对于构件无裂纹寿命的预测很重要，一则可以安排检修周期，二则可以在它寿命终了之前调换下来，以保证整机继续安全运行。

疲劳的损伤容限设计是一种较少保守的设计，它允许构件或系统具有一定缺陷或破损，即带有一定程度的损伤，只要这种损伤还未达到临界值，失效就不会发生，这种设计方法的基础就是寿命预测。要求在两次检修之间的运行不会发生灾难性事故，这就要求有准确的寿命估计。因此，对于改进后的产品或新研制的产品，如何准确地预测其寿命和可靠度，不仅是安排检修计划的依据，也是保证安全运行的必要条件。

（3）制订验收标准

一般对于承受交变载荷的构件在进行验收时，有的要进行无损探伤，有的要抽样进行疲劳试验。进行无损探伤时，要根据缺陷的大小和分布，估计构件的寿命值，检测其是否达到了设计的要求，达到寿命要求的就合格，反之就不合格。因此允许缺陷的最大尺寸就要根据给定的寿命值来定，这就是说，寿命预测是制订构件缺陷验收标准的依据。

对于抽样疲劳试验，要根据该试验结果来推断产品的寿命，检验其是否符合设计的要求，从而判断合格与否。必须综合考虑产品的预测寿命、抽样检验后的推断寿命和安全系数三者之间的关系，结合产品的重要性来制订验收标准。

4.3.2　等幅交变载荷下疲劳强度和寿命的可靠性预测

（1）疲劳寿命服从对数正态分布的情况

设：$x=\lg N$，$\mu=E(\lg N)$，$\sigma=\sqrt{V\lg N}$，μ，σ 分别为对数寿命 x 的数学期望和标准差。

则其破坏概率为

$$P_f(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx \quad (5.3\text{-}400)$$

引进标准变量

$$t=\frac{x-\mu}{\sigma},\ dt=\frac{1}{\sigma}dx$$

代入上式得

$$P_f(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\exp\left(-\frac{t^2}{2}\right)dt \quad (5.3\text{-}401)$$

按照定义，失效概率为

$$P_f(t<t_f)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t_f}\exp\left(-\frac{t^2}{2}\right)dt \quad (5.3\text{-}402)$$

式中，$-t_f=\dfrac{x_f-\mu}{\sigma}$

所以

$$x_f=\mu-t_f\sigma \quad (5.3\text{-}403)$$

称上式中的 x_f 为与失效概率 P_f 相联系的安全寿命，注意公式中 t_f 本身取正值。

按定义，可靠度为

$$R(t>t_R)=\frac{1}{\sqrt{2\pi}}\int_{t_R}^{\infty}\exp\left(-\frac{t^2}{2}\right)dt \quad (5.3\text{-}404)$$

$$t_R=\frac{x_R+\mu}{\sigma}$$

$$x_R=\mu+t_R\sigma \quad (5.3\text{-}405)$$

称上式中的 x_R 为与可靠度 R 相联系的安全寿命，注意公式中的 t_R 本身取自然值。

（2）疲劳寿命服从威布尔分布的情况

正态分布存在一个明显的缺点，即无论把应力取得如何低，总存在一个失效概率（尽管它的值很小），这等于说不存在失效概率为零的应力值，显然这与疲劳的物理背景不相符，因为材料或构件事实上存在疲劳极限，当所加的交变应力低于这个极限时，就不会发生疲劳失效。三参数的威布尔

分布就能很好地描述这一特性。

三参数威布尔分布的概率密度式如下：

$$f(N)=\frac{b}{T-N_0}\left(\frac{N-N_0}{T-N_0}\right)^{b-1}e^{-\left(\frac{N-N_0}{T-N_0}\right)^b} \quad (5.3\text{-}406)$$

式中，N_0 为位置参数，在寿命分布中称最小寿命；T 为尺度参数，在寿命分布中称为特征寿命；b 为形状参数，又称为威布尔斜率。

累积失效斜率为

$$F(N)=1-\exp\left[-\left(\frac{N-N_0}{T-N_0}\right)^b\right] \quad (5.3\text{-}407)$$

可靠度为

$$R(N)=\begin{cases}\exp\left[-\left(\frac{N-N_0}{T-N_0}\right)^b\right] & 当\ N\geqslant N_0\\ 1,0 & 当\ N<N_0\end{cases} \quad (5.3\text{-}408)$$

在威布尔分布的应用研究中，关键是要解决形状参数 b 的估计、可用作图法、最佳线性无偏估计（BLUE）、最佳线性不变估计（BLIE）、最小二乘法、最大似然法等。但从工程应用的角度有一种简便的方法，它找出了形状参数 b 与寿命的变差系数

$$c_N=\frac{\sigma_N}{\overline{N}}$$

之间的关系，即由下面两式

$$\overline{N}'=\overline{N}-N_0=(T-N_0)\Gamma\left(1+\frac{1}{b}\right) \quad (5.3\text{-}409)$$

$$\sigma_N=(T-N_0)\left[\Gamma\left(1+\frac{2}{b}\right)-\Gamma^2\left(1+\frac{1}{b}\right)\right]^{\frac{1}{2}} \quad (5.3\text{-}410)$$

式（5.3-410）除以式（5.3-409）得

$$c'_N=\frac{\sigma_N}{\overline{N}'_1}=\frac{\left[\Gamma\left(1+\frac{2}{b}\right)-\Gamma^2\left(1+\frac{1}{b}\right)\right]^{\frac{1}{2}}}{\Gamma\left(1+\frac{1}{b}\right)} \quad (5.3\text{-}411)$$

当 N_0 远小于 N 时

$$c'_N=\frac{c_N}{\overline{N}'}\approx\frac{\sigma_N}{\overline{N}}=c_N \quad (5.3\text{-}412)$$

所以 b 与 c'_N 或 c_N 存在式（5.3-411）的关系。这个关系可近似用下式表示。

$$b=c_N^{-1.08}\ 或\ c_N=b^{-0.926} \quad (5.3\text{-}413)$$

则式（5.3-408）的可靠度 R 可表示为

$$R(N)=\exp\left[-\frac{N'}{\overline{N}}\Gamma(1+c_N^{1.08})\right]^{c_N^{-1.08}} \quad (5.3\text{-}414)$$

式中，$N'=N-N_0$；$\overline{N}'=\overline{N}-N_0$。

式（5.3-414）可以解决两类问题。

1）已知等幅交变应力作用下的寿命变差系数 c_N（一般等幅应力下的疲劳寿命 $c_N=0.3\sim0.7$），并已知材料（或构件）的疲劳 $P-S-N$ 曲线，求可靠性寿命 N_R 及寿命的可靠度 $R(N)$。

利用式（5.3-407）和式（5.3-413），可以得到

$$N_R=(T-N_0)\left[\ln\frac{1}{1-F(N)}\right]^{c_N^{1.08}}+N_0$$

式中可靠度 $R=1-F(N)$，c_N 为已知，只要求得到位置参数 N_0 及特征参数 T，就可以从上式求得可靠寿命 N_R。

寿命 N 的可靠度可由式（5.3-408）或式（5.3-414）（当 N_0 远小于 N 时）之一求得。关于 N_0 的确定有许多方法，如作图法、最小二乘法、矩法和最大似然法等。

2）已知可靠性寿命 N_R 和可靠度 R，求平均寿命（或设计寿命）。

设可靠寿命为 N_R，可靠度为 $R(N_R)$，$\overline{N}_d$ 为设计的平均寿命，令

$$\gamma_L=\frac{\overline{N}'}{N'}=\frac{\overline{N}'_d}{N'_R}$$

为寿命系数。

式中，$N'=N-N_0$，$\overline{N}'=\overline{N}-N_0$。

由式（5.3-414）考虑寿命系数 γ_L 得

$$\gamma_L=\frac{\Gamma(1+c_N^{1.08})}{\left[\ln\frac{1}{R(N_R)}\right]^{c_N^{1.08}}}\approx\frac{\Gamma(1+c_N^{1.08})}{[F(N_R)]^{c_N^{1.08}}} \quad (5.3\text{-}415)$$

由上式求得 γ_L 后即可求得平均寿命（或设计寿命）：

$$\overline{N}_d=\gamma_L N_R \quad (5.3\text{-}416)$$

4.3.3 多级载荷下疲劳强度和寿命的可靠性预测

（1）载荷（应力）累积频次分布图

上一节讨论的是等幅交变载荷作用下的疲劳寿命预测问题，实际工程中构件所受到的载荷，其频率和幅值都随时间在变化，这一节要讨论的多级载荷下的疲劳问题，既是前一节问题的发展，又是随机载荷下疲劳问题的近似处理的基础。

对于随机载荷，可以通过适当的计数方法将它在整个记录周期内所出现的峰值载荷频次加以确定，画出载荷（应力）累积频次分布图，如图5.3-65所示。图5.3-65曲线1是代表等于或大于给定载荷出现次数。有时，为了得到构件在其总寿命中可能出现的、更为真实的载荷－时间历程，可以把这条自实际采集样本所得到的分布曲线，扩展到 10^6 次循环，即相当于疲劳极限寿命时的分布曲线，见图5.3-65曲线2。有了这种扩展的现场载荷频次分布图，就可以把它分成若干级（一般分为8级），成为如图5.3-66所示的阶梯载荷直方图，这在疲劳试验中和寿命估计中均有重要的应用。

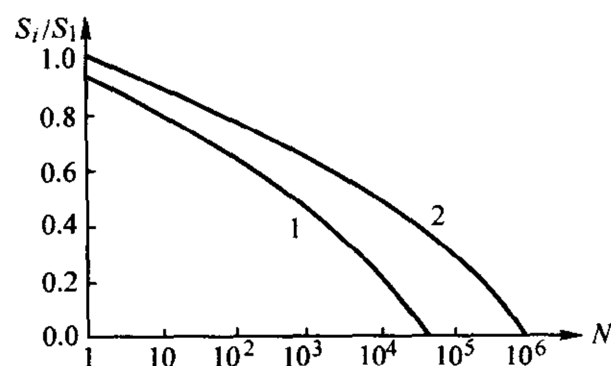


图5.3-65 载荷累积频次分布图

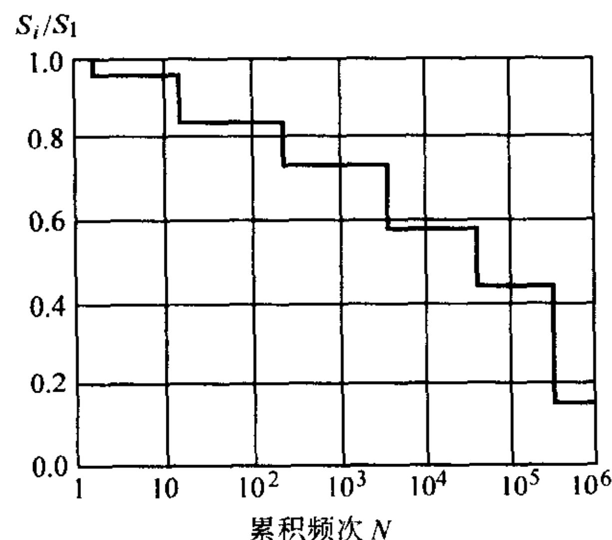


图5.3-66 阶梯载荷直方图

（2）疲劳累积损伤理论

1）迈纳（Miner）线性累积损伤理论　这个理论假定，在试样受载过程中，每一载荷循环都消耗一定的有效寿命分

数；又假定疲劳损伤是与试样中所吸收的功成正比的，而且又认为这个功与应力的作用循环次数和在该应力值下达到破坏的循环次数之比成比例，此外还假定试样达到破坏时的总损伤量（总功）是一个常量，它是载荷的简单函数，并且损伤与载荷作用的次序无关。最后，假定各循环应力产生的所有损伤分量之和等于 1 时，试样发生破坏。因此，归纳起来可得出如下的基本关系式：

$$d_1 + d_2 + \cdots + d_k = \sum_{i=1}^{k} d_i = D$$

$$\frac{d_i}{D} = \frac{n_i}{N_i} \text{或} d_i = \frac{n_i}{N_i} D \qquad (5.3\text{-}417)$$

式中，D 为总损伤量；d_i 为损伤分量；n_i 为为试样在应力级 S_i 作用下的工作循环次数；N_i 为材料在应力级 S_i 下直至破坏的循环数。

$$\frac{n_1}{N_1} D + \frac{n_2}{N_2} D + \cdots \frac{n_k}{N_k} D = D$$

即

$$\sum_{i=1}^{k} \frac{n_i}{N_i} = 1 \qquad (5.3\text{-}418)$$

根据大量试验数据的统计，实际达到破坏时的总累积损伤量 D 值约为 0.61 ~ 1.45；而在交变载荷下，D 值不但与载荷幅值而且与作用次序有关，且差别更大。此外，按迈纳理论，累积损伤只产生在应力大于材料的疲劳极限的范围，小于材料的疲劳极限，则不引起损伤且具有无限寿命。

虽然迈纳理论有一定的局限性，但目前仍被广泛应用，主要原因是因为其十分简便，而且 D 作为一个随机变量其数学期望为 1.0，因此用它来估计寿命是一个比较好的手段。

设 N_L 为在变幅载荷下构件的疲劳寿命，令

$\alpha_i = \dfrac{n_i}{\sum_{i=1}^{k} n_i}$ 为第 i 级应力 S_i 作用下的工作循环数与各级应力下总的循环数之比，则

$$n_1 = \alpha_1 N_L; n_2 = \alpha_2 N_L, \cdots, n_k = \alpha_k N_L$$

代入式（5.3-418）可得

$$N_L \sum_{i=1}^{k} \frac{\alpha_i}{N_i} = 1 \qquad (5.3\text{-}419)$$

又设 N_1 为最大应力级 S_1 作用下材料到破坏循环时的循环数，则根据材料疲劳的 $S-N$ 曲线，有

$$\frac{N_1}{N_i} = \left(\frac{S_i}{S_1}\right)^m$$

代入式（5.3-419），得到估计疲劳寿命的计算式：

$$N_L = \frac{N_1}{\sum_{i=1}^{k} \alpha_i \left(\frac{N_1}{N_i}\right)} = \frac{N_1}{\sum_{i=1}^{k} \alpha_i \left(\frac{S_i}{S_1}\right)^m} \qquad (5.3\text{-}420)$$

2）修正的线性累积损伤理论　由于迈纳理论没有考虑到应力级之间的相互影响和低于疲劳极限一下应力的损伤作用，很多研究者做出了修正。其中以 Corten - Dolan 理论应用较多，这个理论是以最大交变应力下所产生的损伤核数目与疲劳裂纹的扩展速率为依据的，从而推导得出多级载荷作用下疲劳寿命估计的一般公式：

$$N_L = \frac{N_1}{\sum_{i=1}^{k} \alpha_i \left(\frac{S_i}{S_1}\right)^d} \qquad (5.3\text{-}421)$$

可见式（5.3-421）与式（5.3-420）十分相似。因此，Corten - Dolan 理论可以认为是对应于另一种形式疲劳曲线的迈纳公式。一般 $d = (0.8 \sim 0.9)m$。

因此，对于低应力损伤分量的比重较大的场合，应用 Corten - Dolan 理论估计的疲劳寿命，将比用迈纳理论估计的疲劳寿命更短，而且更符合实际。

4.3.4　随机疲劳的寿命估计

（1）随机载荷下疲劳寿命估计的流程和方法

1）随机载荷下疲劳分析的流程分为以下一些环节。

① 信号采集　用电阻应变片通过动态应变仪，将信号放大后输入光电示波器或磁带机，记录构件的响应信号，这就是载荷 - 时间（或应变 - 时间）历程。

② 推断　如果上述载荷 - 时间历程取自于构件的危险点（应力集中区），则可直接加以利用。如果是名义的载荷 - 时间历程，则必须经过推断，将其换算得到局部应变 - 时间历程。

③ 中心化　任何交变载荷均可分解为两部分，一部分是静分量，其值等于平均载荷；另一部分是动分量，其值为某一点的值与平均值之差。

④ 数字处理　这可以从下述三个方面着手：a）幅值域——即从载荷、应力、位移等参量的幅值大小和分布方面来研究随机疲劳过程，这里常用的指标是平均值、均方差、标准差、概率密度函数和概率函数等；b）时间域——即从载荷、应力、位移等参量随时间的变化方面来研究随机疲劳过程，这方面常用的指标是自相关函数和互相关函数，前者描述同一过程的现在与将来之间的关系，后者描述两个不同过程时间前后之间的关系；c）频率域——载荷、应力和位移等的任一波形，均能由不同频率的周期函数叠加而成，确定性的函数是这样，随机函数也是这样。从频率角度来描述随机疲劳时可采用自功率谱密度函数，互谱密度函数和谱相关函数等指标。

⑤ 载荷的累积频次　根据构件的各种工况，分别计数后得到各种工况下的载荷累积频次图，最后把它们合成为一条累积频次曲线。

⑥ 扩展　在信号采集时，子样的时段是有限的，在这有限的时段内，不一定出现载荷的最大值，所以必须将上述载荷的累积频次图加以扩展，得到能代表全貌的扩展后的频次图，一般均扩展至 10^6 次。

⑦ 八级程序　根据载荷频次图，可以编成各种载荷程序，在疲劳研究中用得较普遍的是八级程序，这个八级程序是疲劳试验、名义应力法寿命估计和局部应力 - 应变法寿命估计的依据。

⑧ 程序疲劳试验　用上述八级程序进行构件的疲劳试验，以测定构件的疲劳寿命。

⑨ 名义应力法寿命估计　运用上述八级程序和构件的 $S-N$ 曲线，再选用一种累积损伤理论（一般均用 Miner 线性累积损伤理论）就可以估计构件的寿命。

⑩ 局部应力 - 应变法寿命估计　运用上述八级程序和构件的总应变 - 寿命曲线以及材料的循际应力 - 应变模型，再选用一种累积损伤理论，即可以进行寿命估计。

2）寿命估计方法　随机疲劳的寿命估计，目前有下列几种方法：

① 根据构件的载荷 - 寿命曲线（$P-N_f$ 曲线）计算寿命；

② 根据构件的名义应力 - 寿命曲线（$S-N$ 曲线）计算寿命；

③ 根据光滑试件的应变 - 寿命曲线计算寿命，即局部应力 - 应变的疲劳寿命计算方法；

④ 功率谱法估算寿命。

（2）相当应力方法

谢林生（C.B.Cepihceh）根据 Miner 的线性累积损伤理论，把原来是分级离散型的多项 Miner 计算公式改写成连续性积分式。

$$\int_0^K \frac{\mathrm{d}n_i}{N_i} = 1 \tag{5.3-422}$$

材料的 $\lg S - \lg N$ 疲劳曲线左段斜线，为过载作用下的有限寿命部分，且斜线上每一点（σ_i，N_i）满足下式：

$$\sigma_i^m N_i = \sigma_{-1}^m N_{-1}$$

$$\text{或 } N_i = \left(\frac{\sigma_{-1}}{\sigma_i}\right)^m N_{-1} \tag{5.3-423}$$

式中，$m = \cot\alpha$，为疲劳曲线左段斜线的斜率的倒数，一般钢材 m 为 6～10。

随机载荷作用下，构件的过载应力幅 σ_i 作用的频次数

$$n_i = n_{总} f(\sigma_i)$$

式中，$f(\sigma_i)$为应力 $\sigma \leqslant \sigma_i$ 的累积频率；$n_{总}$ 为疲劳寿命总频次数。

上式两边微分可得

$$\mathrm{d}n_i = n_{总} f'(\sigma_i)\mathrm{d}\sigma_i \tag{5.3-424}$$

式中，$f'(\sigma_i)$为 σ_i 的概率密度函数。

将式（5.3-423）、式（5.3-424）代入式（5.3-422），并整理得

$$\frac{n_{总}}{N_{-1}} = \int_{\sigma_{-1}}^{\sigma_{\max}} \sigma_i^m f'(\sigma_i)\mathrm{d}\sigma_i = \sigma_{-1}^m \tag{5.3-425}$$

构件不产生疲劳破坏的条件为

$$\int_0^K \frac{\mathrm{d}n_i}{N_i} \leqslant 1$$

即上式成为

$$\sqrt[m]{\frac{n_{总}}{N_{-1}} = \int_{\sigma_{-1}}^{\sigma_{\max}} \sigma_i^m f'(\sigma_i)\mathrm{d}\sigma_i} \leqslant \sigma_{-1}$$

不等式左边形式上可看作是一应力幅，称为连续分布应力$f(\sigma_i)$的相当应力幅，记为 σ_e。对于应力比不等于 -1 而等于 r 的非对称循环，可将上式改写为

$$(\sigma_r)_e = \sqrt[m]{\frac{n_{总}}{N_r} = \int_{\sigma_r}^{\sigma_{\max}} (\sigma_r)_i^m f'(\sigma_{ri})\mathrm{d}\sigma_{ri}} \leqslant \sigma_r \tag{5.3-426}$$

上式就是表达构件在随机载荷作用下的疲劳强度条件的谢林生公式。式中，σ_r 为构件在某一相当循环特征 r 下的疲劳极限。$(\sigma_r)_e$ 是随机载荷引起的构件名义应力的相当值。

谢林生通过式（5.3-426）企图把随机变幅载荷下的构件的疲劳寿命与等幅载荷作用下的疲劳寿命联系起来，把随机变幅应力处理成为某一相当循环特征 r 下的等幅应力 $(\sigma_r)_e$ 作用，使问题得到了简化。

（3）局部应力 - 应变法

传统的名义应力法的思想是以构件的名义应力为计算依据，它把缺口处的局部应力看作是名义应力与理论应力集中系数的乘积。然而，实际情况却往往是在名义应力很小时，局部应力已超过了材料的屈服点而产生了塑性变形。如果产生塑性变形的区域较大，则名义应力法显然不适合。为了解决这一问题，国外发展了局部应变法（后来被成为应力 - 应变法），该方法运用弹塑性分析方法，结合构件受载历史和材料在循环载荷下的应力 - 应变关系，求得构件缺口处的局部应力 - 应变响应，然后应用 Miner 累积损伤理论计算构件的裂纹萌生寿命。使用雨流法并与计算技术相结合，局部应力 - 应变法也可以计算出随机载荷对构件造成的损伤，从而估计构件的寿命。

1）循环应力 - 应变曲线及其数学描述　在循环应力的作用下，材料会产生硬化或软化现象。随着循环应变的大小和材料状态的不同，金属材料会出现：循环硬化、循环软化、循环稳定和循环硬化和软化兼有的混合型。

通过控制应变的疲劳试验，将 $\Delta\varepsilon$ 取不同的值，得到幅度大小不同的滞后回线或滞后环，在同一个坐标系中把各个滞后环的顶点连接起来，得到的曲线叫做循环应力 - 应变曲线。

金属材料的应力 - 应变曲线为

$$\frac{\Delta\varepsilon}{2} = \frac{\Delta\sigma}{2E} + \left(\frac{\Delta\sigma}{2K'}\right)^{1/n'} \tag{5.3-427}$$

式中，$\Delta\varepsilon$ 和 $\Delta\sigma$ 称为应变和应力的变程；E 为材料的弹性模量；K'为循环硬化系数；n'为循环硬化指数，一般金属的 n'为 0.1～0.2。

2）应力 - 应变滞后回线的数学描述　在同一坐标系统中，将各个滞后环作平移，使其下尖点都与原点重合，再连接各个上尖点，得到一条曲线。这条曲线的方程基本上是循环应力 - 应变曲线扩大二倍得到的，即上行段的方程可表示为

$$\frac{\varepsilon - \varepsilon_i}{2} = \frac{\sigma - \sigma_i}{2E} + \left(\frac{\sigma - \sigma_i}{2K'}\right)^{1/n'} \tag{5.3-428}$$

式中，（σ_i，ε_i）为各个环的下尖点坐标。

用同样的方法，将上尖点平移至原点，连接各下尖点的曲线方程为

$$\frac{\varepsilon_i - \varepsilon}{2} = \frac{\sigma_i - \sigma}{2E} + \left(\frac{\sigma_i - \sigma}{2K'}\right)^{1/n'} \tag{5.3-429}$$

式中，（σ_i，ε_i）为各个环的上尖点坐标。

所有的滞后环曲线，包括上行段和下行段，均可以用统一的方程描述，其形式如下

$$\frac{\varepsilon}{2} = \frac{\sigma}{2E} + \left(\frac{\sigma}{2K'}\right)^{1/n'} \tag{5.3-430}$$

3）应变 - 寿命曲线　在低周疲劳中，应变 - 寿命曲线的形式一般为

$$\frac{\Delta\varepsilon}{2} = \frac{\sigma'_f - \sigma_m}{E}(2N_f)^b + \varepsilon'_f(2N_f)^c \tag{5.3-431}$$

式中，σ'_f 为疲劳强度系数；b 为疲劳强度指数；ε'_f 为疲劳延性系数；c 为疲劳延性指数。

4）名义应力与局部应力的关系　在构件的应力集中区域，实际应力（应变）与名义应力（应变）并非线性关系，计算损伤时则必须使用局部区域的实际应力（应变）进行。

Neuber 推导得到反平面剪切情况下，在弹塑性范围都适用式（5.3-432）。

$$K_t = \sqrt{K_\sigma K_\varepsilon} \tag{5.3-432}$$

式中，K_t 为应力集中系数（弹性范围内）；K_σ 为实际应力集中系数，$K_\sigma = \sigma/s$，其中 σ 为实际产生的弹塑性应力，又称为局部应力；s 为名义应力；K_ε 为应变集中系数，$K_\varepsilon = \varepsilon/e$，其中 ε 为实际产生的弹塑性应变，又称局部应变；e 为名义应变，$e = s/E$。

对于单向拉伸形式的加载，上式近似成立，又可改写为：

$$\sigma\varepsilon = K_t^2 se = K_t^2 \frac{s^2}{E} \tag{5.3-433}$$

如果知道材料的变形规律：

$$\sigma = f(\varepsilon) \tag{5.3-434}$$

则联立式（5.3-435）、式（5.3-436），则可解得局部的应力 σ 和应变 ε。

5）损伤计算　根据 Miner 线性累积损伤理论，构件的损伤：

$$D = \sum_{i=1}^{k} \frac{n_i}{N_i} \tag{5.3-435}$$

一般当 $D = 1$ 时产生疲劳失效。

对于随机载荷，因为对一段子样进行了采集和分析，所以其损伤 D 为

$$D = B\sum_{i=1}^{k}\frac{n_i}{N_i} \tag{5.3-436}$$

式中，B 为程序块历程块的重复次数。

6）局部应力－应变法　随着对材料疲劳认识的深化，局部的应力－应变法被提出。它的基本观点是：构件和结构的疲劳破坏都是在某些产生应力集中的局部危险区域首先发生的。材料在这些应力集中区域的力学响应都可以在切口试样和光滑试样上模拟复现。只要对标准试样进行仔细的力学性能研究，并准确掌握其在载荷－时间历程作用下的力学响应规律和切口因子影响规律，就可以通过计算分析确定由同种材料制造各种复杂构件上应力集中处的力学响应规律，从而不必对实际构件进行疲劳破坏试验，就能以较高的精确度计算得到它的疲劳寿命。

在运用局部应力－应变法估算构件寿命时，可按照两种方法进行：一种是给出载荷－应变标定曲线、这种载荷－应变标定曲线可由试验测得或用有限元计算得到。具体就是：由载荷－应变标定曲线，得到载荷－时间历程曲线，结合材料的循环应力－应变曲线，得到应力集中区域的局部应变－时间历程 $\varepsilon - t$ 和局部应力－时间历程，这两者结合可得到局部应力－应变的响应曲线；采用一定的循环计数方法（如雨流法）可从材料的应变－寿命曲线得到每种应变幅值循环下的损伤；利用线性累积损伤规律便可估算得到构件的疲劳寿命。

另一种是利用 Neuber 公式的近似计算法，它与前一种方法的不同之处就是利用了 Neuber 公式（5.3-433）来估算局部区域的应力－应变响应，这种方法简便而实用，对于各种不同几何形状的构件，只需要知道应力集中部位的理论应力集中系数 K_t 和疲劳有效应力集中系数 K_f 就可以进行寿命估算，从而不必对构件做复杂的弹塑性分析，这对于设计阶段估算疲劳寿命是很实用的。

（4）功率谱法

功率谱法的依据是随机过程理论，假定实测的应力－时间历程是真实应力过程的一个样本，由此得到随机过程种不同频率成分上的能量或随机应力的强度分布情况，而不直接涉及这个样本的随机幅值序列。当然，由功率谱密度函数也可导出幅值域的统计量，如均方值和零穿越期望值等。

假定我们讨论的是高斯平稳随机过程，$s(f)$为功率谱密度函数，则随机过程 $Y(t)$ 的方差为

$$\sigma^2 = \int_{-\infty}^{+\infty} s(f)\mathrm{d}f \tag{5.3-437}$$

$Y(t)$ 的零穿越期望值，即单位时间内以正斜率穿越零均值的平均次数为

$$f_0 = \sqrt{\frac{\int_{-\infty}^{+\infty} f^2 s(f)\mathrm{d}f}{\int_{-\infty}^{+\infty} s(f)\mathrm{d}f}} \tag{5.3-438}$$

$Y(t)$ 的峰值期望值，即单位时间内正峰值出现的平均次数为

$$n_0 = \sqrt{\frac{\int_{-\infty}^{+\infty} f^4 s(f)\mathrm{d}f}{\int_{-\infty}^{+\infty} f^3 (f)\mathrm{d}f}} \tag{5.3-439}$$

定义：

$\alpha = f_0/n_0$ 为不规则系数；$\varepsilon = \sqrt{1-\alpha^2}$ 为带宽系数。

它们分别代表了随机波形的不规则程度和频带宽窄程度。当 $\varepsilon = 1$ 时为白噪声，$\varepsilon = 0$ 时为窄带。工程上一般认为当 $\varepsilon \leqslant 0.35$ 时可作窄带处理。

平稳随机过程峰值的概率密度函数为

$$h(y) = \frac{\varepsilon}{\sqrt{2\pi}\sigma}\exp\left(-\frac{y^2}{2\sigma^2\varepsilon}\right) + \frac{y\alpha}{2\sigma^2}\left[1+\mathrm{erf}\left(\frac{y}{\sigma}\sqrt{\frac{\alpha^2}{2\varepsilon}}\right)\right] \times \exp\left(-\frac{y^2}{2\sigma^2}\right) \tag{5.3-440}$$

式中，$\mathrm{erf}(x)$ 为误差函数，定义为：

$$\mathrm{erf}(x) = \int_0^x \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\mathrm{d}t$$

在窄带情况下，$\varepsilon = 0$，$\alpha = 1$，则

$$p(y) = \frac{y}{\sigma^2}\exp\left(-\frac{y^2}{2\sigma^2}\right) \tag{5.3-441}$$

可见窄带情况下的峰值服从瑞利分布；

而在白噪声情况下，$\varepsilon = 1$，$\alpha = 0$，则

$$p(y) = \frac{y}{\sigma^2}\exp\left(-\frac{y^2}{2\sigma^2}\right) \tag{5.3-442}$$

峰值服从高斯分布。

工程中的宽带过程的峰值，实际上是高斯分布与瑞利分布的不同组合，不规则系数 α 就是两种分布所占比例的度量。功率谱方法估计疲劳寿命主要控制因素就是峰值的概率密度分布形式，根据这一点推导得到下列寿命估计公式。

利用迈纳线性累积损伤理论，设应力峰值为 s' 时的损伤为

$$D_{s'} = n_{s'}/N_{s'} \tag{5.3-443}$$

式中，$n_{s'}$ 为应力峰值是 s' 时的实际循环数；$N_{s'}$ 为应力峰值是 s' 时材料直至破坏的循环数。当

$$\int_0^{\infty} D_{s'}\mathrm{d}s' = 1 \tag{5.3-444}$$

成立时则破坏发生。用 N_f 表示构件至发生破坏时的总循环数。则令 $p(s') = n_{s'}/N_f$，可得 $D_{s'} = p(s')N_f/N_{s'}$。应用材料等幅载荷下的 $S-N$ 曲线：

$$N_s = CS^{-\beta} \tag{5.3-445}$$

式中，C 和 β 为材料（或构件）的常数。将上式和 $D_{s'}$ 的表达式代入式（5.3-444）可得

$$N_f = \frac{1}{\int_0^{\infty}\frac{p(s')\,s'^{\beta}\mathrm{d}s'}{C}} \tag{5.3-446}$$

对窄带过程，将式（5.3-441）代入可得

$$N_f = \frac{1}{\int_0^{\infty}\left(\frac{1}{\sigma^2 C}\right)s'^{1+\beta}\exp\left(-\frac{s'^2}{2\sigma^2}\right)\mathrm{d}s'} \tag{5.3-447}$$

令 $U = \frac{s'^2}{2\sigma^2}$，$V = \frac{2+\beta}{2}$，代入上式有

$$N_f = \frac{C}{(\sqrt{2}\sigma)^{\beta}\int_0^{\infty}U^{V-1}\mathrm{e}^{-u}\mathrm{d}u} = \frac{C}{(\sqrt{2}\sigma)^{\beta}\Gamma\left(\frac{2+\beta}{2}\right)} \tag{5.3-448}$$

也可写作

$$N_f = \left\{\left(\frac{1}{\sqrt{2}}\right)^{\beta}\frac{C}{\Gamma\left(\frac{2+\beta}{2}\right)}\right\}^{\sigma^{-\beta}} \tag{5.3-449}$$

当用时间表达寿命时，$N_f = f_0 N_T$，则

$$N_T = \frac{1}{f_0}\left\{\left(\frac{1}{\sqrt{2}}\right)^{\beta}\frac{C}{\Gamma\left(\frac{2+\beta}{2}\right)}\right\}^{\sigma^{-\beta}} \tag{5.3-450}$$

对于宽带过程，可在上式除以一个修正因子 λ，λ 与材料性能 β 和带宽系数 ε 有关，可用下述经验公式估计：

$$\hat{\lambda} - a = (1-a)(1-\varepsilon)^b$$

式中，a，b 均为 β 的函数，且为非负数。上式满足如下两种情况：$\varepsilon=0$ 时 $\hat{\lambda}=1$（窄带情况）；当 $\varepsilon=1$ 时 $\mathrm{d}\hat{\lambda}/\mathrm{d}\varepsilon=0$（白噪声）。可用最小二乘法求得 a 和 b 如下：

$$\begin{cases} a(\beta)=0.926-0.033\beta \\ b(\beta)=1.587-2.332\beta \end{cases} \tag{5.3-451}$$

于是得到宽带随机过程的寿命估算式为

$$T=\frac{N_T}{\hat{\lambda}}=\frac{1}{\hat{\lambda}}\left\{\left(\frac{1}{\sqrt{2}}\right)^{\beta}\frac{C}{\Gamma\left(\frac{2+\beta}{2}\right)}\right\}^{\sigma^{-\beta}} \tag{5.3-452}$$

编写：孙永庆（北京航空航天大学）

第4章　安全评定技术和方法

1　概述

锅炉、压力容器、压力管道等是具有潜在爆炸危险的特殊承压设备，一旦发生爆炸或泄漏，就会造成灾难性事故。因此，确保此类设备的安全性，不仅对于保护和发展生产力具有重要的意义，而且对于维护社会安定和促进改革开放更有重要的现实意义。

锅炉、压力容器、压力管道等特种设备遍布我国各行各业。由于历史、科技和管理上的原因，这类设备普遍存在制造质量差、缺陷严重等问题，加上不少设备超期服役、"带病"运行和安全评估技术落后，爆炸和泄漏事故时有发生。因此，依靠科学技术，对此类设备的缺陷进行科学的安全评估，降低事故率、有效保障安全生产，就显得十分重要和迫切。

20世纪70～80年代以来，随着断裂力学、塑性力学、金属疲劳、无损检测、计算机以及产品质量控制等科学技术的发展和应用，压力容器和压力管道的安全控制与爆炸预防技术由以往的经验型逐步发展为定量评估。1975年国际焊接学会（IIW）发行了"按脆性断裂破坏观点建议的缺陷评定方法"（IIS/IIW-X-749—74），1976年日本焊接协会WSD委员会公布了"按脆性评定的焊接缺陷验收标准"（日本WES-2805k—76），1977年美国机械工程师学会（ASME）公布了锅炉压力容器规程第III卷附录G"防止非延性破坏"及第XI卷"核电站部件使用期中检查规程"附录A"缺陷显示的分析"，1977年英国中央电力局（CEGB）公布了"有缺陷结构完整性的评定（R/H/R6-Revl，77），英国标准协会WEE/37焊接验收标准技术委员会发布了"焊接缺陷验收标准若干方法指南"（英国WEE/137—1978），1980年英国标准协会（BSI）颁布了"焊接缺陷验收标准"（英国BSIPD6493—80），1982年美国电力研究院（EPRI）发布了"含缺陷核容器及管道完整性评定方法"（美国EPRI—82），1984年德国焊接协会（DVS）发行了"焊接接头缺陷的断裂力学评定"（德国DVS-240184），同年，中国机械工程学会压力容器分会和中国化工学会机械与自动化分会制定了"压力容器缺陷评定规范"（CVDA—84）等等。这些标准按其理论基础大体可分为四类：

第一类是以美国ASME的标准为代表的，以线弹性断裂理论为基础的评定方法；

第二类是以英国BSI PD6493—80为代表的，按裂纹张开位移COD弹塑性断裂准则，采用COD设计曲线为基础的评定方法。属于这一类评定方法的还有国际焊协会标准IIW-X-749—74，英国焊接标准协会的WEE/137—1978、日本WSD的WES2805—1976、德国DVS的DVS2401—84和我国的CV-DA—84等；

第三类是以英国CEGB1986年版的R6为代表的，采用失效评定图技术为基础的评定方法；但是1977年第一次修正和1980年第二次修正的老R6法中采用的失效评定曲线，仍然是COD理论导出的，即仍是利用理想塑性材料的D-M模型和线弹性断裂理论的关系推导而得的曲线方程，理论上是不严格的，只能看作是一种经验关系；

第四类是以美国EPRI方法为代表的，以J积分的理论为基础的评定方法。

综上所述，20世纪70年代末到80年代初，COD设计曲线法在国际上的压力容器缺陷评定标准中占据主流地位。然而COD法有其固有的缺陷：所谓裂纹尖端张开位移的定义是不严格的和不统一的；在理论分析方面，除了有均匀拉伸中心穿透裂纹板的窄条区屈服的D-M模型解外，没有其他的含裂纹结构具有COD的力学分析解，缺乏严格的力学基础；COD设计曲线完全是一种经验方法，宽板试验的应变测试方法在国际上争议很大，宽板与实际结构相差太大，更未考虑载荷性质、约束情况、材料应变硬化规律等因素。因此到80年代中期后，COD设计曲线方法基本上就停滞不前了。

在弹塑性断裂分析中还有一种方法，就是J积分法。J积分是一个在数学、力学上都非常严格的断裂力学参量。J积分值在理论上有着明确的物理意义。虽然Rice于1968年就提出了J积分的概念，但由于J积分计算比较困难，直到1982年，在计算机迅速发展和日益普及之后，各种基本的含缺陷结构的J积分的计算才成为可能，加之美国EPRI提出了弹塑性分析的J积分的工程方法，给出了各种J积分全塑性解的韧性断裂手册，解决了J积分的工程计算问题，并研究发展了R6失效评定曲线，给出了严格的J积分断裂准则的失效评定曲线，从而促进了压力容器缺陷评定技术进一步的发展。1986年英国CEGB对R6曲线作了彻底的修改，主要表现在如下的两个方面：第一，考虑了材料应变硬化的效应，不再沿用从D-M模型得到的失效评定图，而以J积分理论为基础，提供了建立失效评定曲线的三种选择方法；第二，对裂纹稳定扩展的处理方法作了重大改进，提出了缺陷评定的三种类型的分析方法。这样就使R6方法取K因子理论、COD理论及J积分理论等众家之长，使其自身建立在严格科学的基础之上，并成为90年代以来国际上主流的压力容器缺陷评定标准。目前世界各国的压力容器缺陷评定标准均向新R6方法靠拢，相继采用失效评定图技术。例如，原来以COD理论为依据的英国PD6493—80标准已作彻底修改，新版的PD6493—90采用了三级评定法；国际焊接学会的缺陷评定规程IIS/IIW-X-749-74，也参照PD6493—90的方法，采用了失效评定图技术（IIW/IISSST-1157—90）；美国ASME规范第XI卷IWB3650及IWB3640也改用弹塑性断裂理论，采用了失效评定图技术；1991年瑞典也编制了一个以新R6的选择1曲线和第I类评定方法为蓝本的"含缺陷结构安全评定规程"（SA/Fou Report91）。我国近年来根据国际上颁布的含缺陷结构安全评定规程，针对我国压力容器、压力管道长期存在的共性问题，依据弹塑性断裂理论和安全评定技术的进展，组织国内的高层次的专家，进行系统的含缺陷压力容器和压力管道安全评定技术研究，包括"八五"国家重点科技攻关课题"在役锅炉压力容器安全评估与爆炸预防技术研究（85-924-02）"和"九五"国家重点科技攻关专题"在役含缺陷压力管道安全评定关键技术研究（96-918-02-03）"等项目。这些项目是在总结、继承和吸收国内外关于锅炉压力容器安全评定的理论、技术实践和经验，通过系统、综合和攻关研究后完成的。并对若干关键的技术问题进行了重点攻关和综合研究，制定了"在用含缺陷压力容器安全评定"的工程技术规范。这一标准的问世，无疑将使我国在用含缺陷承压设备安全评定技术提高到当代的国际水平；为降低我国锅炉压力容器压力管道灾难性事故率提供了有效的技术手段；并为提升我国在用锅炉压力容器压力管道科学管理水平，以及我国加入WTO之后与国外正常的技术和经贸关系奠定了一定的技术基础。安全评定标准无一例外地均采用了失效评定图技术，技术路线日趋统一。

除了上述锅炉压力容器、压力管道等承压容器和设备外，还有许多其他的在役机械设备也都面临着可靠性和安全性的问题（尤其是一些大型设备，如飞机、港口装卸设备等)。要保证生产、降低事故发生率，安全分析与评定技术和方法尤其重要。

1.1　安全评定的一般原则

安全评定应包括对评定对象的状况调查（历史、工况、环境)、缺陷成因分析、失效模式判断、材质检验（包括性能、损伤与退化等)、应力分析、必要的实验与计算，并根据国家有关标准的规定对评定对象的安全性进行综合分析和评价。

1.2　失效模式的判别

根据上述安全评定的一般原则，这里首先介绍有关失效模式判别的有关基本概念和方法。

1.2.1　失效模式的类型

在役机械设备的失效形式主要有以下四种：焊缝或母材的断裂，主结构变形过大（包括厚度减薄过大)，腐蚀和磨损。

1.2.2　失效模式的判别

焊缝或母材的断裂主要是由材料质量、焊接缺陷、应力集中、残余应力、疲劳、超载等原因引起的。它将会导致裂纹的扩展直至构件开裂、断裂；主结构（主梁、悬臂梁、支腿等）变形过大主要是由设计储备小、构件初始变形过大、意外碰撞、严重超载等原因引起的。它将会导致结构应力分布不均、承载能力下降、直至局部屈曲和整体失稳；腐蚀和磨损等主要是由于构件进水或积水、油漆养护不当、受腐蚀性物料或气体的腐蚀以及擦碰等原因引起的，它将会导致减小构件承载静面积，直至构件断裂或屈曲。

判断失效模式应依据同类产品或结构的失效分析和安全评定案例与安全评定的经验，对被评定的产品或结构的具体的制造和检验资料、使用工况以及对缺陷的理化检验和物理诊断结果，且对可能存在的腐蚀、应力腐蚀和高温蠕变环境等对失效模式和安全评定的影响也应予以充分的考虑。

具体的失效模式的判别可参考本篇第2章的内容。根据所分析确定的失效模式进行相应的安全评定。

1.2.3　安全评定的方法的选择

通常是以避免在规定工况（包括水压试验）下的安全评定保证期内发生各种失效模式导致事故的可能为原则。一般来说一种评定方法只能评定相应的失效模式，只有对各种可能的失效模式进行综合判断或评价后，才能作出该产品或结构是否安全的结论。

因此在实际工作中，常常要采取多种安全评定方法进行评定，最终得出综合的评定结论。这其中平面缺陷和体积缺陷的断裂安全评定方法和疲劳评定方法（尤其是针对压力容器和压力管道等的安全评定技术和方法）是目前国内外相对成熟的，而且也是应用最广泛的技术和方法。因此本章第2节和第3节将分别对这两部分内容加以介绍。

2　断裂和塑性失效确定性安全评定技术和方法

如前所述，断裂和塑性失效安全评定技术和方法在国内外的研究和应用是相对比较成熟的，因此本部分首先介绍它的有关内容。

由上节安全评定的一般原则可知，安全评定首先对评定对象进行状况调查（历史情况、现使用工况、环境等)，缺陷检测、缺陷成因分析、失效模式分析、应力分析、相应的实验、计算等，主要可分为检测、分析、评定三大部分，(本节主要应用了断裂力学、失效分析、损伤力学、材料力学、弹塑性力学、金属疲劳力学和无损检测技术等多种理论和技术，其相关的基本理论请参见第3章的内容）然后根据相关的理论、标准对评定对象的安全性进行综合分析和评估。

下面从检测及评定程序、安全评定方法的分类、安全系数的选取、平面缺陷和体积缺陷的断裂安全评定方法等几个部分来阐述本节的内容。

2.1　检测及评定程序

这里首先介绍在一般情况下对机械设备进行检测及安全评定的程序。在实际工作中，根据具体情况，对该程序可进行相应的变动。

2.1.1　在调研的基础上编制整个工作大纲

由于各设备之间本身结构的差异较大，且使用情况也不尽相同，因此首先必须查阅有关设备的设计、运行和维修等技术档案，了解金属结构的使用情况、损伤情况等，尤其是关键部位和已发现损伤部位的情况，有的放矢的确定具体检测和评定内容。安全评定所需的参考资料有：

1）产品制造的图样及强度计算书；

2）产品验收的有关资料，包括材料数据、焊接记录、返修记录、无损检测资料、热处理报告、检验记录和压力试验报告等；

3）产品运行状况的有关资料，包括介质情况、工作温度、载荷状况、运行和故障记录及历次检验与维修报告等。

2.1.2　针对不同的失效形式进行检测

评定需要一些基础数据，其中包括材料自身的性能参数，应力参数以及缺陷参数，具体参数包括：

1）缺陷的类型、尺寸和位置；

2）结构和焊缝的几何形状和尺寸；

3）材料的化学成分、力学和断裂韧度性能数据；

4）由载荷引起的应力、残余应力。

根据安全评定的要求，需要对被评定对象可能存在的各种缺陷、材料和结构等合理选择有效的检测方法和设备进行全面的检测并确保缺陷检测结果准确、真实、可靠。由上述资料可以查阅一些所需要数据，但有一些数据还必须通过现场测试才能获得，因此需要针对不同的失效模式进行检测。对于无法进行无损检测的部位存在缺陷的可能性应有足够的考虑，安全评定人员和无损检测人员需要根据已往的工作经验和具体情况作出保守的估计。通常主要检测的位置和所使用的检测手段如下。

(1) 进行（主）结构尺寸和表面大缺陷检测

由于在役设备不同于刚交付使用的新设备，经过长期的运行，载荷越来越重，主梁等主要结构可能产生下挠、侧弯等结构变形问题，因此要对其检测。检测仪器可采用钢丝、水准仪、激光直线仪等（建议采用激光直线仪测量，该方法方便、快捷、准确)。

由于构件进水或积水、涂层养护不当等原因可能引起的构件的锈蚀、腐蚀，导致构件承载静面积减小，因此要对关键部位的构件厚度进行检测，以确定实际承载面积。可采用超声波测厚仪等无损探伤设备来进行检测。

另外还要通过目测和简单的测试来检测一些较大的表面缺陷（如裂纹、主要结构尺寸与图样的一致性等)。

对于一些大型设备，如港口设备，若对其进行全面检测工作量非常大，好像盲人摸象，容易造成漏检；另一方面，人们关心的是设备中的活动缺陷，一些非活动缺陷的存在是允许的。基于以上原因建议在大型设备的检测过程中可采用声发射检测。采用声发射检测可以快速全面地确定活动缺陷的大体位置及其活动程度。它不仅可以提高检出率，还可以减少整个检测的劳动量。

尽管声发射检测具有高效、准确的特点，但它不能指出具体缺陷的位置、大小，因此无法进行具体的评定和维修，因此还必须进行详细的检测。检测方法可以是目测、着色（磁粉）渗透法（表面裂纹）和超声波检测仪、射线检测仪（埋藏缺陷）等方法。

（2）进行力学性能和应力测试

材料性能的斥试和性能数据的获得应按有关标准的规定，并充分考虑材料性能数据的分散性，按偏于保守的原则确定所需的材料性能数值。静态刚度可用经纬仪、直尺、激光直线仪等进行测量；动态刚度和应力可用电阻应变片和动态应变仪来进行检测。如有可能还应检测随机载荷作用下构件的时间历程（每一种工况应不少于 40~50 个工作循环）。

由于长期在恶劣环境下的使用可能导致材料力学性能的退化，因此除静、动刚度外，（若有可能）还应进行材料性能退化测试。该测试需要在现构件上采适量的样品进行实验室分析。

（3）应力分析

考虑各种可能的载荷，并根据具体失效模式的安全评定需要和评定方法，采用成熟、可靠的方法计算评定中所需的应力。对结构的关键部位，采用以上检测的实际结果（尺寸、厚度、力学性能等）进行有限元分析计算，尤其要对位置非常重要、应力较高的部位进行网格细化处理。通过有限元的计算分析首先确定设备（或构件）的整体强度是否满足要求；同时确定局部和整体应力较高的区域。

在应力的计算分析之后，若有条件应进行现场实际工作应力的测试，以便对有限元方法计算的结果进行验证和校和。

通过前面的有限元计算和应力测试，对现（在）役设备的总体强度进行了安全性的初步评定，但它没有考虑到裂纹扩展的潜在危险和疲劳的情况。因此为了更加全面和真实地评价在役设备的安全情况，必须进行总的安全评定。总的安全评定主要从两个方面进行，一个是静载下的断裂安全评定，另一个是疲劳安全评定。

利用前面步骤（2）、（3）的结果对构件的缺陷情况、材料性能情况以及应力情况进行分析，然后根据国家的一些相关标准，对缺陷乃至整体设备进行上述两个方面的安全评定，并给出相应的检测维修周期。

由于理论方法和检测手段的限制，此安全评定不可能做到绝对准确，评定的结果大都是偏安全的，因此需要根据缺陷所在位置的关键程度区别对待。对于安全评定的危险点，有一些需要立即进行改造，而另一些则需要进行监控。

对于有限元计算强度不足和安全评定需要立即进行改造的部位提出加固和改造方案。该方案需根据现场的条件和设备的具体结构情况来制定。

按相关要求提出的检测维修周期对整体设备和局部缺陷进行定期的检测和维修。

上述步骤不一定在每次检测中全部采用，但也切忌以某个单个的检测结果作为最终的判别和评定依据，以免引起评定失误而造成重大事故。

以上程序可简单总结如下：

1）被评定对象的设计、制造、安装、使用等基本情况和数据的收集；

2）缺陷检验数据；

3）材料性能数据测试或选用；

4）应力测试和应力分析；

5）综合安全评价与评定结论。

6）缺陷评定完成后，评定单位应依据国家相关法规、规章的规定，及时出具完整的评定报告并给出明确的评定结论和继续使用的条件。

下面主要给出了压力容器和压力管道平面缺陷和体积缺陷的断裂与塑性失效安全评定方法以及检测及评定程序。

2.2　确定性安全评定方法的分类

20 世纪 80 年代以来，世界各国出版或再版了一系列缺陷安全评定规范。其中的 R6-86 第三次修订版、PD6493-91、IIW/IIS 导则 SST-1157-91 等三个规程在断裂评定中都是采用三级评定路线，但 PD6493 的三级评定和 R6 的三级评定无论是目的还是方法上均有较大差别。

R6 方法的三级评定可简称为三种选择、三种类别。它全部采用失效评定图技术，而失效评定曲线有三种选择：第一选择是按各种材料的选择第二、三选择曲线的下包络线作出的保守的通用失效评定曲线，非常简单，可用于任何材料和任何结构；第二选择是 Ainsworth 提出的仅反映材料性能而忽视结构因素的以参考应力法进行简化的失效评定曲线；第三选择是严格的 $K_r = f(L_r) = \sqrt{J_e/J}$ 的失效评定曲线；三种类别是指评定的三个不同目的，一为起裂评定，二为有限量撕裂评定，三为撕裂失稳评定。R6 的三级评定较全面地反映了以失效评定图技术为基础的断裂评定方法。

PD6493—91 的三级评定也都是采用失效评定图的技术：第 1 级评定实际上是继承了老版的 COD 设计曲线法，但是以失效评定图的形式表示，是初步的筛选方法；第 2 级采用了在新版 R6 中已经放弃了的老 R6 的以 D-M 模型为基础的老通用失效评定图；第 3 级评定采用 R6 的选择 2 曲线，在应力应变曲线不能确定（例如热影响区）时，采用 R6 的选择 1 的通用失效评定曲线。作为它的主要的、新的、全面的评定方法，第 2 级评定和第 1 级评定一样，均以评定点落在失效评定曲线以内还是以外来判断其安全程度，但应该指出的是它并未指出是评定起裂还是有限量撕裂。R6 的第 3 级评定方法，还可以用于具有高应变硬化能力材料需要分析稳定撕裂失稳断裂的场合。

我国对于锅炉压力容器等的有关标准失效评定也采用三级评定的技术路线，与英国 CEGBR6-91（新 R6 方法）的三级路线以及与英国 BSIPD6493-91 的三级评定相似，但又有我国特有的特色和创新之处。我国标准失效评定的三级评定分别是：平面缺陷的简化评定（简称简化评定）、平面缺陷的常规评定（简称常规评定）和平面缺陷的分析评定（简称分析评定），也可称为一级、二级、三级评定。

我国标准的这种处理，较好地实现了既要继承和发展、积极跟踪国际先进技术、与国际规范接轨，又要反映我国成熟的科研成果和实践经验使之具有中国特色的愿望。为了便于使用，每一级评定只给出一种主要方法。

总结起来，依据评定对象的缺陷类型和评定准则的不同，可分为下列类型。

对于平面缺陷有：

1）平面缺陷的确定性简化评定（简称简化评定）；

2）平面缺陷的确定性常规评定（简称常规评定）；

3）平面缺陷的确定性优化评定（简称分析评定）。

对于体积缺陷有：

1）凹坑缺陷的确定性评定（简称凹坑评定）；

2）气孔和夹渣缺陷的确定性评定（简称气孔夹渣评定）。

对于平面缺陷，可采用简化评定与常规评定的方法进行，当二者的评定结果发生矛盾时，以常规评定结果为准。在特殊和可能的情况下，也可按分析评定方法进行更为详尽的分析。以上的几种评定方法在后面的内容中将分别加以介绍。

2.3 安全系数的选取和确定

通过分安全系数的选取，解决了压力容器平面缺陷简化评定与常规评定之间、常规评定与分析评定之间合理衔接问题，以防止简化评定通过而常规评定不通过和常规评定通过而分析评定不通过的“逆转”现象，使对平面缺陷的弹塑性三级评定方法成为一个比较完整的体系。

在平面缺陷的简化评定、常规评定和分析评定中，分别采用了不同的分安全系数，使三级评定方法合理衔接。通过简化评定与常规评定各自在失效评定图上的“安全”区域的比较、各自的多余安全裕度估算比较以及大量的实际或模拟案例的评定结果的比较，均未发现简化评定通过而常规评定不通过的“逆转”情况。不少评定案例用常规评定与分析评定方法同时进行评定的评定结果的比较表明，不存在“逆转”现象。这种合理的衔接安排，为这三级评定方法建立了各级既相对独立，又相互联系和衔接的合理关系。也为用户根据实际情况采用任何一级安全评定方法进行平面缺陷的安全评定以判断其安全性提供了可能。并且用户一般可采用简化评定，但在简化评定不通过时，允许采用常规评定或在可能情况下直接采用分析评定对含缺陷结构做出安全与否的最终评价。或当常规评定不通过时，在可能情况下可再采用分析评定方法对其进行分析评定，以便于科学地得出安全与否的最终结论。

根据相关文献和国内的有关标准规定，常规评定所采用的安全系数见表5.4-1。其他评定方法所采用的安全系数，按各相应的规定选取。

表5.4-1　常规评定安全系数取值

失效后果	缺陷表征尺寸分安全系数	材料断裂韧度分安全系数	应力分安全系数	
			一次应力	二次应力
一般	1.0	1.1	1.2	1.0
严重	1.1	1.2	1.5	1.0

2.4 平面缺陷的确定性安全评定

由本章2.2节的内容可知，平面缺陷安全评定可分为三级，在本节将分别对这三级评定分别加以介绍。由2.1节的内容可知，无论对哪种评定方法而言，都需要三大类数据：缺陷描述、应力情况、材料性能参数。因此在介绍评定方法之前，本节将首先阐述有关这三方面数据的获取方法。

2.4.1 平面缺陷的表征及其处理

(1) 平面缺陷的表征

安全评定时，一般应按规定对实测的平面缺陷进行规则化表征处理，将缺陷表征为规则的裂纹状表面缺陷、埋藏缺陷或穿透缺陷。表征后裂纹的形状为椭圆形、圆形、半椭圆形或矩形。

表征裂纹尺寸应根据具体缺陷情况由缺陷外接矩形之高和长来确定。对穿透裂纹来说，长为 $2a$；对表面裂纹而言，高为 a、长为 $2c$；对埋藏裂纹而言，高为 $2a$、长为 $2c$；对孔边角裂纹，高为 a、长为 c（见图5.4-1）。缺陷外接矩形之长边应与邻近的壳体表面平行。

(2) 表面缺陷的规则化和表征裂纹尺寸

若缺陷沿壳体表面方向的实测最大长度为 l，沿板厚方向的实测最大深度为 h（见图5.4-2），则

当 $h>0.7B$ 时，规则化为长 $2a=l+2h$ 的穿透裂纹（见图5.4-2a）；

当 $h\leqslant 0.7B$ 时：

1) $h<l/2$ 时，规则化为 $c=l/2$、$a=h$ 的半椭圆表面裂纹（见图5.4-2b）；

2) $h\geqslant l/2$ 时，对于断裂评定，规则化为 $c=a=h$ 的半圆形表面裂纹（见图5.4-2c）；对疲劳评定，规则化为 $2c=l$、$a=h$ 的半椭圆表面裂纹（见图5.4-2d）。

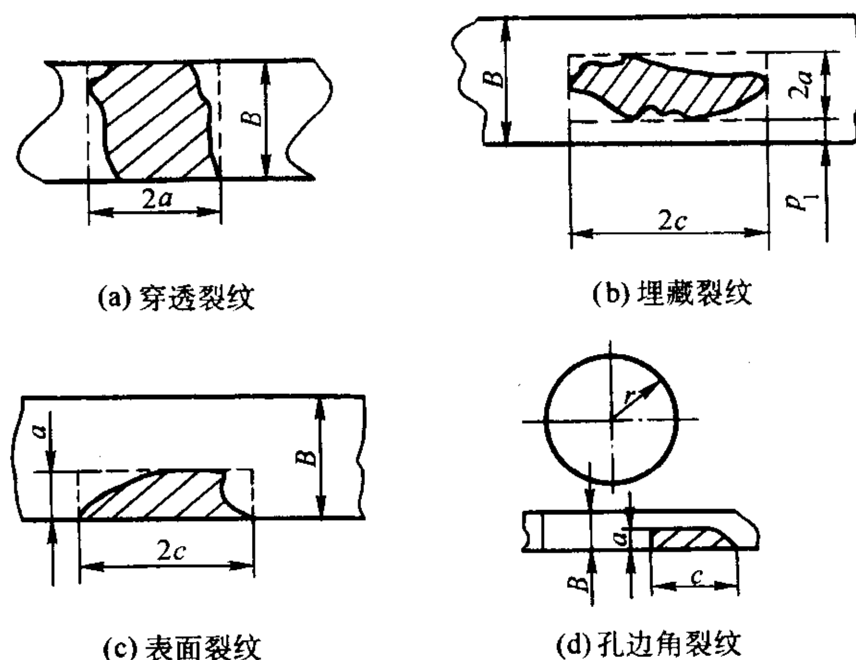


图5.4-1　平面缺陷的表征图例

(3) 埋藏裂纹的规则化和表征裂纹尺寸

若缺陷沿壳体表面方向的实测最大长度为 l，沿板厚方向的实测最大自身高度为 h，缺陷到壳体内外表面的最短距离分别为 p_1 和 p_2，且 $p_1\leqslant p_2$（见图5.4-3），则

当 $p_1\leqslant p_2<0.4\ h$ 时，规则化 $2a=l+2h$ 的穿透裂纹（见图5.4-3a）；

当 $p_1<0.4\ h\leqslant p_2$ 时，规则化为 $a=h+p_1$、$2c=l$ 的半椭圆表面裂纹（见图5.4-3b）；

当 $0.4\ h\leqslant p_1\leqslant p_2$ 时：

1) $h<l$ 时，规则化为 $2c=l$、$2a=h$ 的椭圆形埋裂纹（见图5.4-3c）；

2) $h>l$ 时，对于断裂评定，规则化为 $2c=2a=h$ 的圆形埋藏裂纹（见图5.4-3d）；对于疲劳评定，规则化为 $2c=l$、$2a=h$ 的椭圆埋藏裂纹（见图5.4-3e）；

3) 已表征为表面裂纹的埋藏缺陷，即使 $2a+p_1>0.7B$，也不再表征为穿透裂纹。

4) 穿透缺陷的规则化和表征裂纹尺寸　若穿透缺陷沿壳体表面方向的实测最大长度为 l，则规则化为 $2a=l$ 的穿透裂纹（见图5.4-4）。

(4) 斜裂纹的表征

当裂纹平面方向与主应力方向不垂直时，可将裂纹投影到主应力方向垂直的平面内，在该平面内按投影尺寸确定表征裂纹尺寸。

(5) 裂纹群的处理

若机械设备在局部范围同时存在多个缺陷，当两裂纹或多裂纹相邻时，应考虑裂纹之间的相互影响。在考虑其影响前首先要确定裂纹间的距离 s 和合并临界间距 s_0，然后根据具体情况分别作出合并（复合）、考虑相互影响或忽略相互影响的处理。

裂纹间距 s 及合并间距 s_0 的确定：

共面裂纹的合并规则如图5.4-5所示，其中的几个典型情况如下。

在图5.4-5a中，$s=s_2$、$s_0=2c_2$；

在图5.4-5b、图5.4-5c、图5.4-5d中，若 $\frac{s_1}{2a_2}>\frac{s_2}{2c_2}$，则 $s=s_1$、$s_0=2a_2$；否则 $s=s_2$、$s_0=2c_2$；

在图5.4-5e中 $s=s_2$、$s_0=2a_2$。

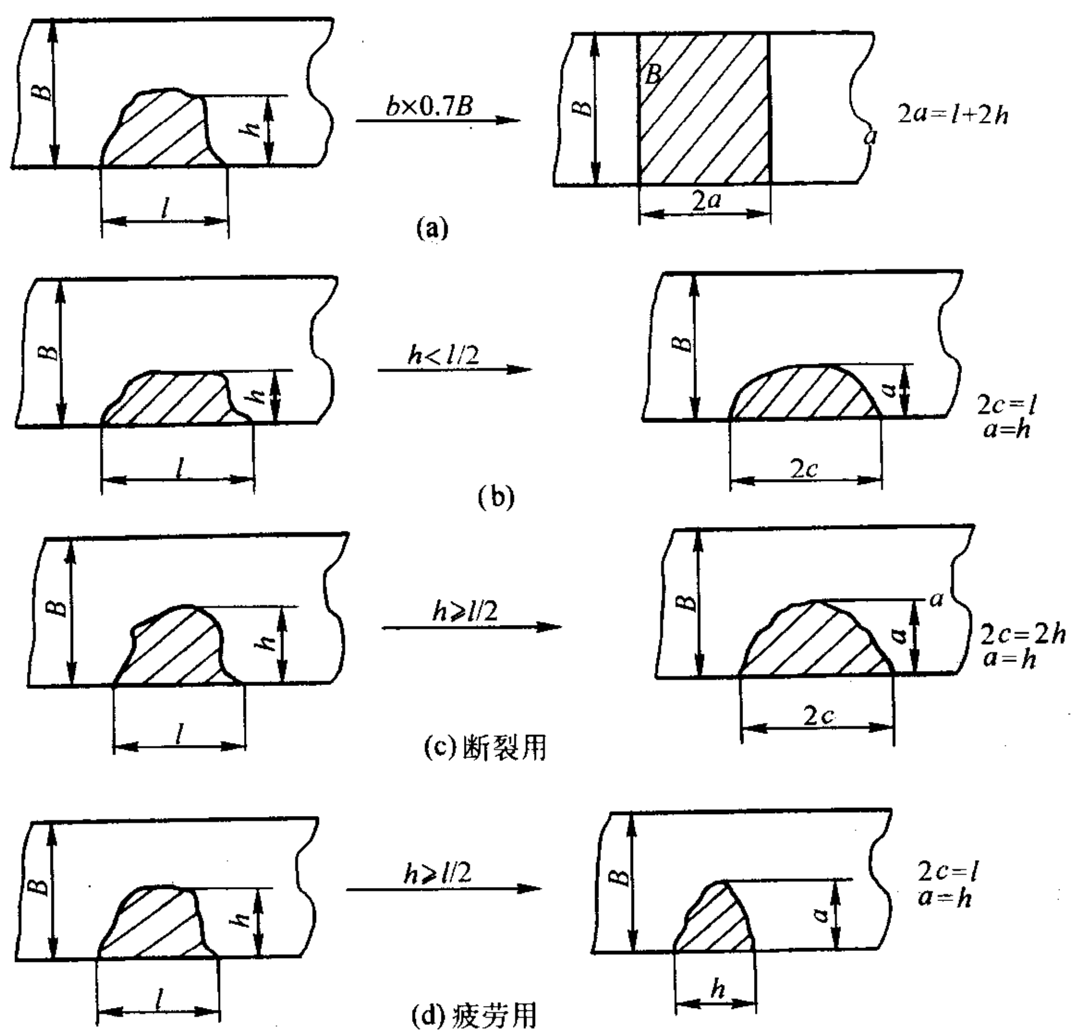


图 5.4-2　表面缺陷的规则化图例

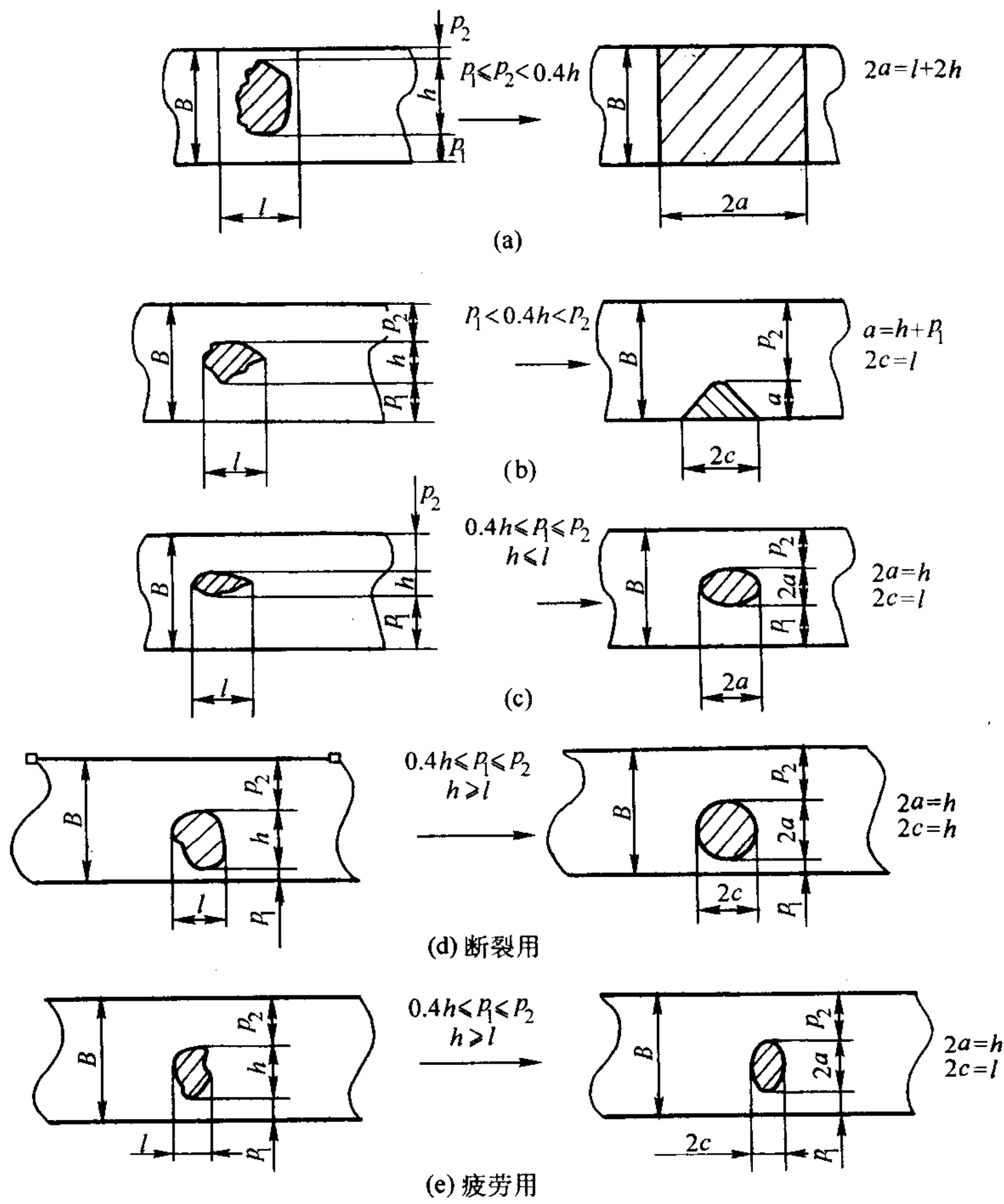


图 5.4-3　埋藏缺陷的规则化图例

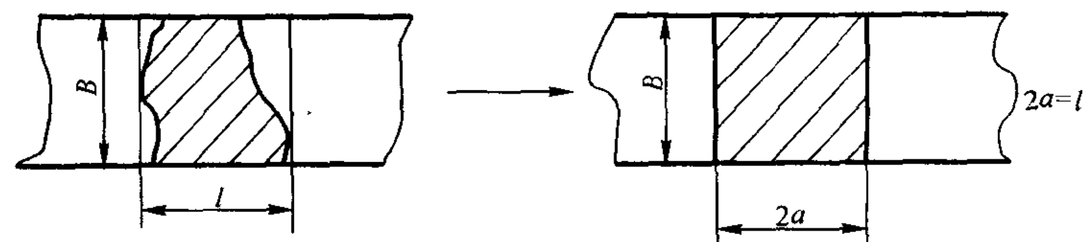


图 5.4-4　穿透缺陷的规则化图例

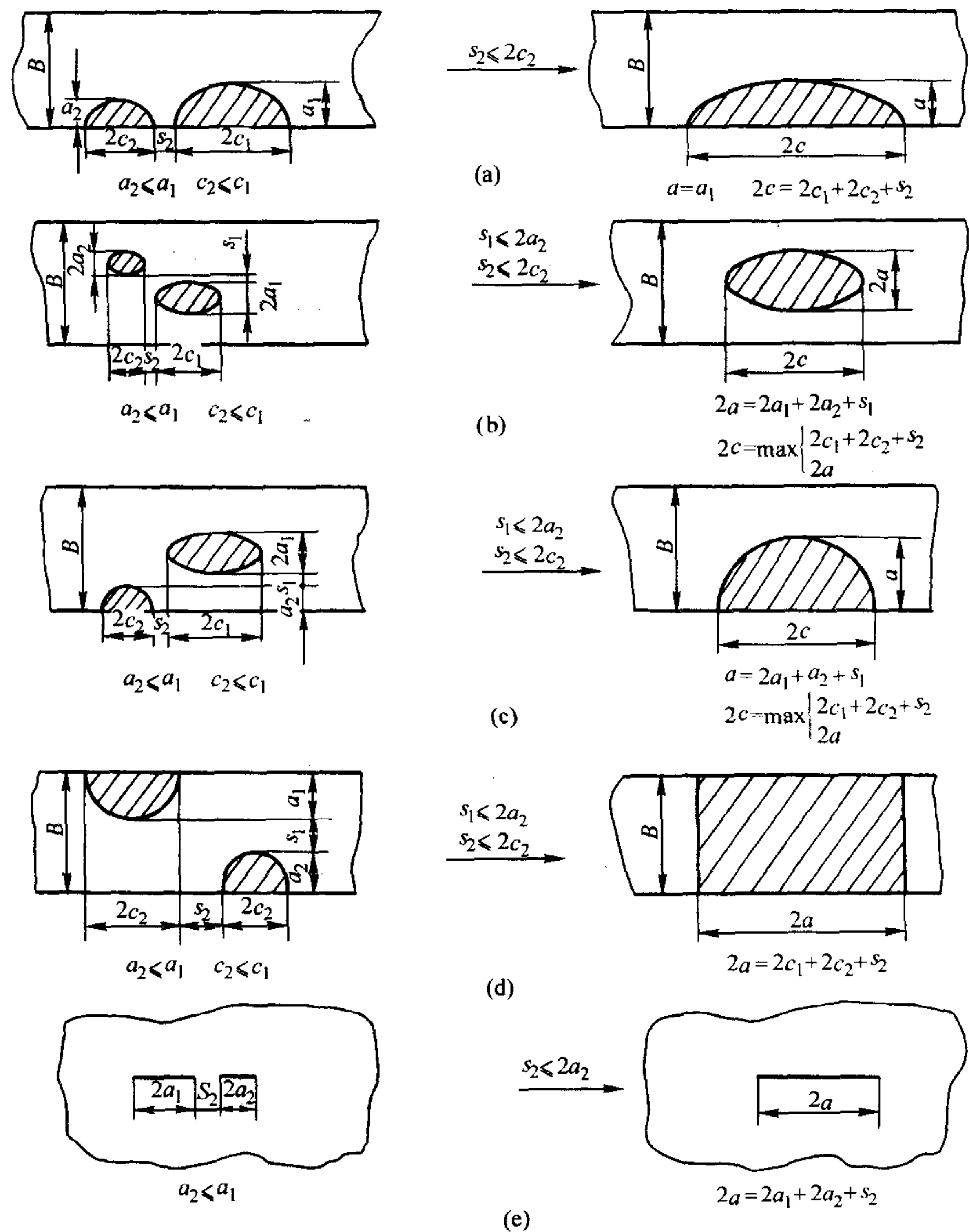


图 5.4-5　共面缺陷的合并规则图例

对于裂纹间的相互影响，这里主要考虑两种情况：共面裂纹和非共面裂纹。

共面裂纹的复合及相互影响处理方法如下。

对于共面裂纹之间的相互影响处理可为三种方式：合并(复合)、采用干涉效应系数和忽略其影响。该三种方式的应用条件如下。

1）当 $s \leqslant s_0$ 时，采用包络该两裂纹（或两个以上 $s \leqslant s_0$ 的裂纹）的外切矩形将其复合，规则化为一个裂纹。对已经复合后的裂纹将不再进行表征，也不再与其他裂纹复合。复合后的裂纹与其他裂纹或复合裂纹之间的相互影响需按 b 或 c 的规定确定。

2）当 $s_0 < s < 5s_0$ 时，则两裂纹不必合并，分别按单个裂纹评定，但要考虑其相互间的影响。即在简化评定中，计算的$\sqrt{\delta_r}$的值需乘以 1.2 的系数；常规评定中，在计算 K_r 时要将应力强度因子乘以弹塑性干涉效应系数 G；疲劳评定中，在计算 ΔK 时要乘以线弹性干涉效应系数 M。下面介绍干涉系数 G，M 的估算方法。

① 线弹性干涉效应系数 M 的估算方法如下：

$$M = 1 + 0.053\frac{s_0}{s} \tag{5.4-1}$$

式中，s 和 s_0 分别为裂纹间距和合并间距。

② 弹塑性干涉效应系数 G 的估算方法　该方法为我国首创。在工程实际中缺陷往往不是孤立存在的。目前世界各国规范都是以相邻缺陷的存在导致应力强度因子增加率达到一定程度时必须将二裂纹看作为已贯穿的一个大裂纹处理。我国“八五”攻关研究发现，裂纹间的弹塑性干涉效应$\sqrt{J_{双}/J_{单}}$比弹性干涉效应 $K_{I双}/K_{I单}$ 大得多，并与材料的本构关系和载荷水平有关。例如两相邻等长裂纹间距超过裂纹长度时，按我国的老规范就不考虑二者间的相互影响了，但计算表明对 A533B 材料来说，在 $L_r > 1$ 后弹塑性干涉效应 $G = \sqrt{J_{双}/J_{单}}$可达到 1.4，对 16MnR 来说，在 $L_r = 1$ 时，G 可达到 2.1，显然忽视相互影响会带来危险的后果。我国

“八五”科技攻关研究表明，只要在单裂纹的应力强度因子上乘以相邻缺陷的弹塑性干涉效应 G，即可利用通用失效评定曲线完成考虑裂纹间弹塑性干涉效应的断裂评定。根据不同应力应变关系材料的计算结果，这里给出 G 值的估算公式，对应力应变关系符合 Ramberg-Osgood（ROR）关系的材料，有

$$G=\begin{cases}M & 0\leqslant L_r\leqslant L_1\\ M+A_1\ (0.076n-0.19)\ (L_r-L_1) & L_1<L_r\leqslant 1.0\\ G_1 & 1.0<L_r<L_r^{\max}\end{cases}\qquad(5.4\text{-}2)$$

式中，n 为材料硬化指数；L_1、A_1、G_1 可由下式获得。

$$L_1=0.24+0.04(n-5),A_1=0.016n-0.056+1.024\left(\frac{s_0}{s}\right)^2-0.048\left(\frac{s_0}{s}\right)^2,G_1=M+A_1(0.076n-0.190)(1-L_1)。$$

上式对 $5\leqslant n\leqslant 12$ 有效，但对 $n<5$ 的材料可保守地取 $n=5$进行计算。

L_r 为载荷比，是指引起一次应力的施加载荷与塑性屈服强度的比值，表示载荷接近于材料塑性屈服强度的程度。这里 L_r 取两个裂纹中较大裂纹的 L_r 值（其具体算法详见安全评估部分 4.4 节的相关内容）。

而对于屈服应变平台长度小于 0.02 的短屈服平台材料，G 需按式（5.4-3）估算。

$$G=\begin{cases}M & 0\leqslant L_r\leqslant L_2\\ M+0.1\ (L_r-L_2) & L_2<L_r\leqslant L_3\\ G_2+B_1\ (L_r-L_3) & L_3<L_r\leqslant 0.90\\ G_3+B_2\ (L_r-0.90) & 0.90<L_1\leqslant 1.00\\ G_4-B_3\ (L_r-1.00) & 1.00<L_r\leqslant 1.05\\ G_5 & 1.05<L_r\leqslant L_r^{\max}\end{cases}\qquad(5.4\text{-}3)$$

式中，L_r 取两个裂纹中较大裂纹的 L_r 值，其他参数可由式（5.4-4）获得。

$$\begin{aligned}L_2&=0.30+0.50\lg\ (s/s_0)\\ L_3&=0.63+0.31\lg\ (s/s_0)\\ B_1&=1.22+0.60\lg\ (s/s_0)\\ B_2&=6.97-4.50\lg\ (s/s_0)\\ B_3&=18.2-12.5\lg\ (s/s_0)\\ G_2&=M+0.1\ (L_3-L_2)\\ G_3&=G_2+B_1\ (0.90-L_3)\\ G_4&=G_3+0.1B_2\\ G_5&=G_4-0.05B_3\end{aligned}\qquad(5.4\text{-}4)$$

不同 s/s_0 下的 L_2、L_3、M、G_2、G_3、G_4、G_5 亦可通过相关表格查阅。

③ 若 $S\geqslant 5s_0$，则可忽略其相互影响，分别作为单个裂纹进行评定。

2.4.2 平面缺陷确定性评定中所需应力的确定

如前所述，在评定中应全面考虑载荷和应力。下面首先介绍评定中主要（经常）涉及的应力，包括：

1）介质的压力及其产生的应力；

2）介质和结构的重力载荷及其产生的应力；

3）外加机械载荷及其产生的应力；

4）振动、风载等载荷及其产生的应力；

5）焊接引起的焊接残余应力；

6）错边、角变形、壁厚局部减薄、不等厚度等结构几何不连续在载荷作用时所产生的应力；

7）温度差、热胀冷缩不协调等所产生的热温差应力或热应力；

8）其他应该考虑的载荷或应力。

根据应力的作用区域和性质，将其划分为一次应力 P、二次应力 Q。除下列特殊规定外，按 JB/T 4732—1995《钢制压力容器—分析设计标准》确定应力分类的规则：

1）由于管系的热膨胀在接管处引起的应力，按一次应力考虑；

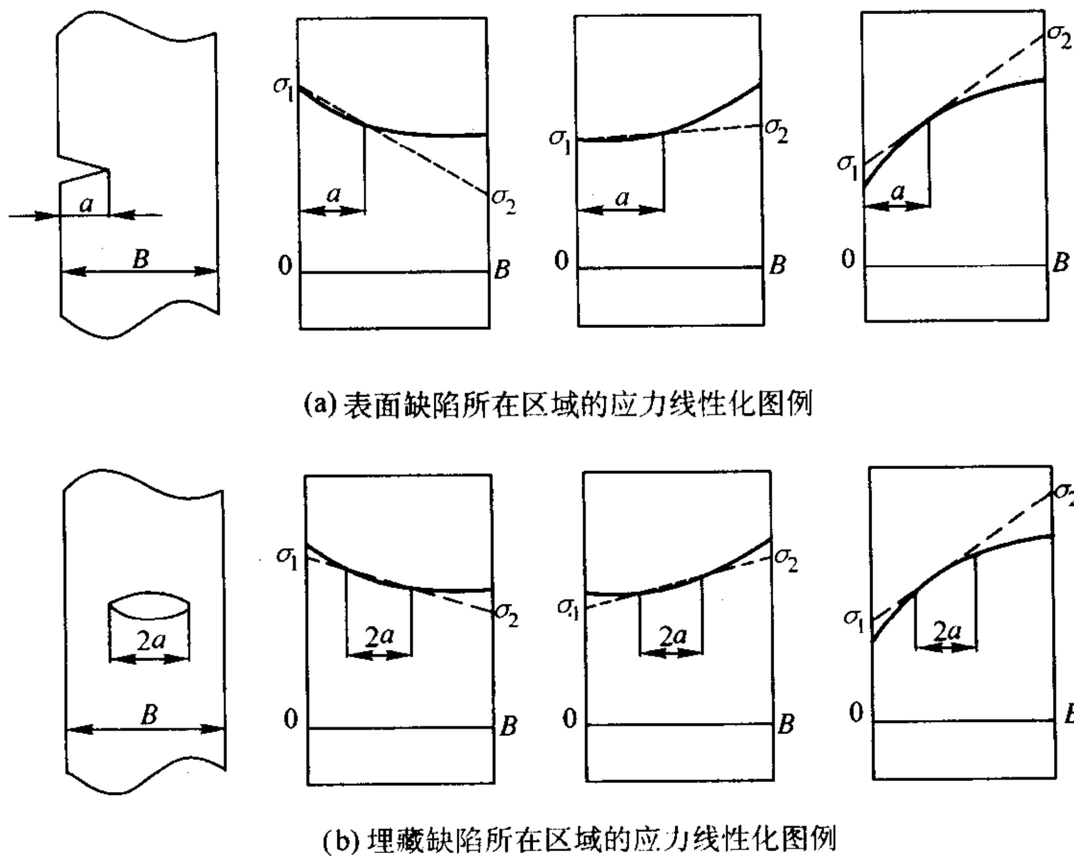
2）焊接产生的残余应力，按二次应力考虑；

3）由错边、角变形、局部厚度差所引起的局部应力，按二次应力考虑；

4）由壁温之差或材料热胀系数不同引起的热应力，按二次应力考虑。

（1）确定性评定应力的确定

在评定中所取用的应力是缺陷部位的主应力，计算该主应力时通常采用线弹性计算方法，并假设结构中不存在缺陷。同时要考虑各种载荷，进而分别计算被评定缺陷部位结构厚度截面上一次应力及二次应力分布，然后将非线性分布的应力针对不同的缺陷形式（表面缺陷和埋藏缺陷）进行缺陷区域的应力线性化处理，如图 5.4-6 中的实线所示。



(a) 表面缺陷所在区域的应力线性化图例

(b) 埋藏缺陷所在区域的应力线性化图例

图 5.4-6 断裂评定中缺陷所在区域的应力线性化图例

对于沿厚度非线性分布的应力，应根据保证整个缺陷长（或深）度范围内各处的线性应力值均不低于实际应力值的原则确定沿缺陷部位截面的线性分布应力，如图5.4-6中虚线所示。

(2) 应力的分解和 P_m、P_b、Q_m、Q_b 的确定

对于圆筒形容器来说，内压会造成容器壁产生周向膜应力，薄壁容器膜应力可由式（5.4-5）获得。

$$\sigma_m = \frac{P_f D_i}{2t} \tag{5.4-5}$$

式中，P_f 为工作压力；D_i 为容器的平均直径；t 为壁厚。

对于厚壁容器来说，其最大周向膜应力为

$$\sigma_{max} = \frac{P_f\ (K^2+1)}{(K^2-1)} \tag{5.4-6}$$

$$K = \frac{D_0}{D_1} \tag{5.4-7}$$

式中，D_0、D_1 分别为容器的外、内径；其最小周向力为

$$\sigma_{min} = \frac{2P_f}{K^2-1} \tag{5.4-8}$$

对于沿厚度直线分布或经线性化处理后的应力，可按式（5.4-9）和式（5.4-10）分解为薄膜应力分量 σ_m 和弯曲应力分量 σ_B。

$$\sigma_m = (\sigma_{max} + \sigma_{min})\ /2 \tag{5.4-9}$$

$$\sigma_B = (\sigma_{max} - \sigma_{min})\ /2 \tag{5.4-10}$$

由一次应力分解而得的 σ_m、σ_B，分别为 P_m、P_b；由二次应力分解而得的 σ_m、σ_B，分别为 Q_m、Q_b。

(3) 残余应力分析

如已知焊接残余应力 σ_R 的实际分布，则按其实际分布确定 σ_m、σ_B，并用于评定计算。如不能得到焊接残余应力的实际应力分布，可参照一些相关资料来确定应力分布或 Q_m、Q_B。这里给出几种典型焊接接头的残余应力的估算：

对于壁厚为 $B<25$ mm 的对接焊接头来说，若力作用在垂直于焊缝的平面上，σ_R 分布（用于垂直焊缝的缺陷）为

$$\frac{\sigma_R}{\sigma_R^{max}} = [1-4\left(\frac{x}{6B}\right)^2]\exp[-2\left(\frac{x}{6B}\right)^2] \tag{5.4-11}$$

并假设其沿厚度均匀分布，取拉伸应力区宽度为 $6B$，B 为评定设备（容器）的厚度；若力作用在平行于焊缝的平面上的 σ_R 分布（用于平行焊缝的缺陷）为 $\sigma_R = 0.3\sigma_R^{max}$，均布在截面上，即

$$Q_m = 0.3\sigma_R^{max},\ Q_b = 0 \tag{5.4-12}$$

对于 $B \geqslant 25$ mm 的对接焊接头来说，筒体环焊缝等低约束对接焊缝，σ_R 沿板厚的分布为

表面裂纹 $a/B \leqslant 0.5$ 时　$Q_m = -\sigma_R^{max}$，$Q_m = 2\sigma_R^{max}$　(5.4-13)

其他情况按线性化规则确定；球罐、厚壁高压容器，σ_R 沿板厚的分布为

表面裂纹 $a/B \leqslant 0.5$ 时

$$Q_m = 0,\ Q_b = \sigma_R^{max} \tag{5.4-14}$$

其他情况按线性化规则确定。

以上情况的 σ_R^{max} 可分为几种情况：

对于焊态结构

$$\sigma_R^{max} = \max\ (\sigma_s^W,\ \sigma_s) \tag{5.4-15}$$

对于经炉内整体消除应力退火热处理的焊接结构

$$\sigma_R^{max} = (0.3 \sim 0.5)\ \max\ (\sigma_s^W,\ \sigma_s) \tag{5.4-16}$$

对于经局部消除应力退火热处理或现场整体热处理的焊接结构，可实测确定或依据经验确定。对焊接修补区、高拘束度焊缝区或焊接残余应力分布情况不明区域，可取焊接残余应力引起的二次应力

$$Q_m = \sigma_s,\ Q_b = 0 \tag{5.4-17}$$

由错边、角变形所产生的应力为二次应力 Q_b，可查阅相关标准。当已知内外壁温差时，温差应力可按下式估算：

对铁素体钢　$Q_b = 1.6\Delta T$

对奥氏体钢　$Q_b = 1.8\Delta T$

2.4.3　材料性能数据的确定

评定过程中主要需要力学、物理性能以及断裂韧度两大类材料性能参数。其中力学和物理性能包括：屈服点 σ_s（或条件屈服强度 $\sigma_{0.2}$）、抗拉强度 σ_b、弹性模量 E 等表征材料力学性能和物理性能的参数；断裂韧度包括：CTOD 断裂韧度（δ_c 或 δ_1，简化评定需要），J 积分断裂韧度（J_{IC}、J_C，常规评定需要），分析评定还要求获得包括 J_{IC} 的整条 J_R（Δa）阻力曲线。

由于每种材料的力学性能参数分散性比较大，因此在评定中优先采用实测数据且实测数据所用的试样尽可能取自被评定缺陷部位的材料，也可取自在化学成分、力学性能、冶金和工艺状态以及使用条件等方面能真实反映缺陷所在部位材料性能的试样。

在无法获得实测数据时，经有关各方协商，在充分考虑材料化学成分、冶金和工艺状态、试样和试验条件等影响因素且保证评定的总体结果偏于安全的前提下，可选取代用数据。

断裂韧度实测试样中的裂纹面和裂纹扩展方向应同被评定结构中的情况一致，也可选取能获得该材料最低断裂韧度数据的其他取样方法。对取自影响区的试样，应考虑裂纹尖端所在部位组织结构类型和晶粒尺寸等的影响。

(1) 材料力学物理性能数据的测试和选取

材料的拉伸性能可按 GB/T 228—2002《金属拉伸试验方法》测定。屈服强度取于0.2%条件屈服强度。焊接热影响区的拉伸性能取母材和焊缝金属中的较低值。在未能实测被评定材料拉伸性能的情况下，可以参照相应钢号的材料标准选取材料的有关拉伸性能指标；没有进行实测且又不能从有关标准中直接查到相应数据时，可通过可靠的方法利用硬度测定值估算材料强度的参考值。弹性模量可参考有关标准或手册。

(2) 材料断裂韧度的测定方法

缺陷评定中所使用的材料断裂韧度应采用标准实验方法来确定。

由第3章的内容可知，在线弹性力学中Ⅰ型裂纹平面应变断裂韧度 K_{IC}，在弹塑性断裂力学中的临界裂纹尺寸 CTOD —— δ_c 和临界 J 积分—— J_{IC} 都是反映材料抗断裂性能的性能参数。下面将分别介绍这三个参量的测试原理及测试方法。

1) 平面应变断裂韧度 K_{IC} 测试原理及测试方法

平面应变断裂韧度是材料在静载下，裂纹尖端处于平面应变状态和屈服范围条件下，Ⅰ型裂纹开始失稳扩展时的应力强度因子。

含裂纹构件的表面层由于没有第三方向的约束，处于平面应变状态；其中心层由于处在周围材料的约束之中，因此处于平面应力状态。当构件厚度达到某一临界值时，平面应力层对其整体断裂韧度的影响将很小，此时其断裂韧度可称为平面应变断裂韧度 K_{IC}，它是与构件厚度无关的材料常数，若构件厚度小于某一临界值时，平面应力层对其整体断裂韧度将起主导作用，这时构件的断裂韧度是平面应力断裂韧度 K_C，它则是与构件厚度有关的量。

由断裂力学的理论可知，随着载荷的增大，裂纹尖端的应力强度因子也将增大，当 K_I 达到某一临界值 K_{cr} 时，裂纹将失稳扩展，此临界值 K_{cr} 即为材料平面应变断裂韧度。根

据此基本理论即可进行 K_{IC} 的测试。目前有关 K_{IC} 的测试方法，很多国家都已制定出相关的规范，最早的 K_{IC} 测试规范是美国材料与试验学会（ASTM）所制订的。在现行的规范中多用弯曲试样和紧凑拉伸试样测试，这里简单介绍一下三点弯曲试样测试 K_{IC} 的方法。

① 测试中对试样尺寸的要求　为了能更好的获得 K_{IC} 值，需要满足小范围屈服（线弹性断裂力学有效性要求）和平面应变两个条件。要满足这两个条件，则必需对裂纹长度 a，试样韧带尺寸 b 以及试样厚度 B 提出要求。

裂纹长度的要求：

为了保证线弹性断裂力学的有效性，对于三点弯曲试样的裂纹长度 $a \geqslant 15\pi r_y$ 时，才能保证 K_I 的相对误差小于 7%，其中 r_y 为塑性区半径，而对于平面应变状态有

$$r_y = \frac{1}{6\pi}\left(\frac{K_{IC}}{\sigma_s}\right)^2 \tag{5.4-18}$$

于是有

$$a \geqslant 2.5\left(\frac{K_{IC}}{\sigma_s}\right)^2 \tag{5.4-19}$$

上述计算方法通过有限元计算和实验结果表明基本上是适当的。

裂纹韧带尺寸 b 的要求：

裂纹韧带尺寸 b 既能影响小范围屈服，也能影响平面应变条件。因此它也必须满足下述条件：

$$b = W - a \geqslant 2.5\left(\frac{K_{IC}}{\sigma_s}\right)^2 \tag{5.4-20}$$

式中，W 为试样宽度。

试样厚度 B 的要求：

如上所述，只有试样的厚度超过某一值后才能将试样看作是处于平面应变状态。大量实验证明，试样需满足下述条件：

$$B \geqslant 2.5\left(\frac{K_{IC}}{\sigma_s}\right)^2 \tag{5.4-21}$$

② K_{IC} 测试装置简述　三点弯曲试样和三点弯曲试样实验装置如图 5.4-7 和图 5.4-8 所示。测试一般在万能试验机上进行。通过载荷传感器和夹式引伸仪采集载荷 P 和裂纹张开位移 V 的信号，经过放大后，再传入 $x-y$ 记录仪，最终获得载荷位移曲线，即 $P-V$ 曲线。

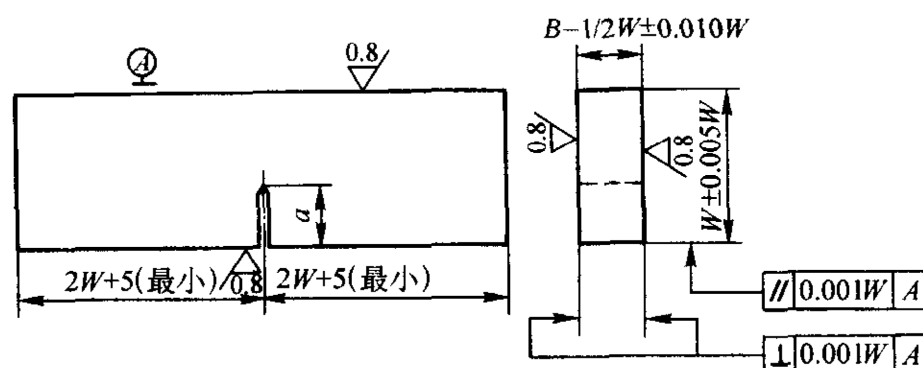


图 5.4-7　三点弯曲试样

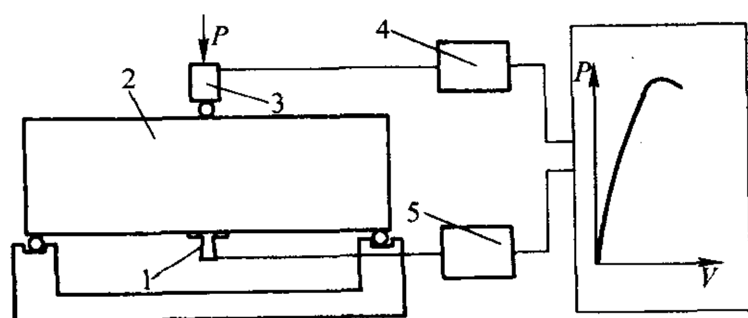


图 5.4-8　三点弯曲试样实验装置示意图

1—夹式引伸仪；2—试样；3—载荷传感器；4—应变仪；5—$x-y$ 记录仪

③ 测试结果的分析计算

a）裂纹长度的确定　用断裂后试样的截面测量裂纹长度 a，试样厚度的 1/4，2/4，3/4 三个位置的裂纹长度分别为 a_1，a_2，a_3，求其平均值，

$$\bar{a} = \frac{1}{3}(a_1 + a_2 + a_3) \tag{5.4-22}$$

若三个值中的最大值与最小值之差大于 0.025 W，或任一处的疲劳裂纹长度小于 0.025 W 或 1.5 mm；或表面处裂纹长度小于 $\bar{a}$；或 $\bar{a}$ 与试样的宽度之比超出了 0.45～0.55 的范围，则该试样无效，应重新制作试样。若满足上述条件，则 $a = \bar{a}$。

b）条件载荷 P_Q 的确定　获得的 $P-V$ 曲线通常为图 5.4-9 所示的三种典型类型。条件载荷可按如下规则确定：作一斜率比 $P-V$ 曲线的线性段斜率低 5% 的割线，它与 $P-V$ 曲线的交点定义为 P_5。

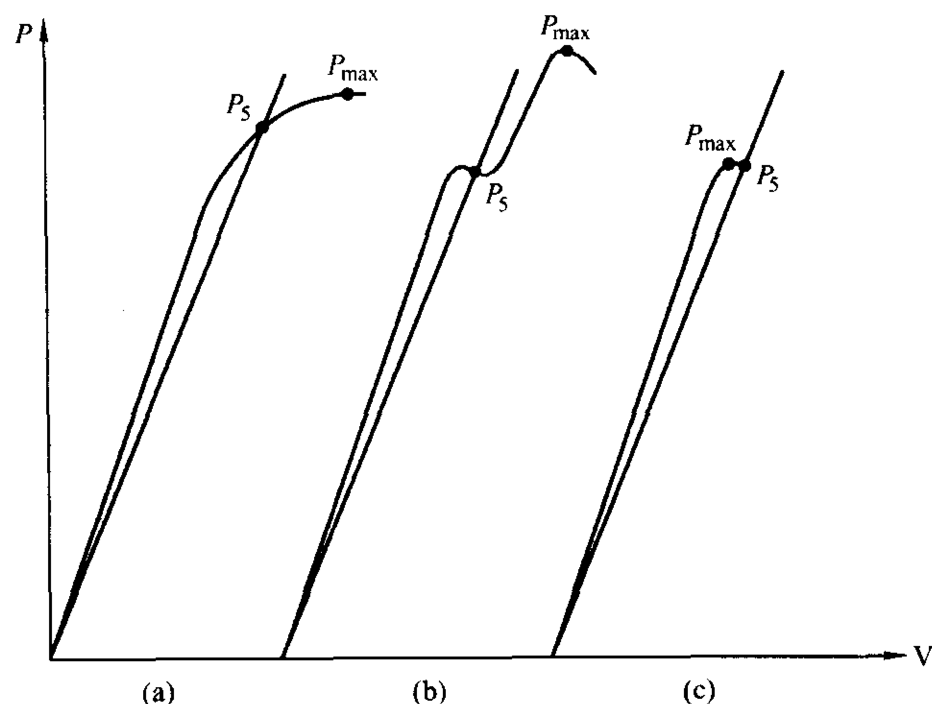


图 5.4-9　$P-V$ 曲线的三种典型类型

对于材料韧性较好的或试样厚度较小的时候，试样中处于平面应变状态的部分相对较少（见图 5.4-9a），看不出明显的迸发平台，$P-V$ 线上所有点的载荷都低于 P_5，这时取 $P_Q = P_5$。

对于材料韧性一般或试样不很厚的情况，试样中平面应力层占一定比例。在加载过程中，当载荷达到某一值时，试样中心层因处于平应变状态首先达到临界状态，裂纹在中心层开始扩展，而试样表面层却因处于平面应力状态裂纹尚不扩展，中心层裂纹的扩展很快被表层的扩展阻力所止住，这时反映在 $P-V$ 线上即为一个平台，这种现象称为迸发。试验中经过迸发阶段后，试样的整体仍有承载能力，直到最大载荷 P_{max}，试样才发生断裂（见图 5.4-9b）。在 P_5 以前存在有载荷大于 P_5 的点，这时取其中最高点的载荷作为条件载荷 $P_Q = P_{max}$。

而对于韧性较差的材料或试样厚度足够大时，试样的平面应力层较小，试样几乎完全处于平面应变状态。在失稳断裂前几乎不发生扩展，当载荷增大到最大值时，试样突然发生脆断（见图 5.4-9c）。同样在 P_5 以前存在有载荷大于 P_5 的点，这时取其中最高点的载荷作为条件载荷 $P_Q = P_{max}$。

c）K_Q 的计算　对应于 P_Q 而计算得到的应力强度因子为条件断裂韧度值 K_Q。根据上面所获得的各数据即可计算条件断裂韧度值 K_Q：

$$K_Q = \frac{P_Q L}{BW^{1/2}} f\left(\frac{a}{W}\right) \tag{5.4-23}$$

$$f\left(\frac{a}{W}\right) = \frac{3(a/W)^{1/2}[1.99-(a/W)(1-a/W)(2.15-3.93a/W+2.7a^2/W^2)]}{2(1+2a/W)(1-a/W)^{3/2}}$$

式中，L 为试样长度。

d）K_Q 的有效性判断　由式（5.4-23）得到的 K_Q 是否

即为材料的平面应变断裂韧度，还需进行有效性判断。其判据主要有两个方面：载荷判据，平面应变与 K 主导判据。

载荷判据：
$$\frac{P_{max}}{P_Q} \leqslant 1.1 \tag{5.4-24}$$

式中，P_{max} 为 $P-V$ 曲线上的最大载荷值。此判据是由实验中总结出来的。

平面应变与 K 主导判据：
$$B \geqslant 2.5\left(\frac{K_{IC}}{\sigma_s}\right)^2 \tag{5.4-25}$$
$$a \geqslant 2.5\left(\frac{K_{IC}}{\sigma_s}\right)^2 \tag{5.4-26}$$
$$b \geqslant 2.5\left(\frac{K_{IC}}{\sigma_s}\right)^2 \tag{5.4-27}$$

这一判据是由上述试样要求中得来的，由于此时材料的 K_{IC} 还未知，因此只能用 K_Q 代替。

若 K_Q 能满足上述两个判据，则所得到的 K_Q 即为材料的平面应变断裂韧度 K_{IC}。否则就应该更改试样尺寸重新做试验。

具体的实验参数请参见相关标准或规范。

2）裂纹张开位移 δ_c 的测试原理及测试方法　CTOD 的测试方法和测试技术同 K_{IC} 一样都已比较完善，很多国家都有自己的测试标准或规范。最早的 CTOD 测试规范是英国标准协会（B.S.I.）于 1972 年发表了一项草案 DD-19，后在 1979 正式颁布了 BS5762 规范。我国现行的 CTOD 断裂韧度的测试标准有 GB/T 2358－1994《金属材料裂纹尖端张开位移试验方法》等。结合该标准对 CTOD 的测试原理和测试方法分别作简要介绍。

① 测试的基本原理　对Ⅰ型裂纹试样进行加载时，随着载荷的增加，裂纹尖端张开位移 δ 也随之增大，当载荷增大到某一临界值时，裂纹起裂，这时对应的 δ 值即为所要测量的临界值 δ_c。试验仍采用三点弯曲试样，其形状与 K_{IC} 试样相同。试样的厚度 B 取被检测材料的厚度（即全厚度试样），试样的宽度 $W=2B$；试样长度 $s=4W$；裂纹长度 $a=0.45W \sim 0.55W$。

测试有试验步骤和所用的测试装置也同 K_{IC} 的测试基本相同，也同样需要材料试验机、载荷传感器和夹式引伸仪及 $x-y$ 记录仪等；在压断三点弯曲试样的过程中，试验机配备的载荷传感器向 $x-y$ 记录仪输入载荷 P 信号，由夹式引伸仪向 $x-y$ 记录仪输入裂纹起裂时切口间位移信号 V，从而获得 $P-V$ 曲线。不同之处在于试验结果的计算有所差异。

② δ_c 的计算公式　由于直接测量 δ_c 比较困难，一般都是采用间接测量法，即通过试样切中的张开位移 V 或加载点位移 Δ 来换算裂纹尖端张开位移 δ_c。测试的计算模型如图 5.4-10 所示，为转动因子模型。

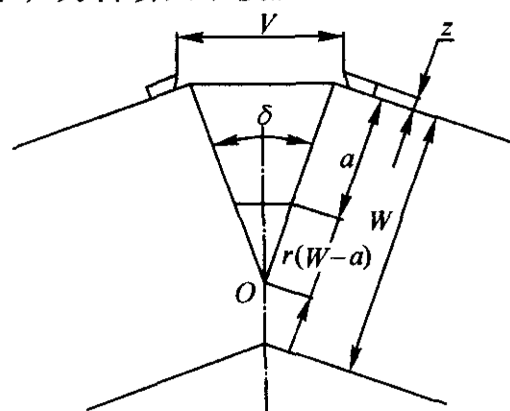


图 5.4-10　转动因子模型

本模型假设在加载过程中裂纹面以裂纹前端某一点 O 为中心张开，即试样两半以 O 点为铰链中心作刚性转动。图中 a 为裂纹长度，W 为试样宽度，L 为试样长度，z 为裂纹两侧刀口（装引伸计用）的厚度，V 为试样切口位移，Δ 为加载点位移，转动中心 O 到裂纹尖端的距离为 $r(W-a)$，其中 r 为转动因子，在加载过程中它是变化的。基于这一模型，有以下几种常用公式。

由图 5.4-10 中的相似性比例关系有：
$$\frac{\delta}{V}=\frac{r(W-a)}{r(W-a)+a+z} \tag{5.4-28}$$
$$\frac{\delta/2}{\Delta}=\frac{r(W-a)}{L/2} \tag{5.4-29}$$

由式（5.4-28）、式（5.4-29）可得 δ 的以下两个表达式：
$$\delta=\frac{r(W-a)}{r(W-a)+a+z}V \tag{5.4-30}$$
$$\delta=\frac{4r(W-a)}{L}\Delta \tag{5.4-31}$$

由式（5.4-30）、式（5.4-31）可见若能确定转动因子 r，则只要测量出开裂时试样的切口位移 V_c 或加载点位移 Δ_c，即可求出临界值 δ_c。

另由 Wells 表达式亦可得到 δ_c 的表达式。Wells 认为裂纹尖端张开位移实际上应当为减去裂纹表面弹性位移后的值，且转动因子为 0.45，即有
$$\delta=\frac{0.45(W-a)}{0.45(W-a)+a+z}(V-V') \quad (V \geqslant 2V')$$
$$\delta=\frac{0.45(W-a)}{0.45(W-a)+a+z}\frac{V^2}{4V'} \quad (V<2V') \tag{5.4-32}$$

式中，V' 为裂纹表面弹性位移引起的裂纹切口位移。
$$V'=rW\frac{\sigma_s}{E_1} \tag{5.4-33}$$

式中，r 为修正系数，与 a/W 有关；E_1 为弹性模量，对于平面应力状态，$E_1=E$，对于平面应变状态，$E_1=\frac{E}{1-\mu^2}$。

由式（5.4-30）可知只要测出临界的中间张开位移 V_c 即可求出临界张开位移 δ_c。

另外为了与线弹性体区别，可用 δ_p 和 V_p 分别表示理想塑性材料的 CTOD 和引伸计位移，用 δ_e 和 V_e 分别表示理想线弹性材料的 CTOD 和引伸计位移。CTOD 可以分解成 δ_e 和 δ_p 的叠加，即
$$\delta=\delta_e+\delta_p \tag{5.4-34}$$

式中，　$\delta_e=\frac{K_I^2}{2E_1\sigma_s}$，$\delta_p=\frac{r(W-a)}{r(W-a)+a+z}V_p$

最终可得
$$\delta_c=\frac{K_I^2(1-\mu^2)}{2E\sigma_s}+\frac{r(W-a)}{r(W-a)+a+z}V_p \tag{5.4-35}$$

式中，r 一般取 0.45。目前 δ_c 多采用式（5.4-35）来进行计算。

③ 临界点 V_c 的确定　由上面的公式可知，确定试样切口位移的临界值 V_c 对于确定 δ_c 来说是非常重要的环节。而 V_c 确定的关键是临界点的确定。由图 5.4-11 可见以下四种典型的 $P-V$ 曲线。

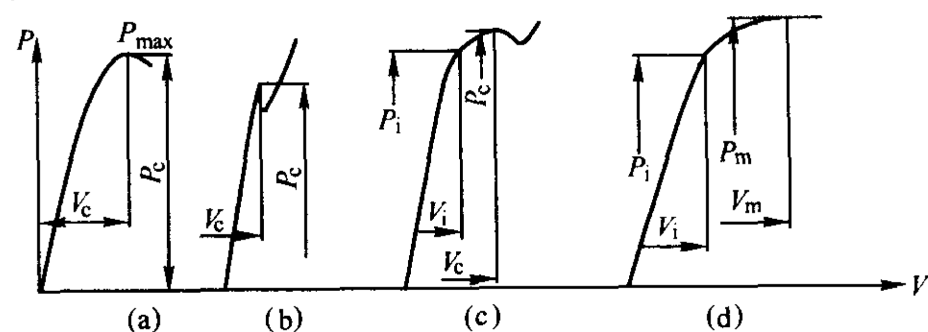


图 5.4-11　P－V 线中四种典型的曲线

对于比较脆的材料，其 $P-V$ 图如图 5.4-11a 所示，载荷达到某一值时，裂纹突然迸裂，没有明显的亚临界扩展阶段，此时 $P_c=P_{max}$；对于试样尺寸较大或材料韧性较差时，其 $P-V$ 图如图 5.4-11b 所示，在失稳后曲线仍有上升部分，此时取第一次迸发时所对应的最大载荷点即为裂纹起裂点；图 5.4-11c 所示的情况是载荷通过最高点后继续下降，不出现斜率突变的现象，也无迸发平台，此类情况不能直接判断，要通过其他方法（如电位法、声发射法等）来确定；而图 5.4-11d 所示情况为，随着载荷增大到最大值后基本保持恒定，而位移增加直至试样断裂。它也没有迸发平台，起裂点也同样无法直接确定，而要依靠其他方法来确定。

通常确定起裂点的方法有电位法、声发射法和阻力曲线法。

a）电位法　电位法又可分为直流电位法和交流电位法。直流电位法利用的是裂纹扩展会导致（$W-a$）区域中电流密度会突然增大的原理，因此开裂点对应的是电位曲线的拐点，如图 5.4-12 所示；交流电位法是根据高频电流的趋肤效应，裂纹尖端附近会产生电阻的变化，裂纹起裂时电阻会产生突变的现象来确定起裂点，通常最低点对应的是开裂点。如图 5.4-13 所示。

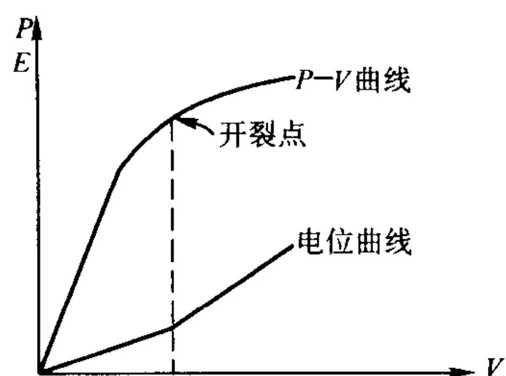


图 5.4-12　直流电位法

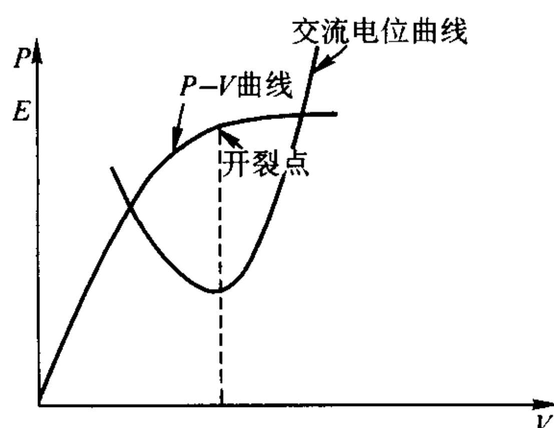


图 5.4-13　交流电位法

b）声发射法　材料在受载变形过程中，缺陷（裂纹）开裂部位会发射声波，这种现象称为声发射。试样在加载过程中，经历了弹性变形、裂纹尖端塑性变形、裂纹起裂并扩展、试样断裂等几个阶段，每一阶段声发射的特征是不同的，故可根据声发射信号的特征来判断起裂点。如图 5.4-14 所示。

c）阻力曲线法　可以采用单试样或多试样法作出材料的 $\delta-\Delta a$ 曲线，并将该曲线外推到 $\Delta a=0$ 处，则在 $\Delta a=0$ 处对应的值即为起裂点起裂时的 CTOD 值 δ_i，从而得到材料的临界 CTOD 值 $\delta_c=\delta_i$。如图 5.4-15 所示。

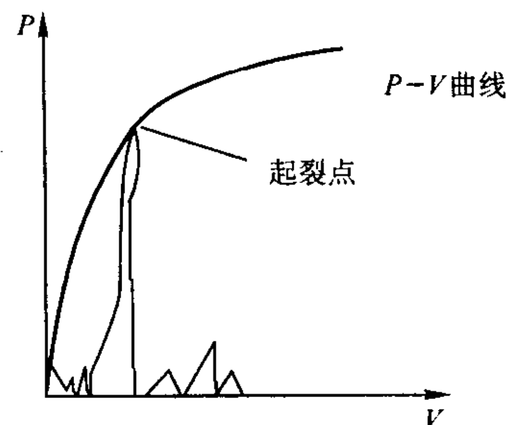


图 5.4-14　声发射法判断起裂点

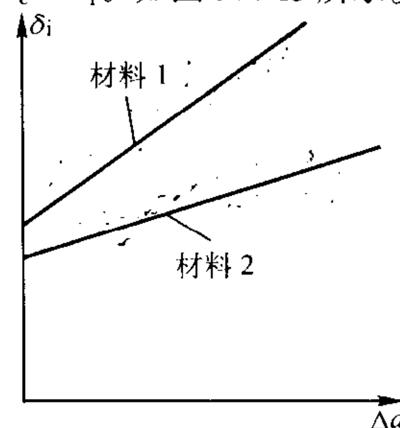


图 5.4-15　由阻力曲线法确定起裂点 δ_i

除了采用上述方法确定临界 CTOD 值外，亦可采用下列两种 CTOD 的临界值：

δ_c——稳定裂纹扩展量 $\Delta a<0.2$ mm 的脆性失稳断裂点或突进点所对应的 CTOD 值。

δ_i——δ_R 曲线上裂纹扩展量 $\Delta a<0.2$ mm 所对应的 CTOD 值。

3）J_{IC}的测试原理和测试方法　J 积分作为弹塑性参数近年来越来越受到重视。同 CTOD 一样目前国内外都有相关的规范或标准，J 积分测试可按 GB/T 2038－1991《金属材料延性断裂韧度 J_{IC}试验方法》进行。它的测试同 CTOD 测试过程中有很多相似之处。例如，它们在试样尺寸要求、多试样绘制阻力曲线法、物理监测开裂点的方法等方面都非常类似。尤其是建立了加载点位移 Δ 与裂纹切口位移 V 的关系后，一组试样可以同时完成 J 积分值和 CTOD 值的测试。在后面的介绍中可以看出这一点。

J_{IC}的测试方法主要有四种：多试样法，单试样法，阻力曲线法和金相剖面法。

① J 积分的测试原理　由第 3 章的内容可知，J 积分有两种定义，一个是回路积分定义，另一个是形变功率定义。在 J 积分的测试过程中一般采用 J 积分的形变功率定义，因为该定义具有比较明确的物理意义，便于进行直接测定。其定义式为

$$J=-\frac{1}{B}\left(\frac{\partial U}{\partial a}\right)_\Delta \tag{5.4-36}$$

式中，U 为外载对试样所作的形变功，即裂纹体总势能；B 为试样厚度；a 为裂纹长度；Δ 为加载点位移。

下面分别介绍上述四种测试方法的原理。

a）单试样法测试 J 积分的原理　目前国内通用的是单试样法，它又可分为计算法和解析法。这里对单试样法解析法的原理进行一下简单介绍。

对于深裂纹、短跨距的三点弯曲试样（$a/W\geq0.4$，$s/W\geq3\sim5$，s 为试样长度，W 为试样宽度），当其厚度和韧带尺寸的比值能够保证韧带的屈服属于平面应变状态时，可以得到如下关系式。

$$J=\frac{2U}{B\ (W-a)} \tag{5.4-37}$$

式中，U 可以通过加载曲线（$P-\Delta$ 曲线，如图 5.4-16 所示）所包围的面积求得。

Sumpter 和 Turner 把 J 积分值分为弹性部分 J_e 和塑性部分 J_p。弹性部分即为弹性应变能释率 G，在线弹性情况下，G 与 K 有下述关系式。

$$J_e=G=\frac{(1-\nu^2)}{E}K_I^2 \tag{5.4-38}$$

式中，K_I 可由 K_{IC}测试中获得的 K_{IC}值进行替代。由此可见，计算 J_e 的问题转变为先测试 K_{IC}值，再换算成 J_e 的问题。

而塑性部分 J_p 的计算公式如下：

$$J_p=\frac{2U}{B\ (W-a)}=\frac{P_L\Delta}{B\ (W-a)} \qquad (5.4\text{-}39)$$

式中，P_L 为全面屈服状态时的极限载荷。最终可得 J 积分的另一测试表达式为

$$J=J_e+J_p=\frac{(1-\nu^2)}{E}K_I^2+\frac{2U}{B\ (W-a)} \qquad (5.4\text{-}40)$$

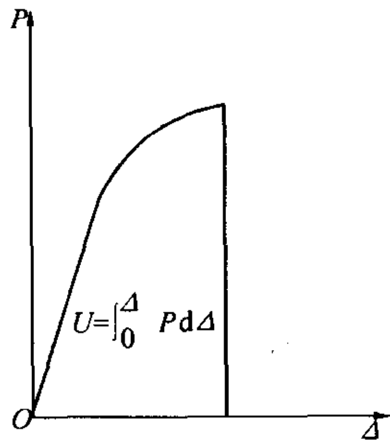


图 5.4-16　J 积分测试中的 $P-\Delta$ 曲线

当采用单试样法测定时，断裂韧度取三个同类试验结果的最低值。对 CTOD 和 J 积分试验，当最低值小于三个试验结果平均值的50%或最大值大于平均值的2倍时，应追加三个同样的试样并采用全部六个试样数据中的第二个最低值作为评定角数据；如果进行了更多的追加试验，则按表 5.4-2 的当量值选取评定用数据。

表 5.4-2　最低值的当量值

试验次数	当量值
3～5	最低值
6～10	第二个最低值
11～15	第三个最低值

b）多试样法测试 J 积分的原理　多试样法测试 J 积分的条件是：首先需要一组试样（至少三个以上），试样的尺寸相同，但裂纹长度不同，加载条件相同，同时记录下每个试样的载荷与加载位移（$P-\Delta$）的关系曲线；然后将每条 $P-\Delta$ 曲线按不同的位移Δ 分为若干份；计算各点位移曲线面积，如图 5.4-17 所示。由图 5.4-17 可见，各段面积为

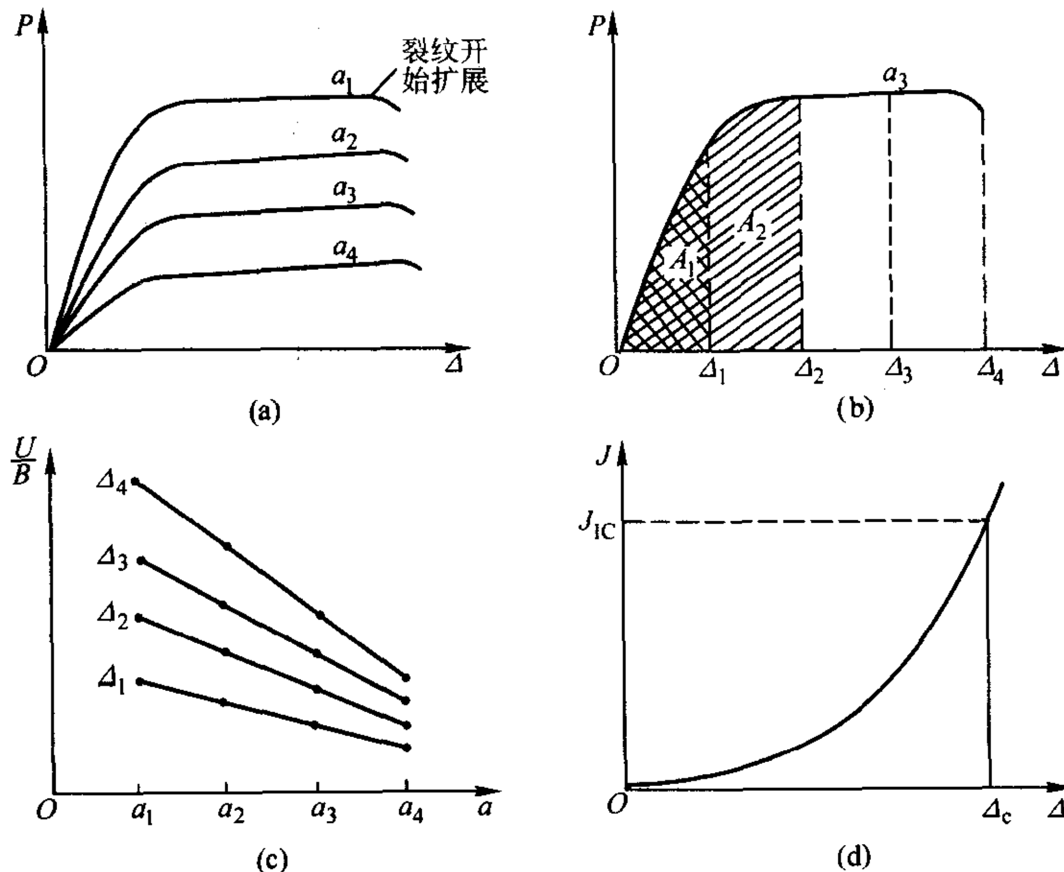


图 5.4-17　多试样法测试 J_{IC}示意图

$$A_i=\int_0^{\Delta_i}P\mathrm{d}\Delta \qquad (i=1,\ 2,\ \cdots\cdots) \quad (5.4\text{-}41)$$

即 A_i 为对应Δ_i 时形变功 U_i；继而可以获得一组 $U/B-a$ 的曲线，如图 5.4-17c 所示。这些曲线近似为直线，其斜率的负值为 $-\frac{1}{B}\left(\frac{\partial U}{\partial a}\right)=J$，即得到 J 积分值。由于该斜率与位移 Δ 有关，因此作 $J-\Delta$ 曲线，见图 5.4-17d。由此，只要确定出起裂点处的加载位移的 Δ_c，则可在 $J-\Delta$ 曲线上得出 J_{IC}的值。

采用多试样法测定时，一般应按最小二乘回归方法获得的数值减去1倍标准偏差所得的值作为评定用数据。

c）阻力曲线法测试 J 积分　与前面采用阻力曲线法求 δ_c 相似，若以起裂点处为临界点，可采用单试样或多试样法作出材料的测得的 $J_R-\Delta a$ 曲线，经外推，在 $\Delta a=0$ 处对应的值即为起裂点起裂时的 J_{IC}的值，如图 5.4-18 所示。

材料断裂韧度 J_{IC} 只是 $J_R-\Delta a$ 曲线上的一个点，根据需要还有以下几种常用的 J 积分临界值的定义：

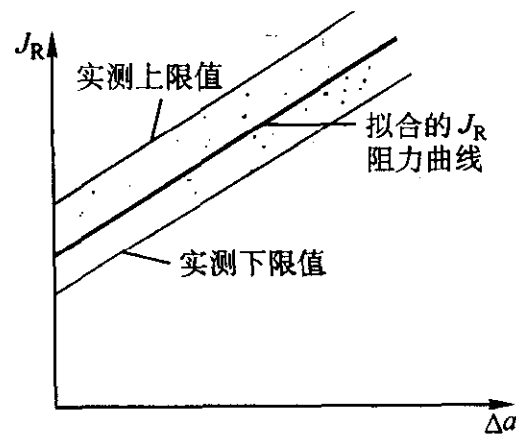


图 5.4-18　J_R 阻力曲线

表观起裂韧性 J_i，为裂纹起裂时材料的断裂阻力曲线 J_R 值，即为 $\Delta a=0$ 时的值；

条件起裂韧性 $J_{0.05}$，表观起裂韧性为 $\Delta a=0.05$ 时的值；

条件起裂韧性 $J_{0.2}$，为 $\Delta a=0.2$ 时的 J_R 值。

d）金相剖面法测试 J_{IC}积分　在不能用标准方法进行断

裂韧度测试时，容许用金相剖面法测定获得CTOD起裂值δ_{is}和估算J起裂值J_{is}。金相剖面法是建立在裂纹钝化原理基础上的。对于金属材料，尤其是延性较好的金属材料来说，在加载过程中由于裂纹尖端的应力集中，使裂纹尖端的材料发生塑性滑移，从而引起裂纹尖端的钝化、张开和延伸，并形成伸张区。伸张区端近似呈半圆形，伸张区高度D即为$\delta/2$，而伸张区宽度与其高度近似相等。当载荷增大时，D也随之增大，最后达到某一值时，裂纹开始扩展，此时的COD值即为δ_{is}。在获得δ_{is}后，对$s=4W$，$a/W=0.4\sim0.7$的三点弯曲试样进行测试以确定其$P-\Delta$曲线（如图5.4-18所示）；之后按GB/T 2038－1991《金属材料延性断裂韧度J_{IC}试验方法》测得$\Delta\delta_p$，并计算

$$(\Delta_p)_F-(\Delta_p)_C=\Delta\delta_p\frac{1}{0.44\ (1-a/W)}\qquad(5.4\text{-}42)$$

在$P-\Delta$曲线上确定起裂点C，最终确定J_{is}。

② 临界值的确定　临界点的确定对J_{IC}的测定有直接的关系，目前对临界点有几种看法，而试验结果表明把起裂点处为临界点是比较合理的。而在$P-\Delta$曲线上确定起裂点的问题在前面已经进行过介绍，如电位法、声发射法等，这里不再赘述。

③ 测试J_{IC}的过程　测试J_{IC}所采用的试样、测试过程、应满足的条件以及计算方法等都同CTOD相似，可参考前文有关内容。

④ J_{IC}、K_{IC}和δ_c的关系　J_{IC}、K_{IC}和δ_c都是表示材料抗断裂性能力的性能参数，在线弹性范围内，三者存在如下关系：

$$\delta_c=\frac{J_{IC}}{\beta\sigma_s}\qquad(5.4\text{-}43)$$

$$K_{IC}=\left(\frac{E\sigma_s\delta_c}{1-\mu^2}\right)^{1/2}=\left[\frac{EJ_{IC}}{\beta\ (1-\mu^2)}\right]^{1/2}\qquad(5.4\text{-}44)$$

式中，β为COD降低系数，取1～2。

如果被测材料较脆，测定J_{IC}值有困难，可按GB/T 4161－1984《金属材料平面应变断裂韧度K_{IC}试验方法》测定平面应变断裂韧度K_{IC}用于评定。

若不能获得实测断裂韧度数据的情况下，对有使用经验的常用钢材，可以从有关手册或资料中选取代用数据，但应对数据的选取作出有足够依据的说明，而且要保证所选用的评定数据使评定结果偏于保守和安全。且在选取代用的断裂韧度数据时，应了解被评定设备的运行环境条件（设备所在的外部环境介质条件、内部介质和操作温度等）。如果被评定的是表面缺陷，则应考虑环境介质的影响；对埋藏缺陷，一般可不考虑这种影响，不论是表面缺陷还是埋藏缺陷，都应考虑温度的影响。对有环境介质或温度影响的容器，还应考虑服役期间材料性能的退化效应。另外，对于被评定部位材料的化学成分、冶金和工艺状态等，在选取代用或参考数据时，均必须予以充分注意。对非均质的焊接接头部位，还应考虑局部显微组织对断裂韧度的影响。

当安全评定所需参数准备完整后，即可以开始对构件进行评定。如上所述，我国安全评定对不同的缺陷类型有不同的方法。这其中各国对压力容器和压力管道的安全评定研究最早，技术也最成熟，因此下面就以压力容器和压力管道为例对不同的安全评定方法进行介绍。对于其他机械构件也可以按照类似的思路进行安全评定，但要根据具体情况做相应的改动。

2.4.4 压力容器平面缺陷的确定性简化评定

参照我国“八五”期间有关锅炉压力容器评定方面的成果，首先介绍压力容器中平面缺陷的简化评定方法。这里主要介绍采用以起裂准则为基础的简化评定方法（也可称为一级评定方法）。

一级简化评定方法继承了我国CVDA－84规范的精华，并采用了BSIPD6493－91的1级筛选评定图的形式。英国PD6493-80与我国CVDA-84都是采用COD设计曲线，但PD6493-91基本上采用了R6技术，这为我们将COD曲线技术转变到失效评定图技术提供了借鉴。因此我国决定继承CVDA-84的精华采用失效评图的形式，作为标准一级平面缺陷的简化评定方法。它比PD6493-91中的筛选评定方法更为先进和偏于安全。

我国的简化评定方法在如下几方面与PD6493-91一致：采用评定点纵坐标为$\sqrt{\delta/\delta_c}$并以符号$\sqrt{\delta_r}$表示的简化失效评定图；考虑安全系数为2，即$\delta=\delta_c/2$为临界条件，$\sqrt{\delta_r}=\sqrt{\delta/\delta_c}=0.7$的呈水平状的断裂失效评定曲线；横坐标用$S_r$，并限制$S_r=0.8$，即$S_r=0.8$的一条截止线；所以评定点以（$S_r$，$\sqrt{\delta_c}$）为参量时的简化失效评定图应该是矩形的。

但是PD6493-91的δ是按Burdikin的COD设计曲线计算的，而我国的标准是按CVDA-84给出的COD设计曲线计算的。理论分析和实践证明，CVDA-84的设计曲线要优于PD6493的设计曲线，主要差别反映在施加应变与材料屈服应变之比在0.8～1.2的区间时，宽板试验表明在此区间的断裂点常落在Burdikin设计曲线之上，但在CVDA曲线之下，说明CVDA-84的COD设计曲线比PD6493-91的COD设计曲线更符合实际，技术更为先进，评定更为安全。下面将具体介绍平面缺陷的简化评定方法。

（1）评定方法

平面缺陷的简化评定方法主要采用简化失效评定图进行评定。简化失效评定图如图5.4-19所示，由纵坐标$\sqrt{\delta_r}$、横坐标S_r以及$\sqrt{\delta_r}$等于0.7的水平线和S_r等于0.8的垂直线所围成的矩形为安全区，该区域之外为非安全区。通过计算所得的（可按本节给定的方法确定）评定点若位于失效评定图的安全区内，则为安全或可以接受；否则，应为不能保证安全或不可接受。

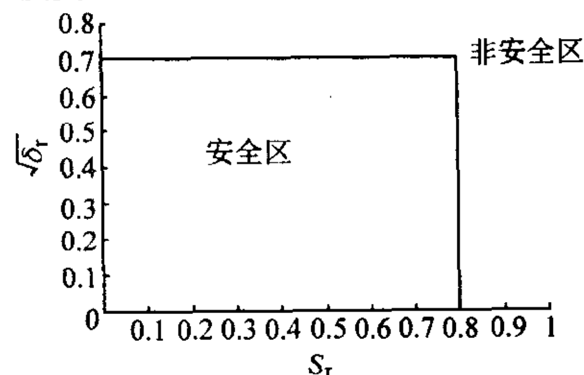


图5.4-19　平面缺陷确定性简化评定的失效评定图

（2）评定程序

平面缺陷的简化评定通常可按下列步骤进行：

1）确定缺陷表征和等效裂纹尺寸；

2）确定应力值；

3）确定材料的性能数据；

4）δ及$\sqrt{\delta_r}$的计算；

5）S_r的计算；

6）进行安全性评价。

其评定程序如图5.4-20所示。

（3）确定性简化评定所需的基本数据和参数的确定

在进行简化评定前首先要计算或选取与评定有关的三大类数据：缺陷的表征尺寸；计算所需的应力值；材料性能参数。下面将分别介绍这三大类数据的确定方法。

1）缺陷表征和等效裂纹尺寸$\bar{a}$的确定　根据缺陷的实

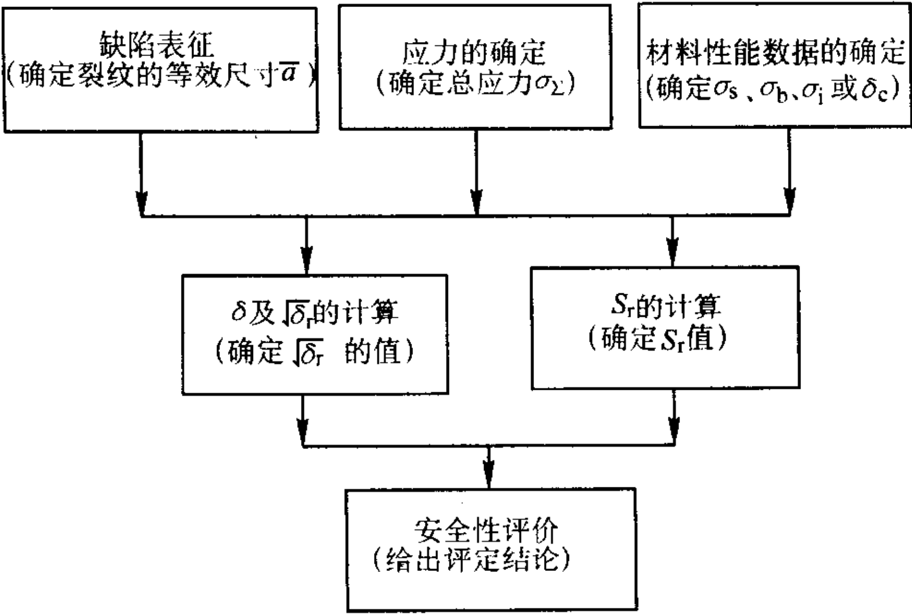


图5.4-20 平面缺陷简化评定程序示意图

际位置、形状和尺寸，按本章2.4.1节的说明（或相关规定）对缺陷进行规则化，从而获得上述表征裂纹计算尺寸a、c，然后针对不同的情况按下列方法计算等效裂纹尺寸$\bar{a}$；$\bar{a}$值也可从相关规定中的表格中查取。

对穿透裂纹（长为$2a$），$\bar{a}=a$；

对埋藏裂纹（长为$2c$、高为$2a$），$\bar{a}=\Omega a$；

$$\Omega=\frac{\left(1.01-0.37\frac{a}{c}\right)^2}{\left[1-\left(\frac{2a/B}{1-2e/B}\right)^{1.8}\left(1-0.4\frac{a}{c}-\left(\frac{e}{B}\right)^2\right)\right]^{1.08}}$$

式中，

$$e=\frac{B}{2}-(a+p_1)$$

上式的适用范围为：$a/B\leqslant0.45$，$a/c\leqslant1.0$。

而对表面裂纹（其长为$2c$、深为a），则$\bar{a}=\left(\frac{F_1}{\varphi}\right)^2a$；

式中

$$F_1=\begin{cases}1.13-0.09\frac{a}{c}+\left(-0.54+\frac{0.89}{0.2+a/c}\right)\left(\frac{a}{B}\right)^2+\left[0.5-\frac{1}{0.65+a/c}+14\left(1-\frac{a}{c}\right)^{24}\right]\left(\frac{a}{B}\right)^4 & 当\frac{a}{c}>0\\ 1.12-0.23\frac{a}{B}+10.55\left(\frac{a}{B}\right)^2-21.71\left(\frac{a}{B}\right)^3+30.38\left(\frac{a}{B}\right)^4 & 当\frac{a}{c}=0\end{cases}$$

$$\varphi=\left[1+1.464\left(\frac{a}{c}\right)^{1.65}\right]^{1/2}$$

上式的适用范围为：$a/B\leqslant0.8$，$a/c\leqslant1.0$。

上式中各参数涵义参见图5.4-2～图5.4-5。

2）总当量应力的σ_Σ确定　简化评定计算所需总当量应力σ_Σ可按式（5.4-45）估算，并保守地假设总当量应力均匀分布在主应力平面上。

$$\sigma_\Sigma=\sigma_{\Sigma1}+\sigma_{\Sigma2}+\sigma_{\Sigma3} \tag{5.4-45}$$

式中，$\sigma_{\Sigma1}=K_tP_m$，$\sigma_{\Sigma2}=X_bP_b$，$\sigma_{\Sigma3}=X_rQ$，此外，K_t为由焊缝形状引起的应力集中系数，X_b为弯曲应力折合系数，X_r为焊接残余应力折合系数，Q为被评定缺陷部位热应力最大值与焊接残余应力最大值σ_R^{max}之代数和。

下面给出了几种常见焊接接头结构K_t取值，表5.4-3给出了X_b和X_r的取值。

表5.4-3　X_b值和X_r值的选取

裂纹种类		X_b	X_r		
			裂纹平行熔合线的情况	裂纹垂直熔合线的情况	填角焊缝裂纹的情况
埋藏裂纹		0.25	0.2	0.6	0.6
穿透裂纹		0.5	0.2	0.6	0.6
表面裂纹	弯曲的拉伸侧	0.75	0.4～0.6	0.6	0.6
	弯曲的压缩侧	0			

① 对接焊接接头结构

$$焊趾处裂纹,K_t=\begin{cases}0.5 & \eta/B\leqslant0.5\\ 1.0 & \eta/B>0.2\\ 0.15\sim0.2 & \eta/B=0.15\sim0.2,可按线性内插求取\\ 1.0 & 无焊缝增高时\end{cases}$$

（其含缺陷结构示意图及各参数的含义参见图5.4-21）

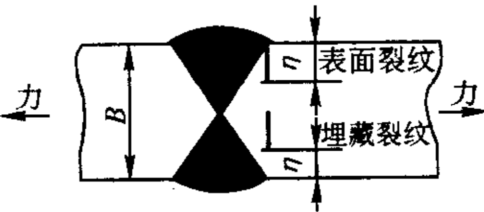


图5.4-21 焊趾处裂纹

有角变形及错边量：

$$K_t=\Gamma\left[1+\frac{3(w+e_1)}{\beta B}\right]$$

$$\beta=\begin{cases}1 & \eta/B\leqslant0.5的表面裂纹\\ 2 & \eta/B>0.5的表面裂纹和埋藏裂纹\end{cases}$$

$$\Gamma=\begin{cases}1.5 & \eta/B\leqslant0.15\\ 1 & \eta/B>0.2\\ 可按线性内插求取 & \eta/B=0.15\sim0.2\end{cases}$$

（其含缺陷结构示意图及各参数的含义参见图5.4-22）

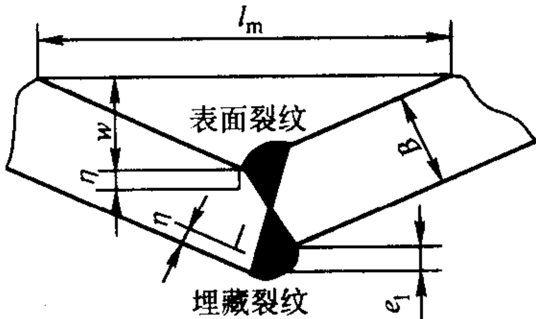


图5.4-22 有角变形及错边量

② 接管处内拐角

球壳及球形封头接管，$K_t=2.0(1+2\sin^2\theta)$

圆柱壳接管，$K_t=3.1(1+2\sin^2\theta)$，用于θ角平面与容器横截面平行时$K_t=3.1\left[1+(\tan\theta)^{4/3}\right]$，用于$\theta$角平面与容器纵截面平行时。

（其含缺陷结构示意及各参数的含义参见图5.4-23）

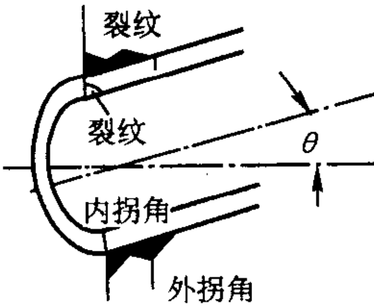


图5.4-23 接管处内拐角含缺陷结构示意图

（用于结构尺寸满足分析设计规范的规定时）

焊接压力容器不可避免地存在焊接残余应力及热应力等

二次应力，由于二次应力是自平衡应力，又是有自限性的，在外力作用下可能局部塑性松弛或再分布，对断裂的影响有时很大有时很小。存在二次应力时，Rice J 积分不再守恒，分析技术难度较大。代表当今国际水平的是：用通用失效评定曲线及 Ainsworth 提出的 ρ 因子来处理二次应力的方法。但是，R6 编者在背景材料中亦不得不承认这个方法用于小裂纹位于高值残余拉伸应力区是不安全的，其他情况又过分保守，引起不少争议和新建议。"八五"攻关中建立了有二次应力存在时守恒的修正 J 积分及其计算程序，揭示了 R6 的 ρ 因子存在问题的症结所在，是其推导过程中采用的一个假设是错误的，从而导出了一套比 R6 法更先进的 ρ 因子，理论和试验都证明要比 R6 的更符合客观实际，具有国际先进水平。此外，用在有二次应力存在下的严格弹塑性断裂力学理论分析计算证明，简化评定中处理二次应力的系数 X_r 不是常数，且变化很大。从既要保持原有评定方法、又要保证安全的角度出发，我国的最新标准对 CVDA-84 中的原二次应力系数值 X_r 的具体规定作了局部调整，计算时可参考此方面的规定。

3）材料性能数据的确定　可按本章说明或相关标准、规范等确定材料的屈服点 σ_s、抗拉强度 σ_b 及断裂韧度 δ_c 值。δ_c 按实际情况可取 δ_i 的值或 δ_{is} 的值（也可保守地取 $\delta_{0.05}$ 的值），并应将所得的材料断裂韧度 δ_c 除以 1.2（值 1.2 为断裂韧度安全系数）后的值用于下面计算中所需的 δ_c 值的计算。

① δ 及 $\sqrt{\delta_r}$ 的计算

$$\delta=\begin{cases}\pi\bar{a}\sigma_s\ (\sigma_\Sigma/\sigma_s)^2M_g^2/E & \text{当 } \sigma_\Sigma<\sigma_s \text{ 时}\\ 0.5\pi\bar{a}\sigma_s\ (\sigma_\Sigma/\sigma_s+1)\ M_g^2/E & \text{当 } \sigma_\Sigma\geqslant\sigma_s\geqslant\sigma_{\Sigma1}+\sigma_{\Sigma2} \text{ 时}\end{cases} \tag{5.4-46}$$

式中，M_g 为鼓胀效应系数，其值在不同的情况下可按下式计算。

$$M_g^2=\begin{cases}1+1.61\bar{a}^2/RB & \text{筒壳轴向裂纹}\\ 1+0.32\bar{a}^2/RB & \text{筒壳环向裂纹}\\ 1+1.93\bar{a}^2/RB & \text{球壳裂纹}\end{cases}$$

$$\sqrt{\delta_r}=\begin{cases}\sqrt{\delta/\delta_c} & \text{(a)}\\ 1.2\ \sqrt{\delta/\delta_c} & \text{(b)}\end{cases} \tag{5.4-47}$$

式（5.4-47）中（a）为对于单裂纹或复合后的单裂纹，以及其他不满足上述考虑干涉效应的裂纹群的情况和 $\sqrt{\delta_r}$ 的计算方法；式（5.4-47）（b）为对于需要考虑干涉效应的裂纹群的情况和 $\sqrt{\delta_r}$ 的计算方法。

② S_r 的计算

$$S_r=\frac{L_r}{L_r^{\max}} \tag{5.4-48}$$

式中，L_r 由 P_m 及 P_b 的值按下面介绍的方法进行计算。在这里不需要考虑载荷应力的安全系数；$L_r^{\max}$ 的值取 1.20 及 $\frac{\sigma_s+\sigma_b}{2\sigma_s}$ 两者中的较小值，即有 $L_r^{\max}=\min\left(1.20,\ \frac{\sigma_s+\sigma_b}{2\sigma_s}\right)$。

③ 载荷比 L_r 参量的计算　这里给出了含缺陷结构的塑性极限载荷解、塑性屈服载荷解和 L_r 基于极限分析方法和净截面屈服准则局部解的计算式。这些计算式不仅适用于平面缺陷的简化评定，也适于平面缺陷的常规评定。

a）L_r 计算中所需参数的确定　在 L_r 的计算中需要确定两类参数，一类是应力参数，另一类是缺陷尺寸参数。其中在确定应力参数时，只考虑一次应力，而不考虑二次应力。在下面计算时需要的应力（P_m 及 P_b）可按本节的方法计算；而在确定缺陷尺寸（a 及 c）时可按本章 2.4.1 节介绍的方法进行确定。

b）下面给出几种不同情况下的 L_r 计算方法　对于平板上半椭圆表面裂纹，假设裂纹 $a\times2c$，板厚 B，板宽 $2W$ 时

$$L_r=\frac{P_b+\sqrt{P_b^2+9\ (1-\xi)^2P_m^2}}{3\ (1-\xi)^2\sigma_s} \tag{5.4-49}$$

式中，$\xi=\frac{ac}{B\ (c+B)}$。

该公式适用范围为 $a/B\leqslant0.8$ 且 $2\ (c+B)\ /W\leqslant1.0$ 的平板及相当尺寸范围界限的压力容器。

对于平板上椭圆形埋藏裂纹，假设裂纹 $2a\times2c$，板厚 B，板宽 $2W$，有

$$L_r=\frac{(3\zeta P_m+P_b)\ +\sqrt{(3\zeta P_m+P_b)^2+9\left[\ (1-\zeta)^2+4\zeta\gamma\right]\ P_m^2}}{3\left[\ (1-\zeta)^2+4\zeta\gamma\right]\ \sigma_s} \tag{5.4-50}$$

式中，$\zeta=\frac{2ac}{B\ (c+B)}$；$\gamma=\frac{P_1}{B}$；$P_1$ 为缺陷距板表面最近处距离（指被评定裂纹尖端位置）。

该公式适用于 $2a/B\leqslant0.8$ 且 $(c+B)\ /W\leqslant1.0$ 平板及相当尺寸范围界限的压力容器。

对于内压球壳上长 $2a$ 穿透裂纹，当板厚为 B，内径为 R_i 时，有

$$L_r=\left(\frac{P_m}{\sigma_s}\right)\frac{1+\sqrt{1+8a^2/\left[R_iB\cos^2\ (a/R_i)\right]}}{2} \tag{5.4-51}$$

上式适用于 $B/R_i\leqslant0.1$ 的情况。

对于平板中长 $2a$ 穿透裂纹，假设零件板厚为 B，板宽为 $2W$，有

$$L_r=\frac{P_b+\sqrt{P_b^2+9P_m^2}}{3\ (1-2a/W)\ \sigma} \tag{5.4-52}$$

上式适用于 $a<<W$ 的情况。

对于内压圆筒体上长 $2a$ 纵向穿透裂纹，若板厚为 B，内径为 R_i，有

$$L_r=\frac{P_m}{\sigma_s}\sqrt{1+1.6a^2/\ (R_iB)} \tag{5.4-53}$$

上式适用于 $B/R_i\leqslant0.1$ 的情况。

对于容器接管拐角裂纹，假设存在 1/4 圆形裂纹，则有

$$L_r=\frac{P_b+\sqrt{P_b^2+9\ (1-\zeta)^2P_m^2}}{3\ (1-\zeta)^2\sigma_s} \tag{5.4-54}$$

式中，$\zeta=a/B_{\min}$；$B_{\min}$ 为接管内拐角至外拐角间的距离；P_m、P_b 为垂直裂纹平面方向在 $B_{\min}$ 厚度上分布的应力。

上式适用于 $a/B_{\min}\leqslant0.8$ 的情况。

对于内压圆筒体上表面裂纹，假设存在 $a\times2c$ 的裂纹，且其板厚为 B，内径为 R_i，有

$$L_r=\left(\frac{1.2P_m}{\sigma_s}\right)\frac{1-a/\ (BM_R)}{1-a/B} \tag{5.4-55}$$

式中，$M_R=\sqrt{1+1.6\left[c^2/\ (R_iB)\right]}$。

而对于仅受内压 p 的容器接管拐角裂纹，假设存在的拐角裂纹为 1/4 圆形裂纹，则有

$$L_r=\left(\frac{P}{\sigma_s}\right)\frac{A_p}{A_s-\pi a^24} \tag{5.4-56}$$

式中，A_p 为内压 p 的作用面积；A_s 为接管承载面积，其定义可见图 5.4-24、图 5.4-25。图中：

$$l_s=0.4\ \sqrt{R_{in}B_n}$$

$$l_m=\begin{cases}0.4\ \sqrt{R_iB}\\ 2R_{in}/3\end{cases}\quad \text{两者中较小者，适用于圆筒形容器}$$

$$l_m=\begin{cases}0.4\ \sqrt{R_iB}\\ R_{in}\end{cases}\quad \text{两者中较小者，适用于球形容器}$$

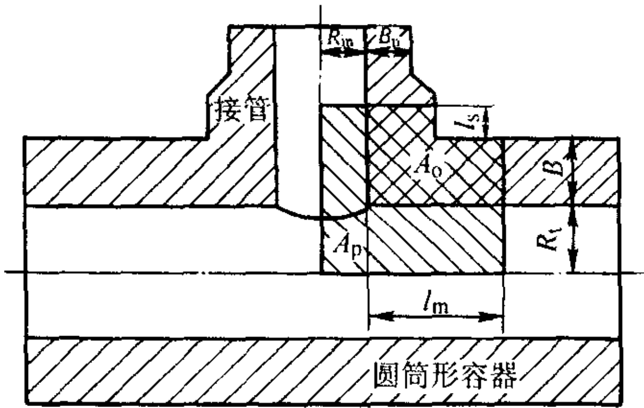


图 5.4-24　圆筒形容器接管拐角裂纹
（在仅受内压作用计算 L_r 时需要采用的受压面积 A_p 和承载面积 A_s 的定义）

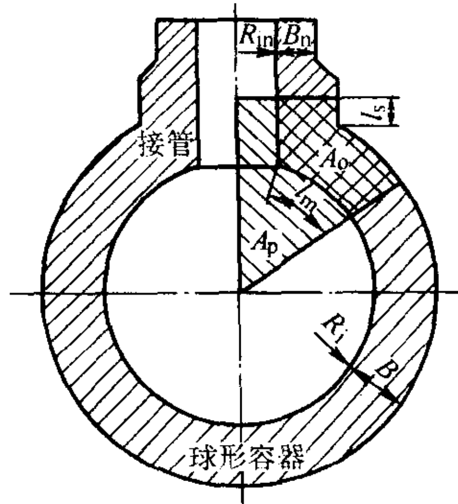


图 5.4-25　球形容器接管拐角裂纹
（在仅受内压作用计算 L_r 时需要采用的受压面积 A_p 和承载面积 A_s 的定义）

式（5.4-56）适用范围为 $a/\sqrt{4A_s/\pi}\leqslant 0.8$ 的情况。

(4) 安全性的评价

将上一部分计算得到的评定点（S_r，$\sqrt{\delta_r}$）绘在图 5.4-19 中，如果评定点落在安全区内，则评定结论为安全或可以接受；否则，认为不能保证安全或不可接受。

(5) 最大容许等效裂纹尺寸的确定

除了上述方法外，对平面缺陷进行简化评定也可以采用给出最大容许等效裂纹尺寸 $\bar{a}_m$ 的评定方法，最大容许等效裂纹尺寸 $\bar{a}_m$ 按下式计算：

$$\bar{a}_m=\begin{cases}\dfrac{E\delta_c}{2\pi\sigma_s(\sigma_\Sigma/\sigma_s)^2M_g^2} & \text{当 } \sigma_\Sigma<\sigma_s \text{ 时}\\[2ex] \dfrac{E\delta_c}{\pi\sigma_s(\sigma_\Sigma/\sigma_s+1)M_g^2} & \text{当 } \sigma_\Sigma\geqslant\sigma_s\geqslant(\sigma_{\Sigma1}\sigma_{\Sigma2})\text{时}\end{cases}\tag{5.4-57}$$

一般认为，当 $\bar{a}\leqslant\bar{a}_m$ 且 $S_r\leqslant 0.8$ 时，该缺陷是容许的。

2.4.5　压力容器平面缺陷的确定性常规评定

常规定是平面缺陷的2级评定。从科学性、成熟性、工程实用性综合考虑，在常规评定中建议采用 R6 的通用失效评定图来进行防止起裂的评定。主要有以下几个理由：第一，20世纪90年代世界各国发表的所有规范均采用 R6 的通用失效评定图，说明它是世界公认的工程评定的适用方法。采用技术上较成熟的 R6 通用失效评定图，既可与国际规范接轨，又便于与国外规范进行对比，有利于对一些细节作出评价。第二，我国"八五"科技攻关 85-924-02-02 专题制作了我国常用压力容器用钢的母材、焊缝、各种试板、容器、焊接接头、各种穿透裂纹和表面裂纹、高应变区裂纹、应变时效或温度影响下的严格的 J 积分失效评定曲线（即 R6 的选择 3）800 多条，与 R6 通用失效评定曲线对比说明，在绝大多数情况下，R6 通用失效评定曲线是偏于安全的，虽然在个别情况下严格失效评定曲线会比 R6 通用失效评定曲线要低一点，但所有曲线的下包络线在 $L_r<1$ 范围内与通用失效评定曲线的误差（以载荷计）均不大于 10%，这就为采用 R6 通用失效评定图作为常规评定提供了符合我国国情的实验证据，也为在平面缺陷的常规评定中采用分安全系数提供了重要的依据。下面将具体介绍平面缺陷的常规评定方法。

(1) 确定性常规评定方法

平面缺陷的常规评定采用通用失效评定图的方法进行。该失效图如图 5.4-26 所示。

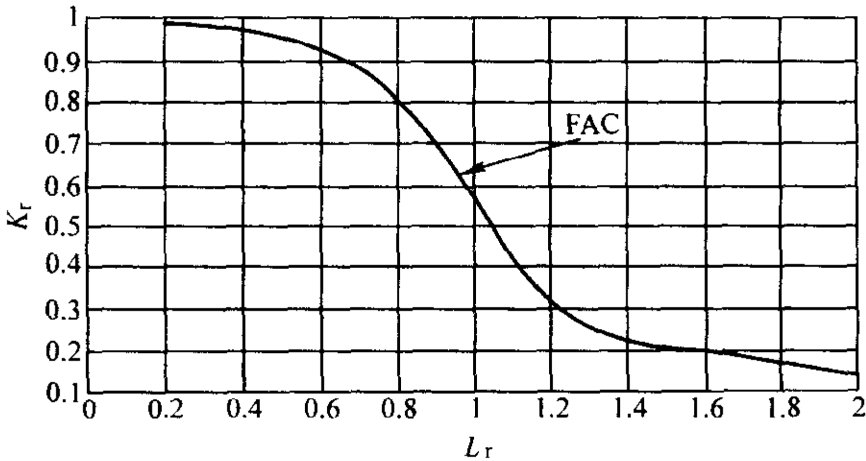


图 5.4-26　通用失效评定图

图 5.4-26 中失效评定曲线（FAC）的方程为

$$K_r=(1-0.14L_r^2)(0.3+0.7e^{-0.65L_r^6})\tag{5.4-58}$$

且通用失效评定图垂直线的方程为

$$L_r=L_r^{\max}$$

式中，$L_r^{\max}$ 的值取决于材料特性，即不同的材料 $L_r^{\max}$ 不同，其确定方法如下：

对奥氏体不锈钢，$L_r^{\max}=1.8$；

对无屈服平台的低碳钢及奥氏体不锈钢焊缝，$L_r^{\max}=1.25$；

对无屈服平台的低合金钢及其焊缝，$L_r^{\max}=1.15$；

对于具有长屈服平台的材料，一般情况下，$L_r^{\max}=1.0$；当材料温度不高于200℃时，$L_r^{\max}$ 可根据 K_r 值及材料屈服强度级别，由表 5.4-4 来确定；

表 5.4-4　温度不高于 200℃时的长屈服平台材料的 $L_r^{\max}$ 值

$L_r^{\max}$ \ K_r	材料屈服强度	
	235 MPa≤σ_s<350 MPa	σ_s>350 MPa
1.25	$K_r\leqslant 0.10$	$K_r\leqslant 0.13$
1.20	$0.10\leqslant K_r<0.12$	$0.13\leqslant K_r<0.15$
1.15	$0.12\leqslant K_r<0.20$	$0.15\leqslant K_r<0.26$
1.00	$K_r\geqslant 0.20$	$K_r\geqslant 0.26$

对于不能按钢材类别确定 $L_r^{\max}$ 的材料，可按式（5.4-59）计算 $L_r^{\max}$ 的值。

$$L_r^{\max}=\bar{\sigma}/\sigma_s=0.5(\sigma_b+\sigma_s)/\sigma_s\tag{5.4-59}$$

图 5.4-26 中，由 FAC 曲线、$L_r=L_r^{\max}$ 直线和两直角坐标轴所围成的区域之内为安全区，该区域之外为非安全区。

在评定点的计算时应注意，相关的输入参量应首先应根据本章表 5.4-1 的规定和现场的实际情况选取相应的分安全系数，然后按下文介绍的常规评定程序进行评定。

(2) 确定性常规评定程序

平面缺陷的常规评定可按下列步骤进行：

1）根据本章 2.4.1 节的内容和规定表征缺陷的尺寸；

2）根据本章 2.4.2 节的内容和规定确定应力的大小；

3）根据本章 2.4.3 节的内容和规定确定材料的性能数据；

4）应力强度因子 K_I^P 和 K_I^S 的计算；

5）根据本章 2.4.3 的内容进行 K_r 的计算；

6）根据本章 2.4.3 的内容进行 L_r 的计算；

7）根据图 5.4-26 通用失效评定图进行安全性评价。

具体常规安全评定程序如图 5.4-27 所示。

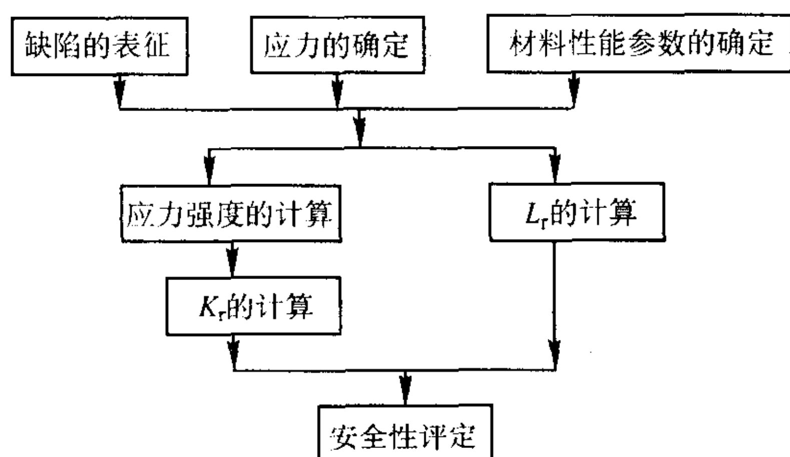


图 5.4-27 平面缺陷常规评定程序示意图

(3) 平面缺陷确定性常规评定所需基本数据的确定

同平面缺陷的简化评定一样，在进行评定前首先要计算或选取与评定有关的三大类数据参数。下面将分别介绍这三大类数据的确定方法。

1）缺陷的表征　如本章 2.1 节所述，目前缺陷的检测有许多方法，如对表面裂纹可以采用简单易行的磁粉法和着色法，对埋藏性缺陷可采用超声波检测仪、X 射线检测仪等进行检测；对近表面缺陷可采用涡流检测仪等。对经以上检测手段检出的缺陷，根据实际位置、形状和尺寸按本章 2.4.1 节的方法进行缺陷的规则化，并得到相应的表征裂纹尺寸 a、c、p_1 等。该表征裂纹尺寸应乘以表 5.4-1 规定的表征裂纹分安全系数后作为计算用的表征裂纹尺寸 a、c 值。

2）一次及二次应力的确定　常规评定中需考虑一次应力和二次应力的应力分量 P_m、P_b、Q_m、Q_b。按本章 2.4.2 节中的计算方法，分别确定各载荷下一次应力、二次应力及各应力分量（亦可以采用直接测量法，进行现场实际测量），并分别计算各类应力分量的代数和，然后乘以表 2.4-1 所规定的应力分安全系数，由此所得到的应力值将用作常规评定计算的应力 P_m、P_b、Q_m、Q_b。

3）材料性能参数的确定　材料性能参数的合理选取对安全评定具有重要意义，因此在评定前应作很多细致的工作以便合理的确定评定所需的材料性能参数。一般可按本章 2.4.3 节的说明或相关标准、规范的规定，确定评定工况下材料的屈服点 σ_s 和抗拉强度 σ_b 及 J 积分断裂韧度 J_{IC}。J_{IC} 值按实际情况可取 J_i 值或 J_{is}值，也可保守地取 $J_{0.05}$的值。

计算 K_r 所需的材料断裂韧度 K_C 可以由测得的 J 积分断裂韧度 J_{IC}按式（5.4-60）求得。

$$K_C = \sqrt{EJ_{IC}/(1-\nu^2)} \quad (5.4\text{-}60)$$

在不能直接得到 J_{IC}值的情况下，可直接测量材料的平面应变断裂韧度 K_{IC}，此时计算 K_r 所需的 K_C 值可用 K_{IC}值代替；也可采用 CTOD 韧度 δ_c 值，按式（5.4-61）估算 K_C 的下限值。

$$K_C = \sqrt{1.5\sigma_s\delta_c E/(1-\nu^2)} \quad (5.4\text{-}61)$$

4）K_I^P 和 K_I^S 的计算　K_I^P、K_I^S 为一次应力 P_m、P_b 和二次应力 Q_m、Q_b 作用下的应力强度因子，该值可按下述方法进行计算。该方法既适用于平面缺陷的断裂常规评定，也适用于疲劳评定（尤其是对一些典型的结构和缺陷情况）。计算中所需的表征裂纹尺寸和应力，分别按前面的说明来确定。

① 计算 K_I 时所需裂纹尺寸和应力值的确定　在平面缺陷断裂常规评定计算 K_I 时，所输入的裂纹尺寸 a、c 和应力值 σ_m、σ_B 或 P_m、P_b、Q_m、Q_b，应按本章 2.4.1 和 2.4.2 节说明的方法进行确定。

② 疲劳评定所需应力强度因子变化范围 ΔK 的计算规则　这里所给出各种 K_I 计算式形式，均适用于下一节缺陷的疲劳评定时有关 ΔK、ΔK_a、ΔK_c 的计算。

在进行 K_I 计算前，需要计算相应的 ΔK、ΔK_a、ΔK_c 的值，并进行参数的替代，具体计算和替代方法如下：

下面算式中的 σ_m、σ_B 分别以 $\Delta\sigma_m$、$\Delta\sigma_B$ 替代，则所得 K_I 即为 ΔK；

若 σ_m、σ_B 分别为一次应力变化范围 ΔP_m、ΔP_b 时，则所得结果是一次应力变化引起的应力强度因子变化范围 ΔK^P；

若 σ_m、σ_B 分别为二次应力变化范围 ΔQ_m、ΔQ_b 时，则所得结果是二次应力强度因子变化范围 ΔK^S；

在计算裂纹尺寸 a 方向裂尖处的应力强度因子变化范围 ΔK_a 时，应采用 a 方向的裂纹构形因子 f_m^A、f_b^A；

在计算裂纹尺寸 c 方向裂尖处的应力强度因子变化范围 ΔK_c 时，应采用 c 方向的裂纹构形因子 f_m^B、f_b^B。

ΔK 计算中所需输入的裂纹尺寸 a、c 和应力变化范围值 $\Delta\sigma_m$、$\Delta\sigma_B$ 或 ΔP_m、ΔP_b、ΔQ_m、ΔQ_b，应按本章 2.4.1 节和下面有关疲劳评定的方法中的规定确定。

③ 典型结构的 K_I 计算方法　这里给出了常见的几种典型的结构的 K_I 的计算式。下面将分别加以介绍，同时给出其相应的适用范围。

a）对于含长 $2a$ 穿透裂纹的板壳，假设其板宽为 $2W$，板长为 $2L$，有

$$K_I = \sqrt{\pi a}\,(\sigma_m + \sigma_B) \quad (5.4\text{-}62)$$

其适用范围为：$a/W \leqslant 0.15$ 且 $a/L \leqslant 0.15$。

b）对于含半椭圆表面裂纹（$a \times 2c$）的板壳，假定其板宽为 $2W$，板长为 $2L$，板厚为 B，则有

$$K_I = \sqrt{\pi a}\,(\sigma_m f_m + \sigma_B f_m) \quad (5.4\text{-}63)$$

式中

$$f_m^A = \frac{1}{\left[1+1.464\left(\frac{a}{c}\right)^{1.65}\right]^{0.5}}\left[1.13-0.09\frac{a}{c}+\left(-0.54+\frac{0.89}{0.2+\frac{a}{c}}\right)\left(\frac{a}{B}\right)^2+\left(0.5-\frac{1}{0.65+\frac{a}{c}}+14\left(1-\frac{a}{c}\right)^{24}\right)\left(\frac{a}{B}\right)^4\right]$$

$$f_b^A = \left[1+\left(-1.22-0.12\frac{a}{c}\right)\frac{a}{B}+\left(0.55-1.05\left(\frac{a}{c}\right)^{0.75}+0.47\left(\frac{a}{c}\right)^{1.5}\right)\left(\frac{a}{B}\right)^2\right]f_m^A$$

$$f_m^B = \left[\left(1.1+0.35\left(\frac{a}{B}\right)^2\right)\left(\frac{a}{c}\right)^{0.5}\right]f_m^A$$

$$f_b^B = \left[1-0.34\frac{a}{B}-0.11a^2/(cB)\right]f_m^B$$

式中，上标 A 表示求裂纹深度处 K_I 时用的系数；上标 B 表示求裂纹长度方向两端点处 K_I 时用的系数。

f_m^A、f_b^A、f_m^B、f_b^B 亦可分别由相关标准或规范中的表格中直接查得。

上式的适用范围为：$a/B \leqslant 0.8$、$a/c \leqslant 1.0$、$c/L \leqslant 0.15$、$c/W \leqslant 0.15$。

c）对于含椭圆埋裂纹（$2a \times 2c$）的板壳，若其板宽为 $2W$、板厚为 B，则有

$$K_I=\sqrt{\pi a}\ (\sigma_m f_m+\sigma_B f_b) \qquad (5.4\text{-}64)$$

式中

$$f_m^A=\frac{1.01-0.37\frac{a}{c}}{\left[1-\left(\frac{2a/B}{1-2e/B}\right)^{1.8}\left(1-0.4\frac{a}{c}-\left(\frac{e}{B}\right)^2\right)\right]^{0.54}}$$

$$f_b^A=\frac{\left(1.01-0.37\frac{a}{c}\right)\ [2e/B+a/B+0.34a^2/(cB)]}{\left[1-\left(\frac{2a/B}{1-2e/B}\right)^{1.8}\left(1-0.4\frac{a}{c}-\left(\frac{e}{B}\right)^2\right)\right]^{0.54}}$$

$$f_m^B=\frac{1.01-0.37\frac{a}{c}}{\left[1-\left(\frac{2a/B}{1-2e/B}\right)^{1.8}\left(1-0.4\frac{a}{c}-0.8\left(\frac{e}{B}\right)^2\right)\right]^{0.54}}$$

$$f_b^B=\frac{\left(1.01-0.37\frac{a}{c}\right)\ [2e/B-a/B-0.34a^2/(cB)]}{\left[1-\left(\frac{2a/B}{1-2e/B}\right)^{1.8}\left(1-0.4\frac{a}{c}-0.8\left(\frac{e}{B}\right)^2\right)\right]^{0.54}}$$

式中，e 为埋裂纹中心与板厚中心的偏移量。

f_m^A、f_b^A、f_m^B、f_b^B 也可由相关标准或规范中的表格中直接查得。但一般的表格中是以缺陷离表面的最近距离 p_1 表示，而不用参量 e，它们之间的关系是 $p_1+a+e=B/2$。

此种算法的适用范围为：$a/B\leqslant 0.45\ (1-2e/B)$、$a/c\leqslant 1.0$、$c/W\leqslant 0.15$、$c/B\leqslant 0.5$ 的情况。

d) 对于焊趾处的表面裂纹，假设其为深 a，长 $2c$ 的半椭圆形（见图 5.4-28），有

$$K_I=\sqrt{\pi a}\ (\sigma_m f_m^A M_{km}+\sigma_B f_b^A M_{kb})\ /\Phi \qquad (5.4\text{-}65)$$

式中，K_I 为裂纹最深处的值；Φ 为第二类椭圆积分，其近似值为 $\Phi=[1+1.464\times a/c)^{1.65}]^{0.5}$，或由表 5.4-5 查得；$f_m^A$、$f_b^A$ 可通过相关表格查得；M_{km}、M_{kb} 亦可通过相关表格查得。

表 5.4-5　Φ 值的确定

a/c	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Φ	1.00	1.02	1.05	1.10	1.15	1.21	1.28	1.35	1.42	1.50	1.57

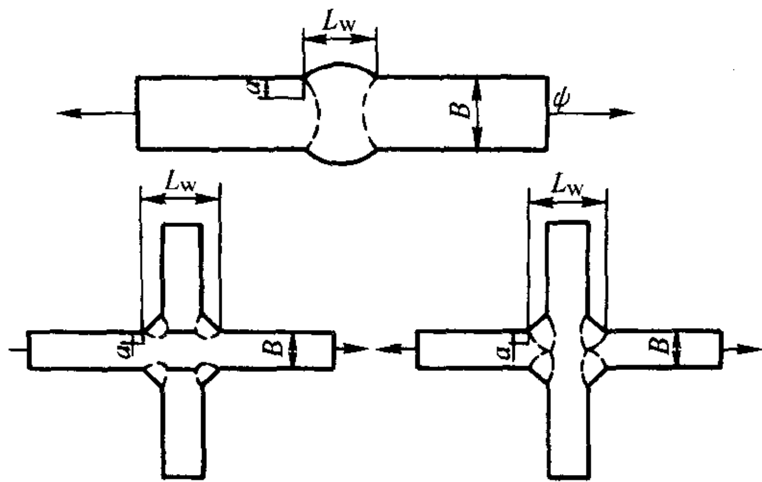


图 5.4-28　焊趾处表面裂纹的结构尺寸

对于十字接头中的焊根裂纹，如图 5.4-29 所示。

$$K_I=\sigma_m M_k\ \sqrt{\pi a\sec\ (\pi a/L_w)} \qquad (5.4\text{-}66)$$

式中，M_K 值可由相关标准或规范中的表格中直接查得。

式 (5.4-66) 的适用范围：$0.2\leqslant h_w/B\leqslant 1.2$、$0.02\leqslant a/B\leqslant 0.7$。

当仅受内压圆筒的 1/4 圆形接管拐角裂纹时（见图 5.4-29），有

$$K_I=P_m\ \sqrt{\pi a}\ \left(1+\sqrt{\frac{R_{in}B_n}{R_iB}}\right)\ f_{cc} \qquad (5.4\text{-}67)$$

式中，f_{cc} 为 a/B 有关的系数，可由表 5.4-6 查得。

表 5.4-6　f_{cc} 的值

a/B	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
f_{cc}	1.88	1.74	1.60	1.49	1.38	1.30	1.22	1.18	1.13

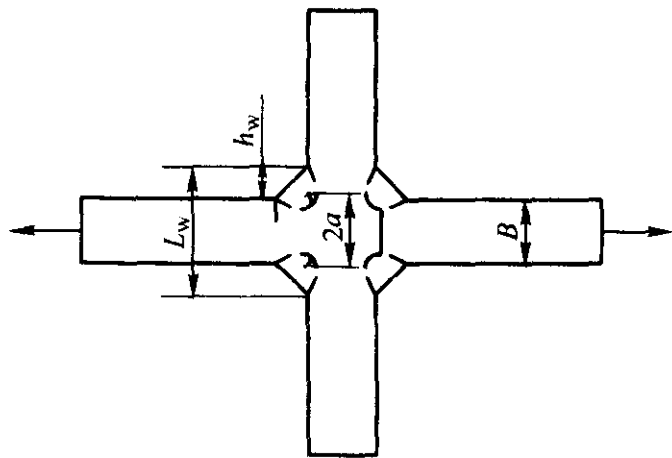


图 5.4-29　十字接头焊根裂纹结构尺寸

式 (5.4-67) 的适用范围为：$a/B\leqslant 0.8$、$R_{in}/R_i\leqslant 0.4$。

对于仅受内压圆筒接管拐角的椭圆裂纹，有

$$K_I=M_\sigma F_c P_m\frac{\sqrt{\pi a}}{\Phi}\left[\left(\frac{4}{\pi}+\frac{ac}{2B_{in}R_{in}}\right)\Big/\left(\frac{4}{\pi}+\frac{ac}{B_{in}R_{in}}\right)\right]^{1/2} \qquad (5.4\text{-}68)$$

式中，a、c 为拐角裂纹在接管侧和容器侧的裂纹长度；R_{in} 为接管内半径；P_m 为容器薄膜应力；Φ 为椭圆裂纹形状因子，可用以下经验式表示：

$$\begin{cases}\Phi=\sqrt{1+1.46\ (a/c)^{1.65}} & \text{当 } a/c\leqslant 1.0 \text{ 时}\\ \Phi=\sqrt{1+1.46\ (c/a)^{1.65}} & \text{当 } a/c>1.0 \text{ 时}\end{cases}$$

M_σ 为应力集中修正因子，$M_\sigma=K_t/3$，其中，K_t 为接管拐角的弹性应力集中系数，该值优先选用实测值；在无法获得无实测值时可用下式计算：

$$K_t=\frac{2+2\frac{D_n}{D}\left(\frac{D_nB_n}{DB}\right)^{1/2}+1.25\left(\frac{D}{B}\right)^{1/2}\left(\frac{D_n}{B}\right)}{1+\frac{B_n}{B}\left(\frac{D_nB_n}{DB}\right)^{1/2}}$$

F_c 为接管拐角裂纹的边界修正因子，可用以下公式计算：

$$F_c=[M_1+M_2(a_{45}/B_{45})^2+M_3(a_{45}/B_{45})^4]g_1g_2g_3g_4f_\varphi$$

式中，$f_\varphi=[(a/c)^2\cos^2\varphi+\sin^2\varphi]^{1/4}$。拐角裂纹在筒内壁处 C 点（图 5.4-30）的 $\varphi=0$，在接管内壁处的 B 点的 $\varphi=\frac{\pi}{2}$；a_{45}、B_{45} 分别为接管拐角平分线方向的拐角裂纹深度和容器壁厚。

式中其他系数计算如下：

对 $a/c\leqslant 1.0$

$$M_1=1.13-0.09a/c$$

$$M_2=-0.54+\frac{0.89}{0.2+a/c}$$

$$M_3=0.5-\frac{1}{0.65+a/c}+14\ (1-a/c)^{24}$$

$$g_1=1+\ [0.1+0.35\ (a_{45}/B_{45})^2]\ (1-\sin\varphi)^2$$

$$g_2=\frac{1+0.358\lambda+1.425\lambda^2-1.578\lambda^3+2.156\lambda^4}{1+0.13\lambda^2}$$

$$\lambda=\frac{1}{1+\ (c/R_{in})\ \cos\ (0.85\varphi)}$$

$$g_3=(1+0.04a/c)[1+0.1(1-\cos\varphi)^2][0.85+0.15(a_{45}/B_{45})^{1/4}]$$

$$g_4=1-0.7(1-a_{45}/B_{45})(a/c-0.2)(1-a/c)$$

对 $a/c>1.0$

$M_1=(c/a)^{1/2}(1+0.04c/a)$

$M_2=0.2(c/a)^4$

$M_3=-0.1(c/a)^4$

$g_1=1+[0.1+0.35(c/a)(a_{45}/B_{45})^2](1-\sin\varphi)^2$

g_2 与 $a/c\leqslant 1$ 时的 g_2 相同

$g_3=(1.13-0.09c/a)[1+0.1(1-\cos\varphi)^2][0.85+0.15(a_{45}/B_{45})^{1/4}]$

$g_4=1$

也可按相关标准或规范中的提供的表格中直接查得拐角裂纹在筒体表面及接管表面的 F_c/g_2 及 g_2 值，由 F_c/g_2 的值乘以 g_2 之值则可得到 F_c，即 $F_c=F_c/g_2\times g_2$

以上算法的适用范围为：$a_{45}/B_{45}\leqslant 0.8$、$a/c=0.2\sim 2.0$、$R_{in}/R_i\leqslant 0.4$。

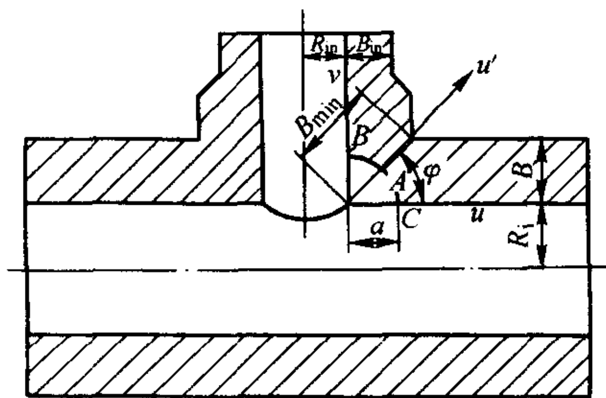


图 5.4-30　内压容器接管 1/4 圆形拐角裂纹的结构的尺寸

(4) K_r 的计算

断裂比 K_r 值是失效评定图 FAC 中的纵坐标值，该值将直接影响到评定的结果，因此其计算的准确性和合理性十分重要，下面介绍它的相关算法。K_r 值可按式 (5.4-69) 计算。

$$K_r=G\ (K_I^P+K_I^S)\ /K_p+\rho \tag{5.4-69}$$

式中，G 为相邻两裂纹间弹塑性干涉效应系数，按本章 2.4.1 节中的说明来确定；K_p 为评定用材料断裂韧度，即按前面已求得的 K_c 值除以表 5.4-1 规定的分安全系数；ρ 为塑性修正因子，按式 (5.4-70) 求得：

$$\rho=\begin{cases}\varphi_1 & \text{当 } L_r<0.8 \text{ 时}\\ \varphi_1\ (11-10L_r)\ /3 & \text{当 } 0.8<L<1.1\\ 0 & \text{当 } L_r>L \text{ 时}\end{cases} \tag{5.4-70}$$

式中，φ_1 的值可以由图 5.4-31 根据 $K_I^S/(\sigma_s\sqrt{\pi a})$ 的值查得。L_r 按本节中的前面的说明来进行计算分析。

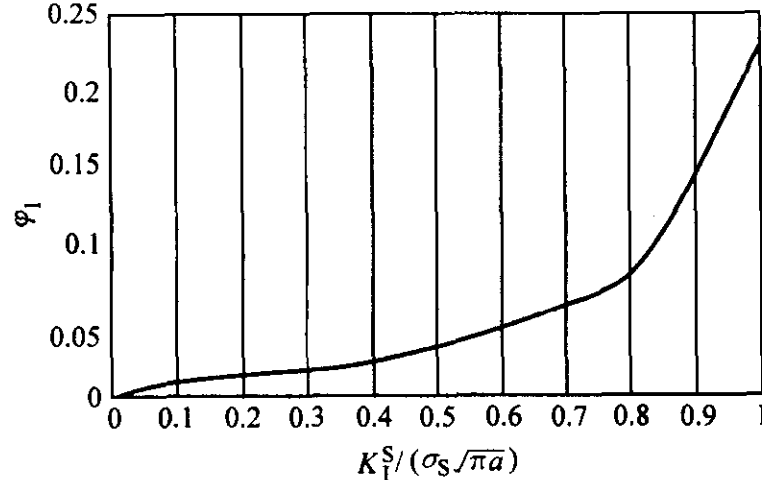


图 5.4-31　φ_1 取值

(5) L_r 值的计算及确定

L_r 为失效评定图中的横轴坐标值，因此同 K_r 值一样在评定过程中非常重要。它的算法可按本章 2.2.4 节平面缺陷简化评定中的计算方法求得。

(6) 安全性评价

将按前面 (4) 和 (5) 计算得的 K_r 和 L_r 值所构成的评定点 (L_r，K_r) 绘在常规评定通用失效评定图 5.4-26 中。如果该评定点位于安全区之内，则认为该缺陷经评定是安全的或可以接受的；否则，认为不能保证安全或不可接受。如果 $L_r<L_r^{max}$ 而评定点位于失效评定曲线上方，则容许采用分析评定方法重新评定。

2.4.6　压力容器确定性分析评定方法

作为第三级评定，分析评定主要用于重要的大型设备或部件在常规评定方法不能通过评定的缺陷又难于返修的场合。平面缺陷的分析评定方法主要是在已经确定裂纹起裂后，在裂纹发生延性稳定扩展至失稳临界尺寸过程中，分析构件所能承受的许用载荷，并对结构的安全性进行评价。目前，压力容器平面缺陷的分析评定方法采用 EPRI-82 工程优化方法，并在优化的方法上有一定的创新。

平面缺陷的分析评定方法直接采用 J 积分为断裂参量，是最严格的弹塑性断裂力学方法，该方法能精确地评定含缺陷材料（如压力容器、压力管道等）从起裂、有限量撕裂、直至撕裂失稳的全过程。PD6493—91 的第 3 级评定采用通用失效评定曲线及 $R6$ 的选择 2 曲线并非是严格的失效评定曲线，难于达到精确评定的目的。$R6$ 的选择 3 失效评定曲线及第 3 类评定方法能达到精确评定的要求，也可以采用与 $R6$ 选择 3 等价的、由 EPRI 建立的稳定性评定图法，以 $J\ (\sigma,\ a)\ =J_R\ (\Delta a)$ 及 $\frac{\partial J\ (\sigma_R,\ a)}{\partial a}=\frac{\partial J\ (\Delta a)}{\partial a}$ 作为失稳的条件，由阻力曲线与推力曲线相切来确定其失稳点，但切点难于判别。这里介绍一种“优化”计算方法，即由给定的 Δa_k 寻求满足平衡条件 $J\ (\sigma,\ a_k)\ =J_R\ (\Delta a_k)$ 的相容应力 σ_k，找诸相容应力的极大值即为失稳应力，并用软件来实施，该方法称为 EPRI 工程优化评定方法，它反映了目前我国在安全评定方面的创新成果。

如上所述，分析评定需要非常多的而且非常详尽的原始资料，例如可靠的 J 积分解和材料的整条 J_R 阻力曲线等，但目前仅有部分含裂纹结构的 J 积分解，因而目前还不是任何情况都能采用分析评定，但将会有越来越多的承压部件能进行分析评定，尤其是难于做出判断和抉择的场合。同时，分析评定方法也是常规评定方法的技术基础，是研究、发展、评价常规评定方法和简化评定方法的重要手段。下面具体介绍平面缺陷的分析评定方法的相关内容。

(1) 适用范围

如上所述，分析评定并非对所有材料在所有工况下都能进行评定，因此在采用这种方法进行分析前需要了解它的使用范围。

平面缺陷的分析评定方法适用于 ROR 关系材料及具有较长屈服平台材料的设备或部件，如容器内表面的环向裂纹、整圈环向裂纹、轴向裂纹及超长轴向裂纹等。

(2) 分析评定程序

同简化评定和常规评定类似，分析评定也有它的评定程序。平面缺陷的分析评定按以下几个步骤来进行：

1) 确定裂纹构形及材料断裂性能；
2) 确定裂纹延性撕裂失稳载荷（应力）；
3) 确定构件塑性极限载荷（应力）；
4) 确定安全系数的选取及许用载荷；
5) 进行含裂纹结构安全性的评价。

(3) 分析评定所需确定的基本数据

如上所述，分析评定时需要比简化评定和常规评定更多和更具体的参数。具体说来在分析评定时所需确定的裂纹构形尺寸和材料性能等基本数据如下：

1) 材料的厚度、表面裂纹原始深度、裂纹原始半长（B、

a_0、c_0）；

2）材料的尺寸参数，如容器内、外半径（R_i、R_0）；

3）具有ROR关系材料的应变硬化系数和指数（α、n）；

4）对具有较长屈服平台材料首先确定其应力应变曲线，然后分两段进行ROR关系拟合，可以得到两对应变硬化系数和指数（α_1、α_2、n_1、n_2），两段拟合曲线交点处应为σ_c；

5）材料屈服强度、抗拉强度、弹性模量、泊松比（σ_s、σ_b、E、ν）等性能参数的确定；

6）材料的J_R阻力曲线数据的确定，参照可靠的标准试验方法进行测试。目前，一般可以采用两种方法进行测试，单试样法和多试样法。如果采用单试样（柔度）法测定时，可以从三个单试样试验所获得的最低曲线中得到评定数据。如果采用多试样法，至少应有四个有效试样的数据，并且从下包络曲线或回归曲线减去一个标准偏差的下分散带来获得数据；

7）焊接残余应力。对于轴向表面裂纹，如确认沿筒体壁厚分布有环向残余应力，则应在评定前，通过实验或其他可靠方法确定残余应力σ_R，并将残余应力沿壁厚方向的分布函数用下式来表示：

$$\sigma_R=\sigma_{(0)}+\sigma_{(1)}\frac{Z}{B}+\sigma_{(2)}\left(\frac{Z}{B}\right)^2+\sigma_{(3)}\left(\frac{Z}{B}\right)^3\quad(\text{在内壁},Z=0)$$

（4）安全系数的选取

安全系数的选取要根据失效模式来选取，通常可分别按裂纹撕裂失稳和塑性失稳失效模式来选取。

1）对于裂纹撕裂失稳失效模式，安全系数n_{ins}取1.5～2.2，具体数值可根据评定要求、失效后果以及评定经验综合确定。

2）对于结构塑性失稳的失效模式，安全系数n_L按下述原则选取：失效后果一般的取1.3；失效后果严重的取1.5。

（5）裂纹延性撕裂失稳载荷的确定（确定应力）

确定裂纹的延性撕裂失稳载荷的具体计算步骤如下：

1）首先需要给定一系列裂纹扩展量$\Delta a(i)$　（$i=1,2,\cdots,n$）；$\Delta a(i)$由评定计算者根据经验给定，数据间距越小计算精度越高；

2）在材料J_R阻力曲线上确定一系列与$\Delta a(i)$相对应的材料撕裂断裂韧度值$J_R\Delta a(i)$（$i=1,2,\cdots,n$）；

3）由式（5.4-71）解出平衡应力$\sigma(i)$（$i=1,2,\cdots,n$）。式（5.4-71）是以$\sigma(i)$为单一未知数的非线性方程，需迭代求解，验敛误差应不大于2%。

$$J(a_0+\Delta a(i),\sigma(i))=\Delta J_R(\Delta a(i))\qquad(5.4\text{-}71)$$

方程左边项为被评定结构含深为（$a_0+\Delta a(i)$）、半长为c_0的裂纹在外载应力$\sigma(i)$作用下的J积分，计算公式由下面按几种不同情况的J积分的算法。

① 一般含裂纹结构J积分计算公式的确定　含裂纹构件的弹塑性J积分等于弹性J积分J_e与全塑性J积分J_p之和，即有

$$J=J_e+J_p\qquad(5.4\text{-}72)$$

这里给出四种典型裂纹构形的J积分表达式。在下面计算全塑性J积分J_p时需要用到全塑性解系数H_1，该值可由相关标准或规范中的表格中查出。

a）材料（如压力容器，压力管道）内表面环向裂纹J积分

$$J_e=\frac{\sigma_s^2a}{E}\left[\pi(1-\nu^2)F_e^2\left(\frac{B}{a}\right)\left(\frac{\sigma}{\sigma_s}\right)^2\right]\qquad(5.4\text{-}73)$$

$$\frac{a_e}{B}=\frac{a}{B}+\frac{n-1}{6(n+1)}\cdot F^2\left(\frac{\sigma}{\sigma_S}\right)^2\Big/\left[1+B_R\left(\frac{\sigma}{\sigma_S}\right)^2\right]\qquad(5.4\text{-}74)$$

$$\begin{cases}F=0.25+0.5298\left(\dfrac{a}{BA_0}\right)+0.3835\left(\dfrac{a}{BA_0}\right)^2 & \text{当}\dfrac{a}{BA_0}\geqslant0.25\text{时}\\ F=3.72\left(\dfrac{a}{BA_0}\right)-13.475\left(\dfrac{a}{BA_0}\right)^2+19.988\left(\dfrac{a}{BA_0}\right)^3 & \text{当}\dfrac{a}{BA_0}<0.25\text{时}\end{cases}\qquad(5.4\text{-}75)$$

$$A_0=(0.25+a/c)^{0.58}\qquad(5.4\text{-}76)$$

$$B_R=\left\{(R_0^2-R_i^2)/\left[R_0^2-(R_i+a)^2+\left(1-\frac{c}{4R_i}\right)(2R_ia+a^2)\right]\right\}^2\qquad(5.4\text{-}77)$$

$$\sigma=P/[\pi(R_0^2-R_i^2)]\qquad(5.4\text{-}78)$$

$$\begin{cases}J_p=\dfrac{\sigma_s^2a}{E}\left[\alpha_1\left(\dfrac{B}{a}\right)H_1\left(\dfrac{\sigma}{\sigma_s}\right)^{n_1+1}\right] & \text{当}\sigma\leqslant\sigma_c\text{时}\\ J_p=\dfrac{\sigma_s^2a}{E}\left\{\alpha_1\left(\dfrac{B}{a}\right)H_1\left(\dfrac{\sigma_c}{\sigma_s}\right)^{n_1+1}+\alpha_2\left(\dfrac{B}{a}\right)H_2\left[\left(\dfrac{\sigma}{\sigma_s}\right)^{n_2+1}-\left(\dfrac{\sigma_c}{\sigma_s}\right)^{n_2+1}\right]\right\} & \text{当}\sigma>\sigma_c\text{时}\end{cases}\qquad(5.4\text{-}79)$$

b）材料（如压力容器，压力管道）内表面长轴向裂纹J积分

$$J_e=\frac{\sigma_s^2a}{E}\left[\pi(1-\nu^2)F_e^2\left(\frac{a_e}{a}\right)\left(\frac{\sigma}{\sigma_s}\right)^2\right]\qquad(5.4\text{-}80)$$

$$\frac{a_e}{a}=1+\frac{n-1}{6(n+1)}F^2\left(\frac{\sigma}{\sigma_s}\right)^2\Big/\left[1+B_R\left(\frac{\sigma}{\sigma_s}\right)^2\right]\qquad(5.4\text{-}81)$$

$$F=1.1+A\left[4.951\left(\frac{a}{B}\right)^2+1.092\left(\frac{a}{B}\right)^4\right]\qquad(5.4\text{-}82)$$

$$\begin{cases}A=[0.125(R_i/B)-0.25]^{0.125} & \text{当}5\leqslant R_i/B\leqslant10\\ A=[0.2(R_i/B)-1.0]^{0.25} & \text{当}10\leqslant R_i/B\leqslant20\text{时}\end{cases}\qquad(5.4\text{-}83)$$

$$B_R=\frac{3}{16}[(1-R_i^2/R_0^2)(R_i+a)/(B-a)^2]\qquad(5.4\text{-}84)$$

$$\sigma=2PR_0^2/[\pi(R_0^2-R_i^2)]\qquad(5.4\text{-}85)$$

$$\begin{cases}J_p=\dfrac{\sigma_s^2a}{E}\left[\alpha_1\left(\dfrac{B}{a}\right)H_1\left(\dfrac{\sigma}{\sigma_s}\right)^{n_1+1}\right] & \text{当}\sigma\leqslant\sigma_c\text{时}\\ J_p=\dfrac{\sigma_s^2a}{E}\left[\alpha_1\left(\dfrac{B}{a}\right)H_1\left(\dfrac{\sigma_c}{\sigma_0}\right)^{n_1+1}+\alpha_2\left(\dfrac{B}{a}\right)H_2\left\{\left(\dfrac{\sigma}{\sigma_s}\right)^{n_2+1}-\left(\dfrac{\sigma_c}{\sigma_s}\right)^{n_2+1}\right\}\right] & \text{当}\sigma>\sigma_c\text{时}\end{cases}\qquad(5.4\text{-}86)$$

c）材料（如压力容器，压力管道）内表面全周长环向裂纹J积分

$$J_e=\frac{\sigma_s^2a}{E}\left[\pi(1-\nu^2)F_e^2\left(\frac{a_e}{a}\right)\left(\frac{\sigma}{\sigma_s}\right)^2\right]\qquad(5.4\text{-}87)$$

$$\frac{a_e}{a}=1+\frac{n-1}{6(n+1)}F^2\left(\frac{\sigma}{\sigma_S}\right)^2\Big/\left[1+\left(\frac{\sigma}{\sigma_s}\right)^2\right]\qquad(5.4\text{-}88)$$

$$F=1.1+A\left[1.984\left(\frac{a}{B}\right)^{1.5}+0.3342\left(\frac{a}{B}\right)^{4.2}\right]\qquad(5.4\text{-}89)$$

$$\begin{cases}A=[0.125(R_i/B)-0.25]^{0.25} & \text{当}5\leqslant R_i/B\leqslant10\\ A=[0.4(R_i/B)-3.0]^{0.25} & \text{当}10\leqslant(R_i/B)\leqslant20\text{时}\end{cases}\qquad(5.4\text{-}90)$$

$$\frac{P}{P_0}=\frac{\sqrt{3}}{2}\frac{R_0^2-R_i^2}{R_0^2-(R_i+a)^2}\left(\frac{\sigma}{\sigma_s}\right)\qquad(5.4\text{-}91)$$

$$\sigma=P/[\pi(R_0^2-R_i^2)]\qquad(5.4\text{-}92)$$

$$\begin{cases} J_p = \frac{\sigma_s^2 a}{E}\left[\alpha_1\left(\frac{B}{a}\right)\left(1-\frac{a}{B}\right)^2 H_1\left(\frac{P}{P_0}\right)^{n_1+1}\right] & \text{当 } \sigma \leqslant \sigma_c \text{ 时} \\ J_p = \frac{\sigma_s^2 a}{E}\left[\alpha_1\left(\frac{B}{a}\right)\left(1-\frac{a}{B}\right)^2 H_1\left(\frac{P_c}{P_0}\right)^{n_1+1} + \alpha_2\left(\frac{B}{a}\right)\left(1-\frac{a}{B}\right)^2 H_2\left\{\left(\frac{P}{P_0}\right)^{n_2+1} - \left(\frac{P_c}{P_s}\right)^{n_2+1}\right\}\right] & \text{当 } \sigma > \sigma_c \text{ 时} \end{cases} \quad (5.4\text{-}93)$$

$$\frac{P_c}{P_0} = \frac{\sqrt{3}}{2}\frac{R_0^2 - R_i^2}{R_0^2 - (R_i+a)^2}\left(\frac{\sigma_c}{\sigma_s}\right) \quad (5.4\text{-}94)$$

d) 材料（如压力容器，压力管道）内表面轴向裂纹 J 积分

$$J_e = \frac{\sigma_s^2 a}{E}\left[\pi(1-\nu^2)F_e^2\left(\frac{B}{a}\right)\left(\frac{\sigma}{\sigma_s}\right)^2\right] \quad (5.4\text{-}95)$$

$$F_e = 0.25 + 0.475\,9\alpha_e + 0.126\,2\alpha_e^2 \quad (5.4\text{-}96)$$

$$\alpha_e = \left[\frac{a}{B} + \frac{n+1}{6\ (n+1)}\frac{F^2\ (\sigma/\sigma_s)^2}{1+B_R\ (\sigma/\sigma_s)^2}\right] \Big/ \ (a/c)^{0.58} \quad (5.4\text{-}97)$$

$$F = 0.25 + 0.475\,9\alpha_0 + 0.126\,2\alpha_0^2 \quad (5.4\text{-}98)$$

$$\alpha_0 = \left(\frac{a}{B}\right)\Big/\left(\frac{a}{c}\right)^{0.58} \quad (5.4\text{-}99)$$

$$B_R = 1.045\,3\left\{\left[1+\frac{a}{B}\Big/\left(1+0.153\,3\frac{C^2}{B^2}\right)^{0.5}\right]\Big/\left(1-\frac{a}{B}\right)\right\} \quad (5.4\text{-}100)$$

$$\sigma = \frac{R_0^2 + R_i^2}{R_0^2 - R_i^2}P \quad (5.4\text{-}101)$$

$$\begin{cases} J_p = \frac{\sigma_s^2 a}{E}\left[\alpha_1\left(\frac{B}{a}\right)H_1\left(\frac{\sigma}{\sigma_s}\right)^{n_1+1}\right] & \text{当 } \sigma \leqslant \sigma_c \text{ 时} \\ J_p = \frac{\sigma_s^2 a}{E}\left\{\alpha_1\left(\frac{B}{a}\right)H_1\left(\frac{\sigma_c}{\sigma_0}\right)^{n_1+1} + \alpha_2\left(\frac{B}{a}\right)H_2\left[\left(\frac{\sigma}{\sigma_0}\right)^{n_2+1} - \left(\frac{\sigma_c}{\sigma_s}\right)^{n_2+1}\right]\right\} & \text{当 } \sigma > \sigma_c \text{ 时} \end{cases} \quad (5.4\text{-}102)$$

以上 J 积分值的计算是在没有考虑残余应力的情况下所获得的。下面介绍一种考虑残余应力时 J 积分的计算方法。

② 考虑残余应力时的 J 积分计算公式　在 J 积分计算中，只需考虑残余应力对线弹性部分 J_e 中的附加影响即可。

$$J_e = \frac{\sigma_s^2 a}{E}\left[\pi\ (1-\nu^2)\ F_e^2\left(\frac{B}{a}\right)\left(\frac{\sigma+\sigma_R^e}{\sigma_s}\right) + \alpha\left(\frac{B}{\alpha}\right)H\left(\frac{\sigma}{\sigma_s}\right)^{n+1}\right] \quad (5.4\text{-}103)$$

式中，等效残余应力 σ_R^e 由实际残余应力 σ_R 分布曲线（式 (5.4-70)）根据“应力强度因子相等”原则换算。

$$K_I = \sqrt{\pi B}\sum_0^3 \sigma_{(i)} G_i$$

等效的均匀残余应力 σ_R^e 产生的应力强度因子为

$$K_I = \sqrt{\pi B}\cdot\sigma_R^e\cdot G_0$$

$$\sigma_R^e = \left(\sum_0^3 \sigma_{(i)} G_i\right)/G_0$$

$$G_i = (A_1\alpha_i + A_2\alpha_i^2 + A_3\alpha_i^3 + A_4\alpha_i^4 + A_5\alpha_i^5)\Big/\left(0.102\frac{R_i}{B} - 0.02\right)^{0.05} \quad (5.4\text{-}104)$$

$$\alpha_i = \left(\frac{a}{B}\right)\Big/\left(\frac{a}{c}\right)^m \quad (5.4\text{-}105)$$

式 (5.4-104)、式 (5.4-105) 中的系数 $A_1 \sim A_5$ 及 m 的取值见表 5.4-7。

表 5.4-7　计算等效残余应力时各系数的取值

	A_1	A_2	A_3	A_4	A_5	m
G_0	1.776 7	−2.597 5	2.752 0	−1.323 7	0.236 3	0.58
G_1	0.104 5	0.418 9	0	0	0	0.22
G_2	0.020 4	−0.004 0	0.421 3	0	0	0.10
G_3	0.072 8	−0.360 1	0.66 88	0	0	0.05

4）由 n 对离散点（σ（i），Δa（i））描述的 $\sigma \sim \Delta a$ 曲线上，确定 $\frac{\partial\sigma}{\partial\Delta a} = 0$ 处的 σ 和 Δa；也可用（σ（i），Δa（i））数据系列中的最大 σ（i）所对应的位置近似作为 $\sigma \sim \Delta a$ 曲线上 $\frac{\partial\sigma}{\partial\Delta a} = 0$ 的位置；

5）上步中确定的 σ 即为裂纹撕裂失稳应力 σ_f，相应的载荷即为延性撕裂失稳载荷 P_f，Δa 即为裂纹由稳定扩展至失稳扩展的临界裂纹扩展量 Δa_f。

(6) 构件塑性极限载荷的确定

计算裂纹扩展至临界尺寸时含裂纹构件的塑性极限载荷 P_L 或 p_L。与前面相对应的这里同样给出四种典型的裂纹构形的塑性极限载荷表达式，并需以 $a_0 + a_f$ 作为裂纹深度 a 代入公式计算。

1）材料（如压力容器或压力管道等）内表面环向裂纹

$$P_L = 2\pi RB\bar{\sigma}\left[\frac{2a_L}{\pi} - \left(\frac{a}{B}\right)\left(\frac{\theta}{\pi}\right)\left(2-2\zeta+\frac{a}{B}\zeta\right)/\ (2-\zeta)\right] \quad (5.4\text{-}106)$$

$$a_L = \arccos\ (A_1\sin\theta) \quad (5.4\text{-}107)$$

$$A_1 = \frac{a}{B}(1-\zeta)\left[\left(2-2\zeta+\frac{a}{B}\zeta\right)^2 + \left(1-\zeta+\frac{a}{B}\zeta\right)^2\right]\Big/ \frac{1}{2}[1+(2-\zeta)(1-\zeta)] \quad (5.4\text{-}108)$$

$$\theta = \frac{\pi c}{4R_i} \quad (5.4\text{-}109)$$

$$\zeta = B/R_0 \quad (5.4\text{-}110)$$

2）材料（如压力容器或压力管道等）内表面全周长环向裂纹

$$P_L = \left(\frac{2\pi}{\sqrt{3}}\right)\bar{\sigma}\ [R_0^2 - \ (R_i+a)^2] \quad (5.4\text{-}111)$$

$$\bar{\sigma} = \frac{1}{2}\ (\sigma_s + \sigma_b) \quad (5.4\text{-}112)$$

3）材料（如压力容器或压力管道等）内表面轴向裂纹

$$P_L = \bar{\sigma}\left(\frac{B}{R_0}\right)\left[\left(1-\frac{a}{B}\right)\Big/\left(1-\frac{a}{B}/M_g\right)\right] \quad (5.4\text{-}113)$$

$$M_g = \ [1+1.61c^2/\ (RB)]^{0.5} \quad (5.4\text{-}114)$$

4）材料（如压力容器或压力管道等）内表面长轴向裂纹

$$P_L = \frac{2}{\sqrt{3}}\bar{\sigma}\left(\frac{B}{R_i}\right)\left\{\left(1-\frac{a}{B}\right)\Big/\left[1+\frac{a}{B}/\ (R_i/t)\right]\right\} \quad (5.4\text{-}115)$$

(7) 许用载荷的确定

如前所述，在分析评定中需要确定裂纹撕裂失稳破坏模式下的许用载荷，通过以上的方程可以确定不同缺陷情况下的构件塑性极限载荷的情况为

$$[P_f] = \frac{P_f}{n_{ins}} \text{或} [P_f] = \frac{P_f}{n_{ins}} \quad (5.4\text{-}116)$$

结构塑性失稳破坏模式下的许用载荷为

$$[P_f]=\frac{P_f}{n_L}\text{或}[P_f]=\frac{P_f}{n_L} \quad (5.4\text{-}117)$$

结构的许用载荷为

$$[P]=\min\{[P_f],[P_L]\}\text{或}[P]=\min\{[p_f][p_L]\} \quad (5.4\text{-}118)$$

(8) 安全性评价

通过以上的分析计算，可以获得极限工作载荷和许用载荷，如果工作载荷小于许用载荷，则认为该缺陷是安全或可以接受的；否则，不能保证安全或不可接受。

以上为三级评定，即分析评定的整个过程。

2.4.7　压力管道确定性安全评定方法

作为现代五大运输系统（公路、铁路、水运、航空、管道）之一的管道运输在现代工业生产和城市建设中的应用已经遍及各个领域，几乎所有流体的生产、加工、运输及其使用过程都使用压力管道。压力管道工程日益复杂，它正向大型化、整体化和自动化的方向发展。它在石油、化工、冶金等行业占有越来越重要的地位。

由于压力管道所输送的介质经常具有易燃、易爆、有毒、强腐蚀性及高温高压的特点，一旦出现泄漏、爆炸、燃烧和中毒事故，将给国民经济和人民生命财产造成重大损失。而另一方面，由于制造、运输、安装等原因，压力管道中的缺陷是不可避免的，若一有超标缺陷就返修或更换，不仅会带来很大的经济损失，而且可能引入新的缺陷，甚至可能造成更多的危害。同时，由于压力管道的失效模式，受力情况等多个方面与压力容器不同，因此有关压力容器安全评定的很多成果不能直接应用到压力管道的安全评定过程中。

因此从20世纪70年代开始，西方发达国家对压力管道做了大量的研究工作，例如美国退化管道研究计划、国际完整性研究计划、美国管道及其焊缝短裂纹研究计划、德国的管道研究计划和日本原子能研究所管道可靠性研究计划等。我国对压力管道的研究起步比较晚，90年代才开始进行系统地研究，尤其在“九五”期间全面系统地开展了有关压力管道完整性评定的大量研究工作。国内外对压力管道的完整性研究主要包括以下几个方面：

1）压力管道的失效模式和失效准则的研究；

2）压力管道安全评定规程；

3）压力管道的可靠性研究。

下面结合国内外对压力管道完整性评定的研究现状和我国在此方面的研究成果，介绍有关压力管道直管段平面缺陷安全评定方法。

压力管道周向面型缺陷安全评定采用了我国“九五”重点科技攻关创新性研究成果—— U 因子工程评定方法。这里主要采用了国家“九五”科技攻关专题（96-918-02-03）提出可适用于任意应力应变关系材料、任意材料断裂韧度、能自由选择安全系数的简化因子工程评定方法——一种以起裂为断裂判据，同时可以完成含周向面型缺陷压力管道起裂和塑性失效安全评定的 U 因子工程评定方法。该方法评定过程极为简便，只需要查两个表格进行简单的算术运算就可以完成拉、弯、扭、内压联合载荷作用下的周向面型缺陷安全评定。国际上至今还没有适应面如此之广、评定过程如此简便、且评定精度很高的工程评定方法。

这里主要给出了在内压、拉压、弯矩为主的组合载荷作用下，含裂纹、未熔合、未焊透等平面缺陷的钢制压力管道直管的安全评定方法

附加的特殊主要符号及其含义如下：

K_{eff}为有效应力强度因子，$N/mm^{3/2}$；K_{III}为撕裂型应力强度因子，$N/mm^{3/2}$；L_r^F 为起裂时的 L_r 值，量纲为1；M_B 为缺陷处不考虑缺陷存在时外载荷引起的弯矩，kN·m；M_c 为缺陷处不考虑缺陷存在时外载荷引起的扭矩，kN·m；N 为轴向拉力，kN；n_p 为规定按载荷取的安全系数，量纲为1；U 即 U 因子，为塑性极限载荷与起裂断裂载荷之比，量纲为1；θ 为缺陷周长包角为之半角，rad；θ/π 为缺陷的无因次长度，即周向缺陷长度与管周长之比值，量纲为1；$[\bar{\sigma}]$ 为许可流变应力比，即周长缺陷管道塑性极限载荷应力 $(\sigma_m+\sigma_B)_L$ 与材料流变应力 $\bar{\sigma}$ 值之比，量纲为1。

1）压力管道直管段轴向平面缺陷的评定　轴向平面缺陷，应按本章2.4中说明的方法进行评定。

2）压力管道直管段环向平面缺陷的评定　相关材料性能数据确定的特殊规定说明。

① 材料拉伸性能 σ_s 及 σ_b　对非匹配焊缝，取母材与焊材料二者的较低值。

② 评定用材料断裂韧度的下限值 $(J_{IC})_{下限}$ 可取 1.1 A_{KV}，也可根据评定者过去使用该种材料的历史经验决定。

③ 在缺乏在役材料的 A_{KV} 数据的情况下，如评定者能够判断该材料适用于压力管道，且在在役状况下无脆化倾向，使用温度不低于韧脆转变温度，则在评定计算时可取压力管道材料容许的最低值 $A_{KV}=27$ J。

④ 在使用上述断裂韧度值进行评定不能获得保证安全的结论时，可通过实物取样实测 A_{KV} 值或断裂韧度值，重新评定。

⑤ 断裂韧度 J_{IC} 的测定　对大直径的厚壁管，可按本章前面的有关规定进行；对小直径的薄壁管，在取样不能满足测试试件厚度要求时，可用该管材制备周向穿透裂纹的管段试件在四点弯曲加载下进行实测。

对在役工况下无脆化倾向的20钢或奥氏体不锈钢管道及壁厚不大于13 mm的16 Mn管道的未焊透缺陷允许采用压力管道直管段体积缺陷的安全评定方法进行评定。

⑥ 评定计算用 K_C 值，可由 J_{IC} 值按下式进行估算：

$$K_C=\sqrt{EJ_{IC}/(1-\nu^2)} \quad (5.4\text{-}119)$$

3）U 因子评定法

① 应力的计算　根据管系有限元分析确定缺陷处管道横截面的弯矩 M_B、拉力 N 和内压 P，按下式计算评定所需的轴向膜应力 σ_m 及弯曲应力 σ_B：

$$\sigma_m=\frac{N+\pi R_i^2 P}{2\pi RB} \quad (5.4\text{-}120)$$

$$\sigma_B=\frac{M_B}{\pi R^2 B} \quad (5.4\text{-}121)$$

② 起裂时载荷比 L_r^F 的确定　计算 $\sigma_s\sqrt{B}/K_C$，并根据 $\sigma_s\sqrt{B}/K_C$、缺陷无因次长度 θ/π 值和相对深度 a/B 值，由相关表格查得 L_r^F 值。

③ U 因子值的计算

$$U=\frac{\sigma_s+\sigma_b}{2L_r^F\sigma_s} \quad (5.4\text{-}122)$$

当 $U<1$ 时，取 $U=1$。

④ 许可流变力比 $[\bar{\sigma}]$ 的确定　由缺陷尺寸 θ/π 值及 a/B 值，由表5.4-8查得 $[\bar{\sigma}]$。

表 5.4-8　许可流变应力比 $[\bar{\sigma}]$

A/B	θ/π									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	1.254	1.236	1.221	1.211	1.206	1.206	1.206	1.206	1.206	1.206
0.2	1.234	1.196	1.165	1.143	1.131	1.129	1.129	1.129	1.129	1.129
0.3	1.213	1.156	1.107	1.07	1.048	1.041	1.041	1.041	1.041	1.041
0.4	1.193	1.114	1.045	0.992	0.957	0.942	0.942	0.942	0.942	0.942
0.5	1.172	1.071	0.981	0.909	0.858	0.832	0.827	0.827	0.827	0.827
0.6	1.15	1.027	0.914	0.821	0.753	0.712	0.697	0.697	0.697	0.697
0.7	1.128	0.981	0.844	0.729	0.64	0.583	0.564	0.549	0.549	0.549

4）安全性评价　如果式（5.4-123）所表述的判据成立，则评定结论为安全或可以接受；否则，为不能保证安全或不可接受。

$$(\sigma_m+\sigma_B)\leqslant\left(\frac{\sigma_s+\sigma_b}{2}\right)\frac{[\bar{\sigma}]}{Un_p}\qquad(5.4\text{-}123)$$

2.5　体积缺陷的确定性安全评定

前面介绍了平面缺陷的安全评定的三种方法，下面将介绍有关体积型缺陷的安全评定方法。与平面缺陷的评定过程相似，在进行体积型缺陷的安全评定时，首先要进行缺陷的表征，体积型缺陷包括凹坑、气孔及夹渣等，下面分别介绍对不同的体积型缺陷表征处理方法。

2.5.1　体积缺陷的表征及其处理

体积缺陷的表征主要包括凹坑缺陷以及气孔和夹渣缺陷的表征。

（1）单个凹坑缺陷的表征

按相关规定，对表面的不规则凹坑缺陷一般按其外接矩形将其规则化为长轴长度为 $2X$、短轴长度为 $2Y$ 及深度为 Z 的半椭球形凹坑。其中，长轴 $2X$ 为凹坑边缘任意两点之间的最大垂直距离，短轴 $2Y$ 为平行于长轴且与凹坑外边缘相切的两条直线间的距离，深度 Z 取凹坑的最大深度（如图 5.4-32 所示）。

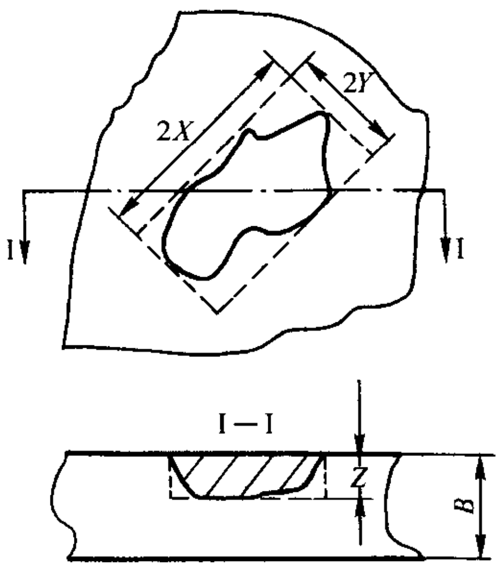


图 5.4-32　单个凹坑缺陷表征示意图

（2）多个凹坑缺陷的表征

在构件的表面通常存在多个缺陷，在处理多个凹坑缺陷的情况时，应考虑到它们之间的相互作用，因此当存在两个以上的凹坑时，首先要对每个凹坑分别按上述单个凹坑进行规则化的方法进行规则化处理，然后确定各自的凹坑长轴。若规则化后相邻两凹坑边缘间最小距离 k 大于较小凹坑的长轴 $2X_2$，则可将两个凹坑视为互相独立的单个凹坑分别进行评定。否则，应将两个凹坑合并为一个半椭球形凹坑来进行评定，该凹坑的长轴长度为两凹坑外侧边缘之间的最大距离，短轴长度为平行于长轴且与两凹坑外缘相切的任意两条直线之间的最大距离，该凹坑深度为两个凹坑的深度的较大值（见图 5.4-33）。

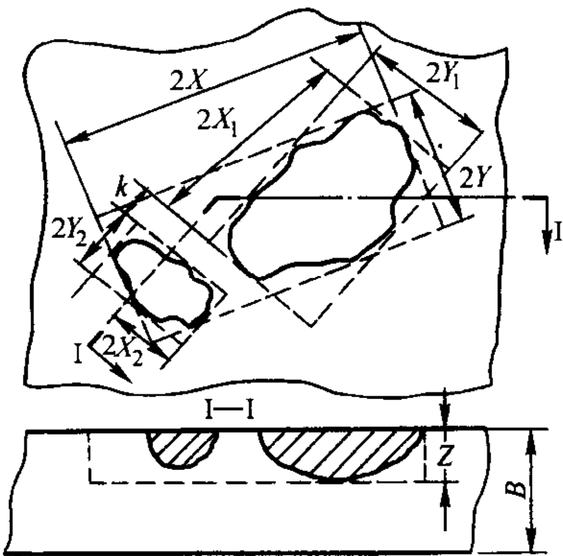


图 5.4-33　多个凹坑缺陷表征示意图

（3）气孔和夹渣缺陷的表征

1）气孔的表征　气孔用一般用来气孔率表征。气孔率是指在射线底片有效长度范围内，气孔投影面积占焊缝投影面积的百分比。射线底片有效长度按 JB/T 4730—1994《压力容器无损检测》的规定确定，焊缝投影面积为射线底片有效长度与焊缝平均宽度的乘积。

2）条形夹渣的表征　条形夹渣以其在射线底片上的长度来表征。当存在多个相邻夹渣时，同样应考虑夹渣间的相互影响，由于夹渣是体积型缺陷，夹渣间的空间位置可以通过是否共面来划分，因此这里介绍两种情况下的夹渣间的复合。

① 共面夹渣间的复合　若两个夹渣间的距离小于图 5.4-34 中的规定值，则将其复合为一个连续的大夹渣。

② 非共面夹渣的处理　当两个非共面埋藏夹渣之间的最小距离 s_3 小于较小夹渣的自身高度的一半时，则这两个夹渣可以视为共面并按①的规定进行复合。否则，均应逐个分别进行评定。

需要注意的是复合后的夹渣不再与其他夹渣或复合夹渣进行复合。

有时，为了简化起见，将气孔和夹渣视为平面缺陷进行评定，这是则应按本章 2.4.1 节的规定的方法进行缺陷的表征。

2.5.2　压力容器凹坑缺陷的确定性安全评定

这里首先介绍有关压力容器凹坑缺陷安全评定的基本方法和步骤（尤其对内压容器壳体表面）。在评定之前，首先要对缺陷区进行处理，将被评定缺陷打磨成表面光滑、过渡

平缓的凹坑，并确认凹坑及其周围无其它表面缺陷或埋藏缺陷。

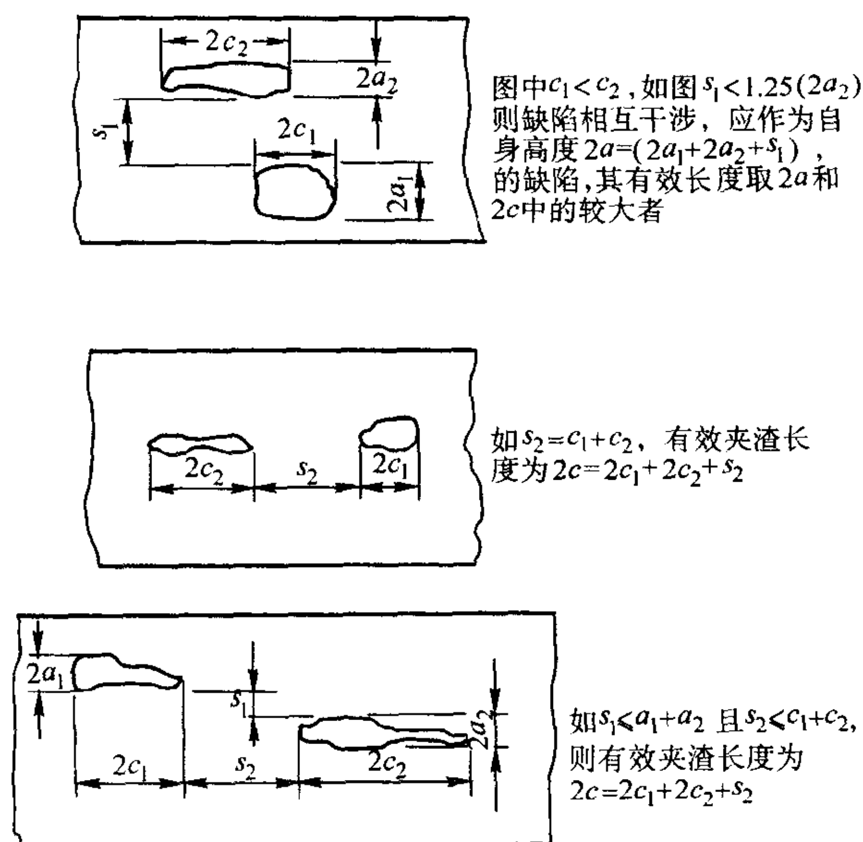


图 5.4-34　多个夹渣的复合准则图例

由于凹坑缺陷的安全评定是一个非常复杂的过程，这里介绍的方法不能对所有情况下的凹坑进行安全评定，而是具有一定的适用范围。该评定方法适用于下述情况的压力容器：

1) $B_0/R<0.18$ 的筒壳或 $B_0/R<0.10$ 的球壳，其中 B_0 为缺陷附近实测容器壳体壁厚，R 为容器平均半径；

2) 材料韧性满足压力容器设计规定，未发现劣化和退化情况；

3) 凹坑深度 Z 小于计算厚度 B 的 60%，且坑底最小厚度（$B-Z$）不小于 2 mm；

4) 凹坑长度 $2X\leqslant 2.8\sqrt{RB}$；

5) 凹坑宽度 $2Y$ 不小于凹坑深度 Z 的 6 倍（容许打磨至满足要求）。

而对于超出上述限定条件或在服役期间表面有可能生成裂纹的凹坑缺陷，目前建议按平面缺陷进行评定。

(1) 凹坑缺陷的评定程序（评定程序如图 5.4-35 所示）

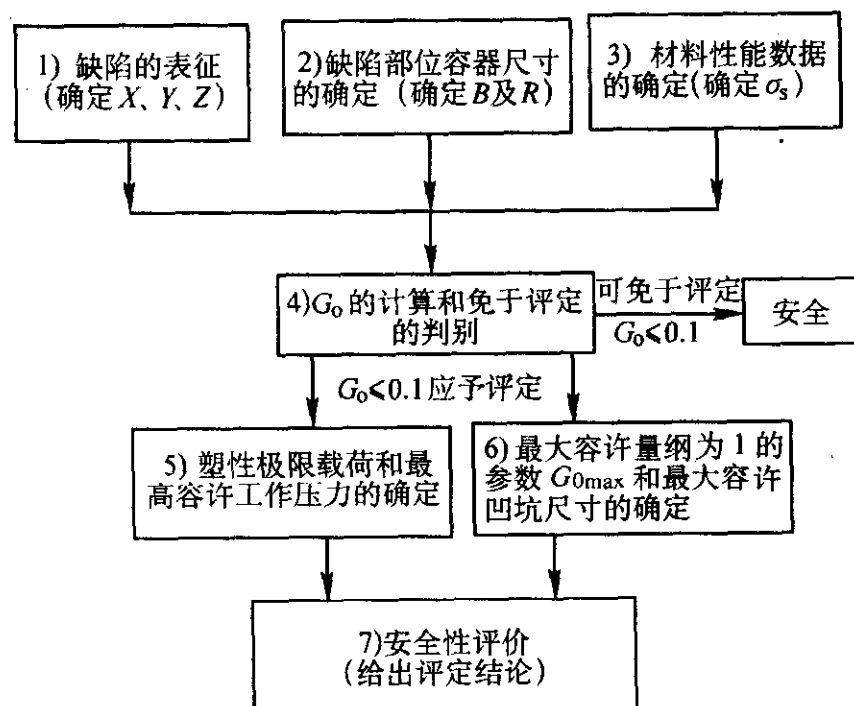


图 5.4-35　凹坑缺陷的评定程序示意图

凹坑缺陷的安全评定一般按下列步骤进行：

1) 缺陷的表征（确定 X，Y，Z 的值）；

2) 缺陷部位容器尺寸的确定（确定 B 和 R 的值）；

3) 材料性能数据的确定（确定 σ_s）；

4) 量纲为 1 的参数 G_0 的计算和免于评定的判别；

5) 塑性极限载荷和最高容许工作压力的确定；

6) 最大容许量纲为 1 的参数 G_{0max} 和最大容许凹坑尺寸的确定；

7) 安全性评价。

(2) 缺陷的表征与缺陷部位材料（容器）尺寸的确定

对通过各种检测手段检测查明的凹坑缺陷，根据其实际位置、形状和尺寸，按 2.5.1 节的方法将其规则化，并确定凹坑所在部位容器的计算厚度 B 和平均半径 R。

(3) 材料性能数据的确定

在进行凹坑的安全评定时主要采用屈服应力 σ_s 和材料流动应力 σ'。按本章 2.4 节中提供的方法，确定在评定工况下材料的屈服应力 σ_s。评定中所需的材料流动应力 σ' 可按下述方法选取：

$\sigma'=\sigma_s$　　用于非焊缝区凹坑

$\sigma'=\phi\sigma_s$　　用于焊缝区凹坑

其中 ϕ 为焊接接头系数，该值按容器的实际设计要求选取；当无法得到容器的设计要求时，也可按 GB 150—1998《钢制压力容器》或其他相关标准确定。

(4) 量纲为 1 的参数 G_0 的计算和免于评定的判别

如前所述，在对凹坑安全评定时，需要进行免于评定的判断，而在判断时需要计算量纲为 1 的参数 G_0。容器表面凹坑缺陷的量纲为 1 的参数 G_0 按式（5.4-124）计算。

$$G_0=\frac{Z}{B}\frac{X}{\sqrt{RB}} \tag{5.4-124}$$

若 $G_0\leqslant 0.1$，则该凹坑缺陷可免于评定，认为是安全的或可以接受的，不需要再做进一步的评定；

若 $G_0>0.1$，则该凹坑缺陷应按下述方法进行进一步的安全评定。

(5) 塑性极限载荷和最高容许工作压力的确定

由前面凹坑安全评定的步骤可知，若不能通过免于评定的判断，则需要进行塑性极限载荷和最高容许工作压力的确定。下面介绍几种情况下塑性极限载荷的计算方法。

1) 无凹坑缺陷壳体塑性极限载荷 P_{L0} 的计算

对球形容器　$$P_{L0}=2\bar{\sigma}'\ln\left(\frac{R+R/2}{R-B/2}\right) \tag{5.4-125}$$

对圆筒形容器　$$P_{L0}=\frac{2}{\sqrt{3}}\bar{\sigma}\ln\left(\frac{R+R/2}{R-B/2}\right) \tag{5.4-126}$$

式中，$\bar{\sigma}'$ 为评定温度下材料的流变应力。

2) 带凹坑缺陷容器极限载荷 P_L 的计算

对球形容器　$P_L=(1-0.6G_0)\ P_{L0}$　(5.4-127)

对圆筒形容器　$P_L=(1-0.3\sqrt{G_0})\ P_{L0}$　(5.4-128)

3) 带凹坑缺陷容器最高容许工作压力 P_{max}　带凹坑缺陷容器最高容许工作压力 P_{max} 按式（5.4-129）确定。

$$P_{max}=\frac{P_L}{1.8} \tag{5.4-129}$$

(6) 最大容许量纲为 1 的参数 G_{0max} 和最大容许凹坑尺寸的确定

由凹坑安全评定的步骤可知，在进行完塑性极限载荷和最高容许工作压力的确定时，需要确定最大容许量纲为 1 的参数 G_{0max} 和最大容许凹坑尺寸，这里分别对球形容器和圆筒形容器的情况进行简单的介绍。

1) 最大容许量纲为 1 的参数 G_{0max} 的确定

对球形容器　$G_{0max}=(1-1.8P/P_{L0})\ /0.6$　(5.4-130)

式中，P 为安全评定要求的容器工作压力。

对圆筒形容器　$G_{0max} = [(1-1.8P/P_{L0})/0.3]^2$　(5.4-131)

式中，P 为容器工作压力。若 $G_{0max} < 0.1$ 或 $p \geqslant 0.5p_{L0}$，取 $G_{0max} = 0.1$。

2) 最大容许凹坑尺寸 X_{max} 及 Z_{max} 的确定　凹坑最大容许长度（长轴）

$$2X_{max} = 2G_{0.max}B\sqrt{RB}/Z \ (X_{max} = G_{0.max})\ B\sqrt{RB}/Z \quad (5.4\text{-}132)$$

凹坑最大容许深度 $Z_{max} = G_{0.max}B\sqrt{RB}/Z$　(5.4-133)

(7) 安全性的评价

根据式（5.4-124）~式（5.4-133）的计算结果，符合下述情况之一，且实测凹坑尺寸满足本节开始时的规定条件要求，则认为该凹坑缺陷是安全的或可以接受的；否则，是不能保证安全或不可接受的。

1) $G_0 \leqslant 0.1$；

2) $P \leqslant P_{max}$；

3) $X \leqslant X_{max}$、$Z \leqslant Z_{max}$、$Y \leqslant 6Z$。

2.5.3　压力容器气孔和夹渣缺陷的确定性安全评定

(1) 评定方法与限定条件

这里主要介绍了内压容器壳体中气孔、夹渣缺陷安全评定的基本方法和步骤。该评定方法适用于符合下述条件的压力容器。

1) $B_0/R < 0.18$ 的压力容器；

2) 材料性能满足压力容器设计制造规定，且对铁素体钢，$\sigma_s < 450$ MPa，并且在最低使用温度下V形夏比冲击试验中3个试样的平均冲击功不小于40 J、最小冲击功不小于28 J；对其他材料，该气孔、夹渣所在处的 K_{IC} 大于1 250 $MPa \cdot mm^{1/2}$；

3) 未发现材料劣化和退化；

4) 气孔、夹渣未暴露于器壁表面；

5) 气孔、夹渣无明显扩展情况或可能；

6) 缺陷附近无其他平面缺陷。

对于暴露于器壁表面的气孔、夹渣可打磨消除。打磨成凹坑时，应按2.5.2节中提供的方法进行安全评定。对于超出本节中其他限定条件或在服役期间有可能生成裂纹的气孔、夹渣，建议按平面缺陷进行评定。

(2) 安全性评价

1) 气孔的安全性评价　如果在满足本节中前面所提出的限定条件，同时又能符合下列的条件，则该气孔是容许的；否则，是不可接受的。

① 气孔率不超过6%；

② 单个气孔的长径小于0.5 B，并且小于9 mm。

2) 夹渣的安全评定　若夹渣的尺寸满足本节中前面的限定条件和表5.4-9的要求，则该夹渣是容许的；否则，该缺陷是不能通过安全评定的。

表5.4-9　夹渣的容许尺寸

夹渣位置	夹渣尺寸的容许值	
球壳对接焊缝、圆筒体纵焊缝、与封头连接的环焊缝	总长度≤6 B	自身高度或宽度≤0.25 B，并且≤5 mm
	总长度不限	自身高度或宽度≤3 mm
圆筒体环焊缝	总长度≤6 B	自身高度或宽度≤0.30 B，并且≤6 mm
	总长度不限	自身高度或宽度≤3 mm

前面气孔和夹渣的评定过程中评为不可接受的气孔或夹渣，可表征为平面缺陷并按平面缺陷的评定方法重新进行安全评定。

2.5.4　压力管道体型缺陷确定性安全评定方法

压力管道体型缺陷安全评定采用了我国“九五”重点科技攻关创新性研究成果——含局部减薄缺陷压力管道塑性极限载荷工程评定方法。

与含凹坑压力容器安全评定方法相比，含局部减薄压力管道安全评定更为复杂，其原因如下：①压力管道除内压外还同时承受拉、弯等组合载荷；②安全评定所需要的管系内力和管道应力一般只能通过数值计算分析的方法求得，这对于一般的检验和评定人员来说具有很大的难度。

为解决上述难题，采用了以国家“九五”科技攻关专题(96-918-02-03）提出的以塑性极限载荷理论和含局部减薄压力管道塑性极限载荷拟合计算公式为基础的工程评定方法，尤其是采用了无须进行复杂的管系内力和管道应力计算，应用极为简便的免于评定条件的方法。

局部减薄缺陷安全评定方法理论严谨，方法简单，是我国首创的一种管道缺陷安全评定方法。将此方法偏保守地应用于含其他体型缺陷管道，提出了气孔、夹渣和失效模式为塑性失稳的未焊透等缺陷的安全评定方法。

压力管道直管段体积缺陷安全评定方法主要给出了在内压和弯矩的组合载荷作用下，含气孔、夹渣和壁厚局部减薄等体积缺陷及特定未焊透缺陷的钢制在役压力管道直管段的安全评定方法。

(1) 该评定方法的适用范围

该评定方法的适用范围主要从两个角度来考虑，一是对管道的适用范围，另一个是缺陷的适用范围。该评定方法适用于如下条件的钢制在用压力管道。

1) 在实际工况下，材料具有良好的延性，未发现劣化及劣化倾向；

2) 对于一般管材，管道最低工作温度不低于-20℃；若管道最低工作温度低于-20℃管材材料应为奥氏体不锈钢；

3) 管道径比 $R_0/R_1 \leqslant 1.4$，其中 R_1 为管道内半径，R_0 为管道外半径。

该评定方法适用于下列缺陷。

1) 气孔、夹渣、壁厚减薄、凹坑等体积型缺陷（其周围无裂纹、未熔合等面型缺陷）；

2) 在役工况下无脆化倾向的20钢或奥氏体不锈钢管道，或壁厚不大于13 mm的16Mn管道的未焊透缺陷；

3) 体积缺陷的深度 C 小于壁厚 T 的70%，且缺陷底部最小壁厚（$T-C$）不小于2 mm。

对于不符合上述限定条件或在服役期间表面有可能产生裂纹的含体积缺陷压力管道，应按平面缺陷评定方法进行评定。

(2) 管道安全评定程序

评定过程按下列步骤进行：

1) 缺陷部位管道尺寸的确定；

2) 缺陷的规则化；

3) 材料性能数据的确定；

4) 无缺陷管道在纯内压或纯弯矩下的塑性极限载荷的确定；

5) 进行免于评定的判别；

6) 典型工况载荷组合的确定与缺陷处弯矩的计算；

7) 含缺陷管道在纯内压或纯弯矩下的塑性极限载荷的确定；

8) 安全性评价。

(3) 缺陷部位管道尺寸的确定及缺陷的规则化、量纲为1化处理

根据实际情况确定缺陷附近管道的计算壁厚 T 和外径 R_0。管道的计算壁厚应取管道实测壁厚与至下一个检验周期末的腐蚀量之差。

如前评定方法的适用范围中提到的，该方法适用于气孔、夹渣、壁厚减薄、凹坑等体积型缺陷（其周围无裂纹、未熔合等面型缺陷），下面就分别介绍这几种不同缺陷的规则化问题。

1）壁厚局部减薄缺陷的规则化　对通过各种检测手段检测查明的壁厚局部减薄缺陷，根据其实际位置、形状和尺寸，按图5.4-36所示，将其规则化为轴向半长 A、环向半长 B、深度 C 的表面缺陷。其中 C 取实测最大减薄深度与至下一检验周期末体积缺陷扩展量之和。

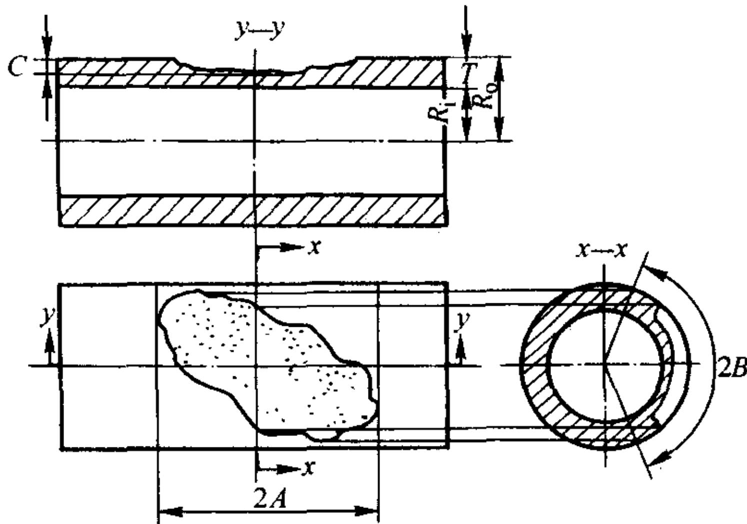


图5.4-36　局部减薄体积缺陷表征示意图

2）未焊透缺陷的规则化　对通过各种检测手段检验查明的未焊透缺陷（仅适用于符合前面限定条件所说明的未焊透），根据其实际位置、形状和尺寸，按图5.4-37所示，将其规则化为轴半长 A、环向半长 B、深度 C 的表面缺陷。

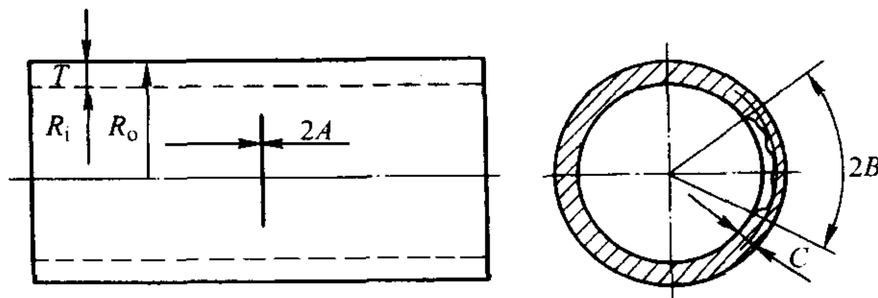


图5.4-37　未焊透缺陷表征示意图

3）气孔、夹渣缺陷的规则化　对于通过各种检测手段检验查明的气孔和夹渣缺陷，首先应按本章2.5.3节的方法分别用气孔率和夹渣长度进行表征。如表征后的气孔、夹渣缺陷不能通过本节中所提到的免于评定判断，则应根据其实际位置、形状和尺寸，按图5.4-38所示，将其规则化为轴向半长 A、环向半长 B、深度 C 的表面缺陷，其中若能够获得可靠的缺陷自身高度 H 值，则 $C=H$，否则可近似取 $C=2A$。

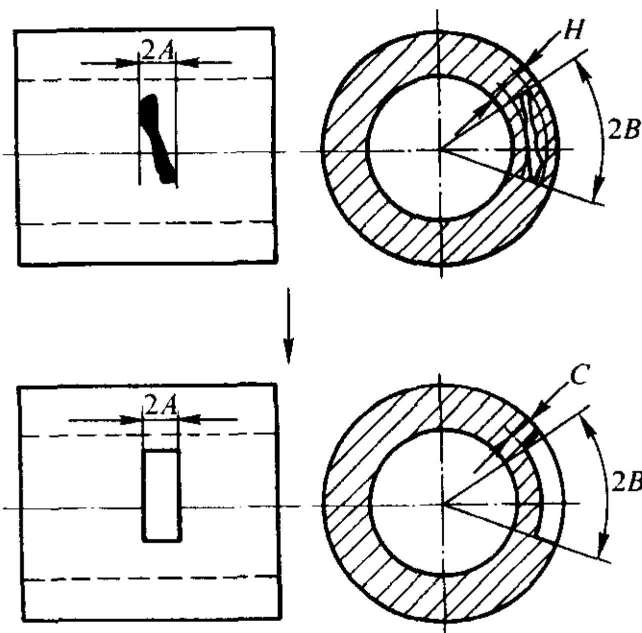


图5.4-38　气孔、夹渣等埋藏缺陷表征示意图

4）缺陷规则化后的尺寸的量纲为1化处理　对经规则化处理得到的缺陷尺寸，需按如下方法进行量纲为1化处理。

相对轴向长度 $a=A/\sqrt{R_0/T}$；

相对环向长度 $b=B/(\pi R_0)$（对于外表面缺陷）；

$b=B/(\pi R_1)$（对于内表面缺陷）；

相对深度 $c=C/T$。

（4）材料性能数据的确定

在进行气孔、夹渣缺陷的安全评定时，需要确定一些材料性能参数，如屈服强度 σ_s 和流动应力 $\bar{\sigma}'$。按本章所介绍的方法或参考相关规定确定在评定工况下材料的屈服强度 σ_s，并按下述方法确定材料的流动应力 σ'：

$$\begin{cases}\bar{\sigma}'=\sigma_s & \text{用于非焊缝区}\\ \bar{\sigma}'=\phi\sigma_s & \text{用于焊缝区}\end{cases}$$

其中，焊接接头系数 ϕ 按管道实际设计制造要求选取，也可参照压力管道或压力容器的有关偏保守地选取。

无缺陷管道在纯内压下的塑性极限内压 P_{L0} 和在纯弯矩下的塑性极限弯矩 M_{L0} 按下式进行计算。

$$P_{L0}=\frac{2}{\sqrt{3}}\bar{\sigma}'\ln\frac{R_0}{R_i} \tag{5.4-134}$$

$$M_{L0}=4\bar{\sigma}'\frac{R_0^3+R_1^3}{3} \tag{5.4-135}$$

（5）免于评定的判别

同凹坑的评定过程相似，气孔夹渣缺陷的评定过程中也有免于评定的判别，若满足此判别，则不需要做进一步的评定，认为是安全的，否则需要做进一步的评定。

1）气孔免于评定的判别　如果同时满足下列条件，则该气孔可免于评定。

① 气孔率不超过6%；

② 单个气孔的长径小于0.5 T，并且小于9 mm。

2）夹渣免于评定的判别　如果夹渣的自身高度或宽度不大于0.4 T，则该夹渣可免于评定。

3）壁厚局部减薄和未焊透缺陷免于评定的判别　如果同时满足下列条件，则该局部减薄或未焊透缺陷可免于评定。

① 管道结构符合设计制造要求；

② $P_{max}/P_{L0}\leqslant 0.3$；

③ 局部减薄或未焊透相对深度 c 不超过表5.4-10规定的容阻值。

表5.4-10　压力管道壁厚局部减薄与未焊透缺陷容限值

缺陷的相对环向长度 b	$b\leqslant 0.25$	$0.25<b\leqslant 0.50$	$0.50<b\leqslant 0.75$	$0.75<b\leqslant 1.00$
容许的最大缺陷相对深度 c	0.45	0.40	0.35	0.25

注：这里所说的未焊透需要满足本节前面的评定范围。

（6）典型工况载荷组合的确定与缺陷处弯矩载荷的确定

根据实际情况，确定需要进行缺陷评定的载荷工况，给出相应的内压载荷 P，并通过有限元方法或其他可靠的方法确定相应评定工况下缺陷处理的弯矩载荷 M。

（7）含缺陷管道在纯内压或纯弯矩下的塑性极限载荷的确定

1）纯内压下的塑性极限内压

$$P_{LS}=P_{LS}\times P_{L0} \tag{5.4-136}$$

式中

$$P_{LS}=\begin{cases}0.95-0.85A_e & a/b\leqslant 7.0\\ 0.95-1.04A_e & 7.0<a/b\leqslant 25.0\\ 0.95-1.47A_e & a/b>25.0\end{cases} \tag{5.4-137}$$

$$A_e = c^3\sqrt{a_e bc} \quad (5.4\text{-}138)$$

$$a_e = \min(3.0,\ a) \quad (5.4\text{-}139)$$

2）纯弯矩下的塑性极限弯矩

$$M_{LS} = m_{LS} \times M_{L0} \quad (5.4\text{-}140)$$

式中

$$m_{LS} = \begin{cases} \cos\left(\dfrac{c\pi b}{2}\right) - \dfrac{c\sin(\pi b)}{2} & c < \dfrac{1-b}{b} \\ (1-c)\sin\left[\dfrac{\pi(1-bc)}{2(1-c)}\right] + \dfrac{c\sin(\pi b)}{2} & c \geqslant \dfrac{1-b}{b} \end{cases} \quad (5.4\text{-}141)$$

(8) 安全性评价

通过对过程（6）的分析计算可以获得工况载荷组合和缺陷处弯矩载荷，再经过工程（7）中式（5.4-136）和式（5.4-140）的计算可以获得塑性极限内压和极限弯矩，若以上计算结果能使式（5.4-142）成立，则该缺陷是安全或可以接受的；否则，认为不能保证安全或不可接受。

$$\left(\frac{P}{P_{LS}}\right)^2 + \left(\frac{M}{M_{LS}}\right)^2 \leqslant 0.44 \quad (5.4\text{-}142)$$

对于本节所讨论的气孔、夹渣等埋藏型体积缺陷，如未能通过本节所规定的安全评定计算，可将其表征为埋藏平面缺陷并按本章2.4.7节中说明方法重新进行评定。

3 疲劳失效确定性评定技术和方法

除了断裂和塑性失效安全评定技术和方法外，疲劳失效评定技术和方法在国内外的研究和应用也比较成熟。而且通过“八五”的攻关研究，使该项评定技术又有了进一步的发展。

由本篇第2章的内容可知，疲劳是比较常见的一种破坏形式。传统的疲劳强度设计，是假设材料的初始状态为连续体，不存在裂纹，经过一定的应力循环后，由于材料的损伤积累，才形成了裂纹。裂纹形成前的应力循环，称为无裂纹寿命。零部件的总寿命 N 由两部分组成，即无裂纹寿命 N_0 和裂纹扩展寿命 N_p。如本篇第3章所述一般可把疲劳裂纹扩展的情况分为三个阶段，即阶段Ⅰ，疲劳裂纹萌生与微观缓慢扩展阶段，阶段Ⅱ宏观裂纹扩展阶段（裂纹稳定扩展阶段）和阶段Ⅲ快速扩展阶段。

由于疲劳失效评定技术是建立在疲劳性能分析和疲劳寿命分析的基础上的，因此首先要进行与之相关的讨论，再介绍相关的疲劳失效评定技术。

3.1 材料疲劳性能参数的获取

经典疲劳理论是建立在应力幅度与疲劳循环次数 $S-N$ 曲线基础上的，在 $S-N$ 曲线基础上引进存活率或失效概率可得 $P-S-N$ 曲线，进而考虑平均应力影响即可得 $P-S_a-S_m-N$ 曲面，再进一步考虑可靠度的置信度可得 $r-P-S_a-S_m-N$ 曲面。

在疲劳试验的基础上可作出表示应力与寿命相互关系的 $S-N$ 曲线。由于材料本身性能的个体差异和分散性，致使其试验获得的通常不止是一条 $S-N$ 曲线，而是一族 $S-N$ 曲线，甚至可以说是一个 $S-N$ 曲面。

3.1.1 $P-S-N$ 曲线

$P-S-N$ 曲线的制作方法在第3章中已经进行了详细的介绍，这里不再赘述。

3.1.2 影响 $P-S-N$ 曲线的因素

如第3章所述，材料成分、晶粒尺寸大小和晶粒方向、热处理状态、焊接等都会影响疲劳强度，在上述条件基本相同时，构件与光滑小试件主要差异在于构件外形的差异、尺寸大小差异和表面质量的差异。综合考虑上述三种因素，构件在对称循环下的疲劳强度为

$$\sigma_e^0 = \frac{\varepsilon_\sigma \beta}{K_\sigma}\sigma_e$$

式中，σ_e 为光滑小试件的疲劳极限；ε_σ 为对称循环下的尺寸系数，由于构件尺寸大，形成疲劳裂纹机会也更多，所以 $\varepsilon_\sigma < 1$；β 为对称循环下的表面质量系数，当构件表面质量低于小试件时，$\beta < 1$，而表面经强化后，$\beta > 1$；K_σ 为对称循环下的有效应力集中系数，一般 $K_\sigma > 1$。

实验证明：构件应力集中、尺寸大小和表面质量等因素，只对属于动应力的应力幅度 σ_a（或用 S_a 表示）有影响，而对属于静应力的平均应力 σ_m（或用 S_m 表示）并无影响。

3.1.3 $P-S_a-S_m-N$ 曲面

如果考虑到平均应力 S_m 的影响，需要将 $P-S-N$ 曲线扩展为 $P-S_a-S_m-N$ 曲线。其中 S_a 表示应力幅度。

在应力幅度 S_a、平均应力 S_m 和疲劳寿命 N 的三维空间中，S_a、S_m 和不同存活率 p 的寿命 N_p [即 $p(N > N_p) = p$]，三者之间用曲面表示，曲面方程的一般形式为

$$f = (N_p,\ S_a,\ S_m) = C_p$$

式中，C_p 为对应存活率 p 的常数。

1）当 S_m 为常数时，曲面方程变成 $f(N_p,\ S_a) = C'_p$（常数），退化为 $P-S_a-N$ 曲线，尤其是 $S_m = 0$ 时，退化为 $P-S-N$ 对称循环曲线。对称循环时，$P-S-N$ 曲线一般表达式为

$$N(S_a - S_0)^\beta = \alpha \quad (5.4\text{-}143)$$

式中，S_0、β、α 为材料常数，当 $S_0 = 0$ 时

$$NS_a^{\beta'} = \alpha'$$

式中，α'、β' 为材料常数。

2）当 N_p 为常数时，曲面方程变成 $f(S_a,\ S_m) = C'_p$，即退化为给定寿命 N_p 的等寿命曲线。等寿命曲线一般是由平均应力 S_m、应力幅度 S_a 以及材料常数 σ_b 或者 σ_s 之间建立的关系来表示的，有代表性的有

$$\text{Goodman 关系：}\ \frac{S_a}{S_{-1}} + \frac{S_m}{\sigma_b} = 1 \quad (5.4\text{-}144)$$

$$\text{Gerber 关系：}\ \frac{S_a}{S_{-1}} + \left(\frac{S_m}{\sigma_b}\right)^2 = 1 \quad (5.4\text{-}145)$$

$$\text{Soderbery 关系：}\ \frac{S_a}{S_{-1}} + \frac{S_m}{\sigma_b} = 1 \quad (5.4\text{-}146)$$

式中，S_{-1} 为曲线 $P-S-N$ 上给定寿命 N_p 的对称循环疲劳强度。在 $N = N_p$ 等寿命曲线上，如果外载应力点（S_a，S_m）落在阴影区时，发生失效，即发生 $N \leqslant N_p$ 事件，如图5.4-39所示。

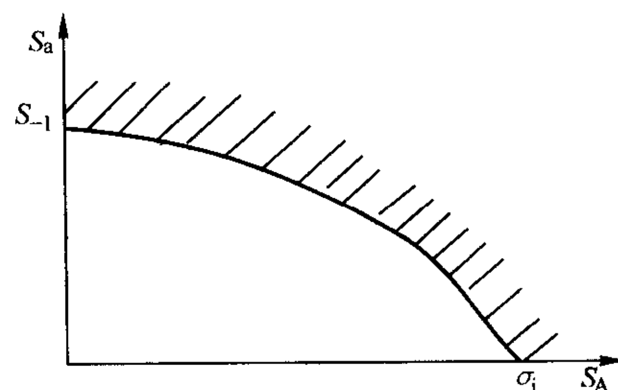


图5.4-39 等寿命曲线示意图

将式（5.4-144）和式（5.4-145）代入式（5.4-143）中，得等寿命曲线推出的 $P-S_a-S_m-N$ 的曲线方程［此时式（5.4-144）中或式（5.4-145）中的 S_{-1} 就是式（5.4-143）中的 S_a］：

$$N_p = \left(\frac{\sigma_{bp}}{\sigma_{bp} - S_m}\cdot S_a - S_{0p}\right)^{\beta p} = \alpha_p \quad (5.4\text{-}147)$$

或

$$N_p=\left(\frac{\sigma_{bp}^2}{\sigma_{bp}^2-S_m^2}S_a-S_{0p}\right)^{\beta_p}=\alpha_p \quad (5.4\text{-}148)$$

式中，σ_{bp}、β_p、α_p、S_{0p}是待定常数，与存活率 p 有关。有时为了方便起见采用

$$N_p=\left(\frac{\sigma_{bp}}{\sigma_b-S_m}S_a-S_{0p}\right)^{\beta_p}=\alpha_p \quad (5.4\text{-}149)$$

即与存活率 p 有关常数只有两个，即 β_p、α_p。而 S_0、σ_b 与存活率无关。

3.1.4　$P-S_a-S_a-N$ 曲面的矩法测定

与在平均应力 $S_m=0$ 情况下测定 $S-N$ 曲线相同，在不同平均应力 S_m 水平下进行4～5个成组试验和一个升降法试验，可得到各种不同应力幅度和平均应力 S_m 组合下的疲劳寿命分布。其中升降法是为了得到疲劳强度的近似分布。

对于 n 个 (S_{a_i}, S_{m_i}) $(i=1, 2, \cdots, n)$，我们可以得到应力 (S_{a_i}, S_{m_i}) 作用下，疲劳寿命分布 $\lg N$ 的均值 μ_i 和标准差 σ_i，因此可计算该点给定存活率 p 条件下寿命 $\lg N_{P_i}$。设 $P-S_a-S_a-N$ 曲面形式为

$$N_p\left(\frac{\sigma_{bp}}{\sigma_{bp}-S_m}S_a-S_{0p}\right)^{\beta_p}=\alpha_p$$

即

$$\lg N_p=\lg\alpha_p-\beta_p\lg\left(\frac{\sigma_{bp}}{\sigma_{bp}-S_m}S_a-S_{0p}\right) \quad (5.4\text{-}150)$$

令

$$w=\sum_{i=1}^{n}(\lg N_p-\lg N_{p_i})^2=\sum_{i=1}^{n}\left\{\left[\lg\alpha_p-\beta_p\lg\left(\frac{\sigma_{bp}}{\sigma_{bp}-S_{mi}}S_{a_i}-S_{0p}\right)\right]-\lg N_{p_i}\right\}^2$$

按最小二乘法，令

$$\frac{\partial w}{\partial(\lg\alpha_p)}=0,\ \frac{\partial w}{\partial\beta_p}=0,\ \frac{\partial w}{\partial\sigma_{bp}}=0,\ \frac{\partial w}{\partial S_{0p}}=0$$

可以得到使残差平方和 w 最小的给定存活率 p 条件下 $P-S_a-S_a-N$ 曲面的最优表达式。

当 $p=50\%$时，表示平均曲面，略去下标 p，并记平均寿命为 $\mu(S_a, S_m)$，则

$$\mu(S_a, S_m)=\lg N_{50}=\lg\alpha-\beta\lg\left(\frac{\sigma_b}{\sigma_b-S_m}S_a-S_0\right)$$

当 $p=84.1\%$时，记下标为1，则

$$\lg N_{84.1}=\lg\alpha_1-\beta_1\lg\left(\frac{\sigma_{b1}}{\sigma_{b1}-S_m}S_a-S_{01}\right)$$

如果 $\lg N$ 服从正态分布，即 N 服从对数正态分布，则有

$$\sigma(S_a, S_m)=\mu(S_a, S_m)-\lg N_{84.1}=\lg\frac{\alpha}{\alpha_1}-\beta\lg\left(\frac{\sigma_b}{\sigma_b-S_m}S_a-S_0\right)+\beta_1\lg\left(\frac{\sigma_{b1}}{\sigma_{b1}-S_m}S_a-S_{01}\right)$$

有了均值曲面 $\mu(S_a, S_m)$ 和标准差 $\sigma(S_a, S_m)$，就可以求任意存活率 p 条件下 $P-S_a-S_a-N$ 曲面。

$$\lg N_p=\mu(S_a, S_m)-\mu_{1-p}\sigma(S_a, S_m) \quad (5.4\text{-}151)$$

式中，μ_{1-p}为标准正态分布的上侧分位数。

3.1.5　$\gamma-P-S_a-S_a-N$ 曲面

对于每个 (S_{a_i}, S_{m_i}) $(i=1, 2, \cdots, n)$ 处的对数寿命分布 $\lg N_i$ 的均值μ_i 和标准差σ_i 是与成组试验样本容量 m 有关，样本容量 m 越大，均值 μ_i 和σ_i 标准差估计越准确，该点处存活率为 p 条件下寿命 $\lg N_i$ 估计越准确。

如果在每个 (S_{a_i}, S_{m_i}) $(i=1, 2, \cdots, n)$ 处，选择单侧容忍下限 $L_{p\gamma}$，满足：

$$P[P(\lg N>L_{P\gamma})\geqslant p]=\gamma \quad (5.4\text{-}152)$$

或

$$P[P(N>10^{L_{P\gamma}})\geqslant p]=\gamma \quad (5.4\text{-}153)$$

即以置信度 γ 保证 $\lg N>L_{P\gamma}$的存活率大于或等于概率 p。把 $N=10^{L_{P\gamma}}$的寿命点连接起来，所拟合而成的曲面就是 $\gamma-P-S_a-S_a-N$ 曲面。它是以置信度 γ 保证对数寿命存活率大于或等于p 的曲面。这就保证了尽管样本容量 m 不同，但我们总能找到置信度为 γ 的存活率为p 的对应不同m 的 $P-S_a-S_a-N$ 曲面。单侧容忍下限 $L_{P\gamma}$与置信度 γ、存活率 p 和样本容量m 有关，具体求法可参考相关文献中的近似求法。下面介绍求单侧容忍下限的 $L_{P\gamma}$二阶矩近似法。

在指定每个 (S_{a_i}, S_{m_i}) $(i=1, 2, \cdots, n)$ 处，设对数疲劳寿命服从正态分布，令随机变量 $Y(S_{a_i}, S_{m_i})=\lg N(S_{a_i}, S_{m_i})$，那么 $Y(S_{a_i}, S_{m_i})\sim N[\mu(S_{a_i}, S_{m_i}), \sigma^2(S_{a_i}, S_{m_i})]$。

对于均值 $\mu(S_{a_i}, S_{m_i})$ 和方差 $\sigma^2(S_{a_i}, S_{m_i})$ 估计分别为：

$$\bar{Y}(S_{a_i},S_{m_i})=\frac{1}{m}\sum_{i=1}^{m}y_i(S_{a_i},S_{m_i})$$

$$S(S_{a_i},S_{m_i})=\sqrt{\frac{1}{N-1}\sum_{i=1}^{m}[y_i(S_{a_i},S_{m_i})-\bar{Y}(S_{a_i},S_{m_i})]^2}$$

式中，$y_i(S_{a_i}, S_{m_i})$ $(i=1, 2, \cdots, n)$ 为 $Y(S_{a_i}, S_{m_i})$ 的 m 个观测值，即成组试验的样本值。所以对总体 p 分位值的估计为

$$\hat{Y}_p(S_{a_i}, S_{m_i})=\bar{Y}(S_{a_i}, S_{m_i})+\mu_p\hat{\sigma}(S_{a_i}, S_{m_i}) \quad (5.4\text{-}154)$$

式中，标准差无偏估计值为 $\hat{\sigma}=(S_{a_i},S_{m_i})=\beta S\ (S_{a_i},S_{m_i})$，$\beta\approx1+\frac{1}{4(n-1)}$ $(n>1)$，即 $E[\hat{\sigma}(S_{a_i}, S_{m_i})]=\sigma(S_{a_i}, S_{m_i})$，$\mu_p$ 为标准正态分布的上侧分位数。

由于 $\bar{Y}(S_{a_i}, S_{m_i})=S(S_{a_i}, S_{m_i})$，为样本函数是随机变量，所以 $\hat{Y}(S_{a_i}, S_{m_i})$ 也是样本函数为随机变量，根据统计理论可知：

$$E[\bar{Y}(S_{a_i},S_{m_i})]=\mu(S_{a_i},S_{m_i})$$

$$D[\bar{Y}(S_{a_i},S_{m_i})]=\sigma^2(S_{a_i},S_{m_i})/m$$

$$E[\hat{\sigma}(S_{a_i},S_{m_i})]=\sigma(S_{a_i},S_{m_i})$$

$$\begin{aligned}D[S(S_{a_i},S_{m_i})]&=E[S(S_{a_i},S_{m_i})]^2-\{E[S(S_{a_i},S_{m_i})]\}^2\\&=\sigma^2(S_{a_i},S_{m_i})-\sigma^2(S_{a_i},S_{m_i})/\beta^2\\&=\sigma^2(S_{a_i},S_{m_i})(1-1/\beta^2)\end{aligned}$$

由此可得随机变量 $\hat{Y}(S_{a_i}, S_{m_i})$ 的均值和方差分别为

$$E[\hat{Y}_p(S_{a_i},S_{m_i})]=\mu(S_{a_i},S_{m_i})+\mu_p\sigma(S_{a_i},S_{m_i})$$

$$D[\hat{Y}_p(S_{a_i},S_{m_i})]=\frac{\sigma^2(S_{a_i},S_{m_i})]}{m}+\mu_p^2(\beta^2-1)\sigma^2(S_{a_i},S_{m_i})$$

如果假设 $\hat{Y}(S_{a_i}, S_{m_i})$ 服从正态分布，可得标准正态变量

$$U=\frac{\hat{Y}_p(S_{a_i}, S_{m_i})-E[\hat{Y}_p(S_{a_i}, S_{m_i})]}{\sqrt{D[\hat{Y}_p(S_{a_i}, S_{m_i})]}}=\frac{[\bar{Y}_p(S_{a_i}, S_{m_i})+u_p\beta S(S_{a_i}, S_{m_i})]-[\mu(S_{a_i}, S_{m_i})+u_p\sigma(S_{a_i}, S_{m_i})]}{\sigma(S_{a_i}, S_{m_i})\sqrt{1/m+u_p^2(\beta^2-1)}}$$

根据统计知识：$\dfrac{(m-1)\ S^2\ (S_{a_i},\ S_{m_i})}{\sigma^2\ (S_{a_i},\ S_{m_i})} \sim X^2\ (m-1)$，且 $S^2\ (S_{a_i},\ S_{m_i})$ 与 U 基本相互独立，所以 t 统计量为

$$t=\frac{U}{\sqrt{X^2\ (m-1)\ /\ (m-1)}}=\frac{[\bar{Y}_b\ (S_{a_i},\ S_{m_i})\ +\mu_p\beta S\ (S_{a_i},\ S_{m_i})]\ -\ [\mu\ (S_{a_i},\ S_{m_i})\ +\mu_p\sigma\ (S_{a_i},\ S_{m_i})]}{\sigma\ (S_{a_i},\ S_{m_i})\ \sqrt{1/m+\mu_p^2\ (\phi^2-1)}}$$

(5.4-155)

设 t_γ 为 t 分布的下侧分位数，即 $P\ (t<t_\gamma)\ =\gamma$，将式（5.4-155）代入得

$$P\ \{\bar{Y}\ (S_{a_i},\ S_{m_i})\ +\mu_p\beta S\ (S_{a_i},\ S_{m_i})\ -t_\gamma S\ (S_{a_i},\ S_{m_i})\ \times \sqrt{1/m+\mu_p^2(\beta^2-1)}<[\mu(S_{a_i},S_{m_i})+\mu_p\sigma(S_{a_i},S_{m_i})]\}=\gamma$$

显然，在给定置信度 γ 条件下，可靠度为 p 的对数疲劳寿命单侧置信度下限为

$$L_{P\gamma}\ (S_{a_i},\ S_{m_i})\ =[\bar{Y}\ (S_{a_i},\ S_{m_i})\ +\mu_p\beta S\ (S_{a_i},\ S_{m_i})\ -t_\gamma S\ (S_{a_i},\ S_{m_i})\ \sqrt{1/m+\mu_p^2\ (\beta^2-1)}]$$

满足

$$P\ [P\ (N>10^{L_{P\gamma}})\ \geqslant p]\ =\gamma$$

所以，$\gamma-P-S_a-S_a-N$ 曲面估计式为

$$N_{P\gamma}=10^{L_{P\gamma}} \tag{5.4-156}$$

式中，$\bar{Y}\ (S_{a_i},\ S_{m_i})$ 和 $S\ (S_{a_i},\ S_{m_i})$ 可以先拟合为曲面，并用解析式表示，例如，采用本章 3.1.4 节提到的方法确定均值曲面和标准差。

3.2　应力变动和累积损伤

3.2.1　应力变动和累积损伤概念

前面提到的疲劳曲线（或曲面）都是恒幅载荷实验得到的，然而多数构件的真实使用条件是承受变动应力，甚至是随机的动载荷，因此存在变循环应力的累积损伤。所谓累积损伤是指：用于循环加载而引起的疲劳损伤累积，材料本身发生变化，因而实际上存在一条新的位置下移 $S-N$ 曲线（或 $P-S-N$ 曲线），即 $S-N$ 曲线是随着应力循环次数的增加而损伤的积累不断下移的。

如图 5.4-40 所示，循环周次造成损伤越大，剩余寿命就越小，即有 $N_4>N_3>N_2>N_1$。损伤越大，给定寿命条件下所剩余强度就越低，$S_4>S_3>S_2>S_1$，断裂失效可能性就越大。因此，在进行疲劳条件下的安全评定，必须考虑累积损伤的效应。

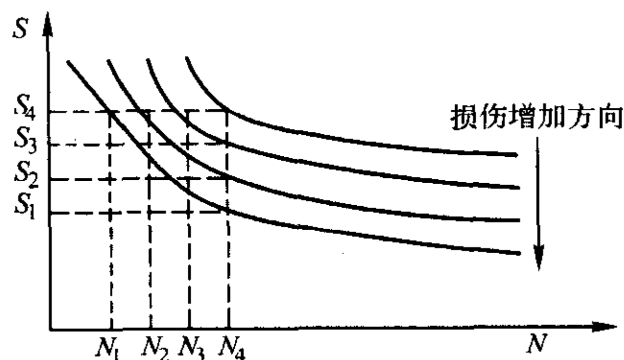


图 5.4-40　累积疲劳损伤示意图

3.2.2　累积损伤

如第 3 章所述对于累积损伤的研究目前得到公认的理论之一是 Miner 理论。Palmgren－Miner 理论认为，在任意应力水平 S_i 下的损伤率为：

$$D_i=n_i/N_i \tag{5.4-157}$$

写成微分形式为：

$$\mathrm{d}D=\mathrm{d}n/N \tag{5.4-158}$$

式中，n_i（或 dn）为循环次数；N_i（或 N）为该应力水平下疲劳寿命。

如果

$$\sum_{i=1}^{k} n_i/N_i\ =1 \tag{5.4-159}$$

或写成积分形式为

$$\int \mathrm{d}D\ =1 \tag{5.4-160}$$

则在变动应力 S_i（$i=1,\ 2,\ \cdots,\ k$）作用下预计会出现疲劳断裂失效。

但这个理论的主要缺点是没有考虑各种不同应力水平的作用顺序的影响。实验表明，如果承受应力为递增应力顺序，则 Miner 数 $\sum_{i=1}^{k} n_i/N_i>1$；如果承受应力为递减应力顺序，则 Miner 数 $\sum_{i=1}^{k} n_i/N_i<1$。根据所采用的是递减型还是递增型应力幅，失效时 Miner 数大约在 1/4～4 之间。如果是各个循环应力幅以一种随机的顺序混合作用，则 Miner 数接近 1。

如果考虑到平均应力 S_m 的影响，可以定义二维 Miner 公式为：

$$\sum_{i=1}^{k}\frac{n_i(S_a,S_m)}{N_i(S_a,S_m)}=1 \tag{5.4-161}$$

或

$$\int_{(S_m)_{min}}^{(S_m)_{max}}\int_{(S_a)_{min}}^{(S_a)_{max}}\frac{nf(S_a,S_m)}{N(S_a,S_m)}\mathrm{d}S_a\mathrm{d}S_m=1 \tag{5.4-162}$$

式中，$n_i\ (S_a,\ S_m)$ 为应力水平 $(S_a,\ S_m)$ 的循环数；$N_i\ (S_a,\ S_m)$ 为应力水平 $(S_a,\ S_m)$ 的疲劳寿命；$f\ (S_a,\ S_m)$ 为二元随机变量 $(S_a,\ S_m)$ 的联合概率密度函数；$N\ (S_a,\ S_m)$ 为应力水平 $(S_a,\ S_m)$ 单独作用下疲劳寿命；n 为随机变量 $(S_a,\ S_m)$ 作用下的疲劳寿命。

有关疲劳累积损伤的另一理论 Marin 累积损伤理论考虑到了小应力造成的损伤较小的特点。在任意应力水平 S_i 下的损伤率为：

$$D=\frac{n_i}{N_i}\left(\frac{S_i}{S_{max}}\right)^d$$

式中，n_i 为在应力 S_i 作用下循环次数；N_i 为在应力 S_i 作用下的疲劳寿命；S_{max} 为最大应力幅度。

当 $d=0$ 时，退化为 miner 损伤理论情形。写成微分形式为

$$D=\frac{\mathrm{d}n}{N}\left(\frac{S_i}{S_{max}}\right)^d$$

如果 $\sum_{i=1}^{k} D_i\ =1$ 或 $\int \mathrm{d}D\ =1$，则预计会出现失效（或断裂）。

3.2.3　变幅载荷作用下疲劳寿命分析

设任意载荷作用下，应力幅度 S_a 和平均应力 S_m 的联合概率密度为 $f\ (S_a,\ S_m)$，在应力水平 $(S_a,\ S_m)$ 的概率为 $f_a\ (S_a,\ S_m)\ \mathrm{d}S_a\mathrm{d}S_m$，令 N_p^* 给定存活率 p 条件下的寿命，则构件在应力水平 $(S_a,\ S_m)$ 内的循环次数为

$$\mathrm{d}n=N_p^*f\ (S_a,\ S_m)\ \mathrm{d}S_a\mathrm{d}S_m \tag{5.4-163}$$

表征材料疲劳性能的 $P-S_a-S_a-N$ 曲面为

$$N_p\left(\frac{\sigma_{bp}}{\sigma_{bp}-S_m}S_a-S_{0p}\right)^{\beta_p}=\alpha_p \tag{5.4-164}$$

因此，在应力水平 $(S_a,\ S_m)$ 作用下，存活率为 p 的

疲劳寿命为

$$N_p = \alpha_p \Big/ \left(\frac{\sigma_{bp}}{\sigma_{bp} - S_m} S_a - S_{0p} \right)^{\beta_p} \quad (5.4\text{-}165)$$

由累积损伤理论得应力水平（S_a，S_m）作用下循环次数 dn，对应的微损伤（采用 Miner 准则）为

$$dD = \frac{dn}{N_p} = \left[\frac{f(S_a, S_m)}{\alpha_p} \left(\frac{\sigma_{bp}}{\sigma_{bp} - S_m} S_a - S_{0p} \right)^{\beta_p} dS_a dS_m \right] N_p^*$$

因为 $\int dD = 1$，所以给定存活率 p 的疲劳寿命为：

$$N_p^* = 1 \Big/ \int_{(S_m)_{min}}^{(S_m)_{max}} \int_{(S_a)_{min}}^{(S_a)_{max}} \frac{f(S_a, S_m)}{\alpha_p} \left(\frac{\sigma_{bp}}{\sigma_{bp} - S_m} S_a - S_{0p} \right)^{\beta_p} dS_a dS_m \quad (5.4\text{-}166)$$

如果材料疲劳性能采用 $\gamma - P - S_a - S_a - N$ 曲面，即以置信度 γ 保证存活率不低于 p 的疲劳寿命为

$$N_{\gamma p}^* \left(\frac{\sigma_{b\gamma p}}{\sigma_{brp} - S_m} S_a - S_{0p} \right)^{\beta_{rp}} = \alpha_{\gamma p}$$

我们就可预测以置信度 γ 保证存活率不低于 p 的疲劳寿命为

$$N_p^* = 1 \Big/ \int_{(S_m)_{min}}^{(S_m)_{max}} \int_{(S_a)_{min}}^{(S_a)_{max}} \frac{f(S_a, S_m)}{\alpha_{\gamma p}} dS_a dS_m \quad (5.4\text{-}167)$$

如果采用材料疲劳性能 $P - S - N$ 曲线为

$$N (S - S_0)^{\beta} = \alpha$$

式中，S_0、β、α 为材料常数，并且应力幅度分布的概率密度函数为 $f(\sigma)$ 时，预测寿命为

$$N_p^* = 1 \Big/ \int_{S_0}^{S_{max}} \frac{f(\sigma)}{\alpha} (S - S_0)^{\beta} dS \quad (5.4\text{-}168)$$

3.2.4 累积损伤理论的疲劳寿命预测修正

若不考虑累积损伤对材料疲劳性能的影响，即对循环应力、应变曲线（或曲面）的影响。在 $P - S_a - S_a - N$ 曲面中，令 N 为常数 N_p，即存活率为 p 的寿命给定，得等寿命疲劳极限图，如图 5.4-39 所示。不发生断裂失效的可靠度 R 可定义为

$$R = \iint_{D_p} f(S_a, S_m) dS_a dS_m \quad (5.4\text{-}169)$$

失效概率为

$$R = 1 - \iint_{D_p} f(S_a, S_m) dS_a dS_m \quad (5.4\text{-}170)$$

式中，D_p 为图 5.4-39 中非阴影区域。也就是应力水平（S_a，S_m）落入等寿命疲劳极限图内的概率。如果采用解析法积分上式，有时很繁琐，尤其是（S_a，S_m）相关情形，一般采用 Monte-Carlo 方法比较方便，方法是根据 f（S_a，S_m）联合概率密度函数，产生随机向量（S_a，S_m），总模拟次数为 N，如果随机向量（S_a，S_m）落入等寿命极限图内的次数为 N'，则不发生疲劳断裂失效的可靠度为

$$R_p = \frac{N'}{N}$$

如果采用 $\gamma - P - S_a - S_a - N$ 曲面的等寿命疲劳极限图，得给定疲劳寿命 N_{rp} 条件下，带置信度 γ 和存活率 p 的不发生疲劳断裂失效的可靠度为

$$R_{\gamma p} = \iint_{D_{\gamma p}} f(S_a, S_m) dS_a dS_m$$

对应的 Monte-Carlo 方法计算可靠度公式为

$$R_{\gamma p} = \frac{N'}{N}$$

累积损伤与载荷顺序、应力幅度与应力均值相关程度等多种因素有关，所以简单用 Miner 公式表达累积损伤是比较困难的，需要根据以往实验数据修正。

为了避开累积损伤等于 1 代表疲劳断裂失效这一不确定假设，采用以往经验资料进行修正。设已知相似谱作用下计算寿命为 N'_{calc}，相似谱作用下实测寿命为 N'_{exp}，以及计算谱的计算寿命 N_{calc}，预测计算谱下的寿命为 N_{exp}，它们之间的关系为

$$N_{exp} = \frac{N'_{exp}}{N'_{calc}} N_{calc} \quad (5.4\text{-}171)$$

给定存活率 p 的预测寿命为

$$(N_{exp})_p = \frac{(N'_{exp})_p}{(N'_{calc})_p} (N_{calc})_p \quad (5.4\text{-}172)$$

式中，给定存活率 p 的相似谱的实测寿命（N'_{exp}）$_p$，可以根据相似谱作用条件下多组实验数据进行估计。

同理，给定置信度 γ 和存活率 p 的预测寿命为

$$(N_{exp})_{\gamma p} = \frac{(N'_{exp})_{\gamma p}}{(N'_{calc})_{\gamma p}} (N_{calc})_{\gamma p} \quad (5.4\text{-}173)$$

3.3 确定性疲劳评定

本篇第 3 章所述，通常根据工作条件的不同，可以将疲劳寿命的估算方法分为三类：高周疲劳寿命的估算方法（包括 BasguinMiner 累计疲劳寿命损伤理论和用疲劳寿命与应力幅之间的关系来估算）、低周疲劳寿命估算（用 Manson-Coffin 公式）及断裂力学疲劳寿命的估算（Paris 公式）。在疲劳评定中主要是考虑疲劳断裂问题，因此主要依据 Paris 理论来进行评定。根据缺陷形式的不同，这里给出了平面缺陷和体积缺陷两种疲劳安全评定方法，下面分别加以介绍。

3.3.1 平面缺陷的确定性疲劳评定

（1）评定方法

平面缺陷的疲劳评定，首先依据疲劳裂纹扩展速率 da/dN 与裂纹尖端应力强度因子变化幅度 ΔK 的关系式 $\frac{da}{dN}C = (\Delta K)^m$，确定在规定的循环周期内疲劳裂纹的扩展量和最终尺寸，然后根据所给出的判别条件和方法，来判断该平面缺陷是否会发生泄漏和疲劳断裂。除 2.4.1 所规定的平面缺陷外，难以明确划分为非平面缺陷的埋藏缺陷也应按平面缺陷进行评定。

（2）评定程序

平面缺陷疲劳评定可按下列步骤进行安全评定：

1）缺陷的表征；

2）应力变化范围的确定；

3）材料性能数据的确定；

4）疲劳裂纹的 ΔK 计算；

5）免于确定性疲劳评定的判别；

6）疲劳裂纹扩展量的计算；

7）容许裂纹尺寸的计算和确定性安全性评价。

评定程序如图 5.4-41 所示。

（3）疲劳评定所需基本数据的确定

1）缺陷的表征　按 2.4.1 节中的规定对平面缺陷进行规则化处理，并确定疲劳评定初始裂纹的尺寸。

2）应力变化范围及循环次数的确定　根据疲劳评定程序，需要确定缺陷的应力变化范围及相应的循环次数。首先要根据外加载荷或温度的变化历程，分别确定被评定缺陷所在截面上垂直于裂纹平面的一次应力二次应力的应力变化范围的分布曲线及其循环次数（平行于裂纹平面的应力变化不予考虑）。然后根据应力变化范围的线性化规则，分别计算出评定所需的薄膜应力变化范围 $\Delta\sigma_s$ 和弯曲应力变化范围 $\Delta\sigma_B$。所谓应力变化范围的线性化规则是定义一条线性的、在缺陷处壁厚范围内各点应力变化范围值均不低于实际分布曲线上的应力变化范围值的应力变化范围分布直线（图 5.4-

42)，并按该线性化的应力变化范围分布直线确定疲劳评定中所需的薄膜应力变化范围值 $\Delta\sigma_m$ 及弯曲应力变化范围值 $\Delta\sigma_B$。

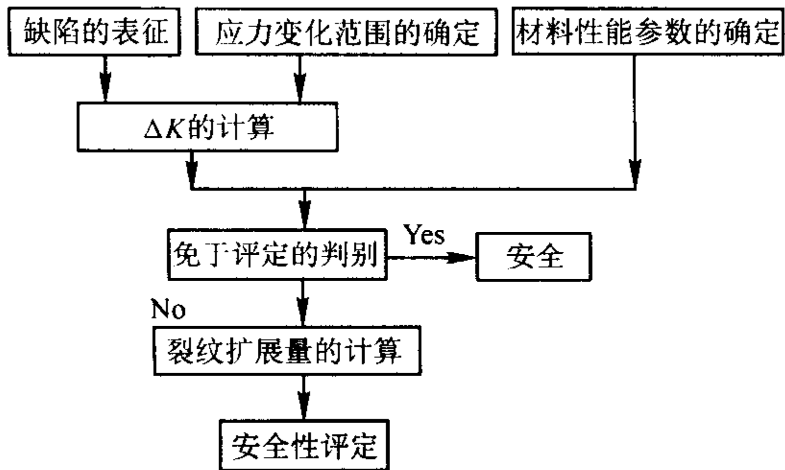


图 5.4-41 平面缺陷的疲劳评定程序示意图

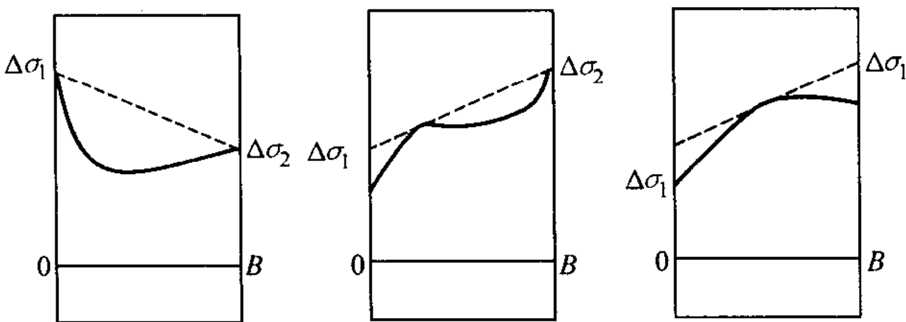


图 5.4-42 疲劳评定中应力变化范围分布的线性化规则图例

$$\begin{cases}\Delta\sigma_m = \dfrac{\Delta\sigma_1 + \Delta\sigma_2}{2} \\ \Delta\sigma_B = \dfrac{\Delta\sigma_1 - \Delta\sigma_2}{2}\end{cases} \quad (5.4\text{-}174)$$

式中，$\Delta\sigma_1$ 和 $\Delta\sigma_2$ 分别为线性化后内、外壁处的应力变化范围值。

其中以一次应力的应力变化范围分布曲线所获得的 $\Delta\sigma_m$ 及 $\Delta\sigma_B$ 为 ΔP_s 及 ΔP_b，以二次应力的应力变化范围分布曲线所获得的 $\Delta\sigma_m$ 及 $\Delta\sigma_B$ 为 ΔQ_s 及 ΔQ_b。若在预期寿命内存在 d 种不同的应力变化范围，则应按评定周期内的载荷作用历程，计算出 $i=1, 2, \cdots, d$ 种不同应力变化范围作用时的 $(\Delta\sigma_m)_i$ 和 $(\Delta\sigma_B)_i$，同时确定其在评定期间内相应的预期循环次数 n_i。

在确定载荷变化范围时需要注意以下几个问题：一是计算中应包括由于操作压力、操作温度和其他外载荷的波动所产生的应力变化范围，并考虑它们的组合效果（但焊接残余应力不予考虑）；二是当容器受双向应力疲劳作用时，其疲劳评定按单向应力疲劳评定方法进行。

3）材料性能数据的确定　在疲劳评定时，需要确定有关的材料性能，其中包括疲劳裂纹扩展速率与 ΔK 的关系式中的系数 C 与指数 m 的值，疲劳裂纹扩展的应力强度因子变化范围门槛值 ΔK_{th} 的值以及材料的 $P-S-N$ 曲线等。其中材料的 $P-S-N$ 曲线在前面部分已经介绍过，下面就前两个方面的取值问题进行讨论。

① 疲劳裂纹扩展速率与 ΔK 的关系式中的系数 C 与指数 m 的取值　在做评定时应尽量采用实测值，及尽可能从实际服役的设备（如容器）上取样，然后按 GB/T 6398—1986《金属材料疲劳裂纹扩展速率试验方法》的规定进行实验。根据实验数据，用最小二乘法回归得到 C 和 m。但在评定中不建议直接采用最小二乘法回归得到的 C 值进行计算，还应乘以一个不小于4.0的系数后才能作为评定所取用的 C 值。按这一规定的统计分析结果表明其存活率在99.99%以上。下面给出几对典型工况和材料的 C 和 m 的推荐值。

对16 MnR钢，在100℃以下的空气环境中，并且 ΔK 在 $300\sim1\,500\ \mathrm{N/mm^{3/2}}$ 范围内时，可取：$m=3.35$，$C=6.44\times10^{-14}$。

对 $\sigma_{0.2}<600$ MPa 的铁素体钢，在不超过100℃的空气环境中，也可取：$m=3.0$，$C=3\times10^{-13}$。

对伴有解理或微孔聚合等具有更高扩展速率的疲劳裂纹扩展机制时，可取：$m=3.0$，$C=6\times10^{-13}$。

② 疲劳裂纹扩展的应力强度因子变化范围门槛值 ΔK_{th} 的取值　当幸存概率为97.5%时，碳钢和碳锰钢在空气中的疲劳裂纹扩展应力强度因子变化范围门槛值 ΔK_{th} 可以用以下方法估算。

$$\text{对于母材：}\begin{cases}\Delta K_{th} = 170 - 214R_\sigma & \text{当 } 0\leqslant R_\sigma\leqslant 0.5 \text{ 时} \\ \Delta K_{th} = 63 & \text{当 } R_\sigma > 0.5 \text{ 时}\end{cases} \quad (5.4\text{-}175)$$

$$\text{对于焊接接头：}\begin{cases}\Delta K_{th} = 214\Delta\sigma/\sigma_s - 44 & \text{当 } \Delta\sigma > \sigma_s/2 \text{ 时} \\ \Delta K_{th} = 63 & \text{当 } \Delta\sigma \leqslant \sigma_s/2 \text{ 时}\end{cases} \quad (5.4\text{-}176)$$

4）应力强度因子变化范围 ΔK 的计算　根据前面得到的表征裂纹尺寸 a_0 和 c_0 和计算的应力变化范围 $\Delta\sigma_m$ 和 $\Delta\sigma_B$，可以按相关文献中的规定计算出各自的应力强度因子变化范围 ΔK_a 和 ΔK_C。

5）免于疲劳评定的判别　与断裂体积型缺陷的评定类似，在疲劳评定过程中也有免于评定的判别，若通过此判别，不需要进一步的评定，即可认为是安全的，若不能通过此判别，则需要做进一步的疲劳评定。免于评定的判别有以下两点，若能满足其中之一即可通过此判别。

① $\Delta K_1 < \Delta K_{th}$，即计算得到的材料应力强度因子小于其疲劳扩展的门槛值，此时认为疲劳裂纹不扩展，因此可以通过免于评定的判别，认为是安全的或可接受的；

② 考虑到大多数设备（如压力容器）承受压力循环次数不多，只要在使用寿命期内，其疲劳扩展量小于无损检测被检出的最小尺寸，仍可免于疲劳评定。为此，这里给出了不同载荷循环的 ΔK_a、ΔK_c 值和对应的预期循环次数（表5.4-11)，并规定如果其结果均小于表5.4-11相应各 ΔK 所对应的容许循环次数，则缺陷可免于疲劳评定，认为是安全的或可接受的。这一规定简化了判别免于疲劳评定时的计算。

表 5.4-11 免于疲劳评定界限

ΔK 值		容许承受循环次数
表面裂纹	埋藏裂纹	
690 ~ 551	460 ~ 368	1×10^3
550 ~ 407	367 ~ 271	2×10^3
406 ~ 321	270 ~ 214	5×10^3
320 ~ 255	213 ~ 170	1×10^4
254 ~ 188	169 ~ 125	2×10^4
187 ~ 149	124 ~ 101	5×10^4
148 ~ ΔK_{th}	100 ~ ΔK_{th}	1×10^5
ΔK_{th}	$<\Delta K_{th}$	不限

6) 疲劳裂纹扩展量和裂纹最终尺寸 a_f 和 c_f 的计算

若不能通过免于评定的判别，则需要进行疲劳裂纹扩展情况的分析。本节给出了疲劳裂纹扩展量和最终裂纹尺寸 a_f、c_f 的循环计算和分段简化计算的方法与步骤。计算中采用了按应力变化范围历程逐个循环计算法。研究表明，基于 Paris 公式及 Newman-Rnjas 应力强度因子解的逐个循环计算法，不仅其理论较为严密，其精度比 WES2805-83 的川原正言法及 BSIPD6493-91 的多级 $S-N$ 曲线法要高。除逐个循环计算法外，还建议采取分段计算法，这可以大大减少计算工作量，甚至可以用手工完成。下面分别介绍这几种计算方法和步骤。

① 逐个循环计算方法与步骤　此方法是按应力变化范围历程逐个循环来计算每一循环后裂纹的增长，然后不断叠加，直至计算出所有循环后裂纹的最终尺寸。其具体步骤如下。

a) 按裂纹初始尺寸 a_0、c_0 和第 1 个循环的 $(\Delta\sigma_m)_1$、$(\Delta\sigma_B)_1$ 分别计算 $(\Delta K_a)_0$、$(\Delta K_c)_0$，并计算第 1 个循环后的裂纹尺寸：

$$\begin{cases} a_1 = a_0 + A\ (\Delta K_a)_0^m \\ c_1 = c_0 + A\ (0.9\Delta K_c)_0^m \end{cases}$$

b) 按第 1 个循环后的裂纹尺寸 a_1、c_1 和第 2 个循环的 $(\Delta\sigma_m)_2$、$(\Delta\sigma_B)_2$ 分别计算 $(\Delta K_a)_1$、$(\Delta K_c)_1$，并计算第 2 个循环后的裂纹尺寸：

$$\begin{cases} a_2 = a_1 + A\ (\Delta K_a)_1^m \\ c_2 = c_1 + A\ (0.9\Delta K_c)_1^m \end{cases}$$

c) 重复以上 a)，b) 步骤，按应力变化范围的历程，用式 (5.4-177) 逐次计算经历第 i 个应力变化循环后的裂纹尺寸。

$$\begin{cases} a_i = a_{i-1} + A\ (\Delta K_a)_{i-1}^m \\ c_i = c_{i-1} + A\ (0.9\Delta K_c)_{i-1}^m \end{cases} \tag{5.4-177}$$

d) 重复以上步骤，直到评定期间预期的最后一个应力变化循环为止，即得到疲劳扩展后的最终裂纹尺寸 a_f 和 c_f。

② 分段简化计算的方法与步骤　此方法是在多种应力变化范围内，各种变化范围大体均匀分布在疲劳评定周期内的前提条件下，将每一种应力变化范围的循环数均匀分段，然后再进行叠加计算的简化方法。具体计算步骤如下。

a) 如果有 d 种应力变化范围，各种应力变化的循环次数分别 n_i $(i=1,\ 2,\ \cdots,\ d)$，则可将每一种应力变化范围的循环数 n_i $(i=1,\ 2,\ \cdots,\ d)$ 均分为 t 段 $(j=1,\ 2,\ \cdots,\ u)$。第 i 种应力变化范围的循环数 n_i 均分为 t 段后，每段的循环次数为 n_i/t。对于恒幅循环，每段的循环次数为 N/t，一般取 $t=5$。

b) 从 $j=1$ 起至 $j=t$ 为止，应用式 (5.4-178) 及式 (5.4-179) 可逐段计算出在各种应力变化范围作用下计算段裂纹尺寸的近似平均值 a_j 及 c_j，直至计算出疲劳扩展的最终尺寸 a_f 和 c_f。此计算方法属于偏安全的计算方法。

$$\begin{cases} a_j = a_{j-1} + A\sum_{i=1}^{d} n_i[(\Delta K_a)_i]_{j-1}^m/t \\ a_f = a_t + 0.5A\sum_{i=1}^{d} n_i[(\Delta K_a)_i]_t^m/t \end{cases} \quad j = 1,2,\cdots,u \tag{5.4-178}$$

$$\begin{cases} c_j = c_{j-1} + A\sum_{i=1}^{d} n_i[(\Delta K_c)_i]_{j-1}^m/t \\ c_f = c_t + 0.5A\sum_{i=1}^{d} n_i[(\Delta K_c)_i]_t^m/t \end{cases} \quad j = 1,2,\cdots,u \tag{5.4-179}$$

式中 $[(\Delta K_a)_i]_{j-1}$和 $[(\Delta K_c)_i]_{j-1}$分别为裂纹尺寸 $a=a_{j-1}$，$c=c_{j-1}$时第 i 种 $(\Delta\sigma_m)_i$ 和 $(\Delta\sigma_B)_i$ 作用下在 a 和 c 方向裂纹尖端的 ΔK 值。

如为恒幅循环，即 $d=1$，则式 (5.4-178) 和式 (5.4-179) 中的 $\sum_{i=1}^{d} n_i$ 简化为 N，$(\Delta K_a)_i$ 简化为 (ΔK_a)，$(\Delta K_c)_i$ 简化为 (ΔK_c)。

③ 可忽略 c 方向扩展时的简化计算方法　在某些工况下，c 方向上裂纹的扩展很小，可以忽略不计。因此在工程上当满足式 (5.4-180) 的条件时，则可以忽略 c 方向的扩展。

$$\Delta K_c/\Delta K_a < 0.5 \tag{5.4-180}$$

忽略 c 方向扩展时，采用本节 1) 和 2) 的计算方法，只计算 a_f 值，但不计算有关 c 方向的所有参数即可。

(4) 确定性疲劳安全性评价

在很多情况下（如盛装危险品或化学、化工产品的压力容器、压力管道等）安全评定不仅包括预防和防止断裂的分析评定，还包括预防和防止泄漏的评定。因此在疲劳安全评定的过程，与前面有关断裂安全评定有所不同，它考虑了疲劳泄漏的评定问题。

1) 疲劳泄漏评定　一般可以根据不同的裂纹类型来进行疲劳泄漏评定，如

对表面裂纹，若 $a_f > 0.7B$，则认为不会发生泄漏。

对埋藏裂纹，若 $\left(\dfrac{p_1 + a_0 - a_f}{a_f}\right) \geqslant 0.8$ 且 $\left(\dfrac{p_1 + a_0 + a_f}{B}\right) \leqslant 0.7$，则认为不会发生泄漏。

2) 疲劳断裂评定　根据前面最终的裂纹尺寸 a_f 及 c_f 值和缺陷所在部位承受的最大载荷应力值，按本章 2.4 节的断裂及塑性破坏评定中平面缺陷简化评定或常规评定的方法来进行断裂和塑性破坏评定，如果评定的结果是安全或可以接受的，则认为不会发生因疲劳断裂和塑性破坏导致的失效。

若疲劳评定结果能同时满足 1) 和 2) 的条件，则认为该缺陷是安全的或可以接受的；否则，是不能保证安全或不可接受的。

3.3.2　体积型（焊接）缺陷的确定性疲劳评定

与前面断裂安全评定相对应，这里也给出了有关体积型缺陷的疲劳安全评定方法和步骤。

满足下述条件的含体积型缺陷（主要是在用压力容器焊接接头的）可以进行基于材料 $P-S-N$ 曲线的评定疲劳评定。具体的适用范围为：

1) 构件（如容器）壁厚等于或大于 10 mm。

2) 对于碳钢、低合金钢等材料（或容器），操作温度低于 375℃。

3) 对于奥氏体不锈钢制材料（或容器），操作温度低于 430℃的。

(1) 体积型缺陷的疲劳安全评定程序

1) 缺陷表征；

2) 应力变化评定的判别；

3) 免于疲劳评定的判别；

4) 使用工况所要求承受的疲劳强度参量 $(S^3N)_x$ 值的确定；

5) 容许承受的疲劳强度参量 $(S^3N)_y$ 值的确定；

6) 疲劳安全性评价。

体积型缺陷的疲劳安全评定程序如图 5.4-43 所示。

(2) 疲劳评定所需基本数据的确定

1) 缺陷的表征　按本章 2.5 节中说明的有关体积型缺陷表征方法对缺陷进行表征。

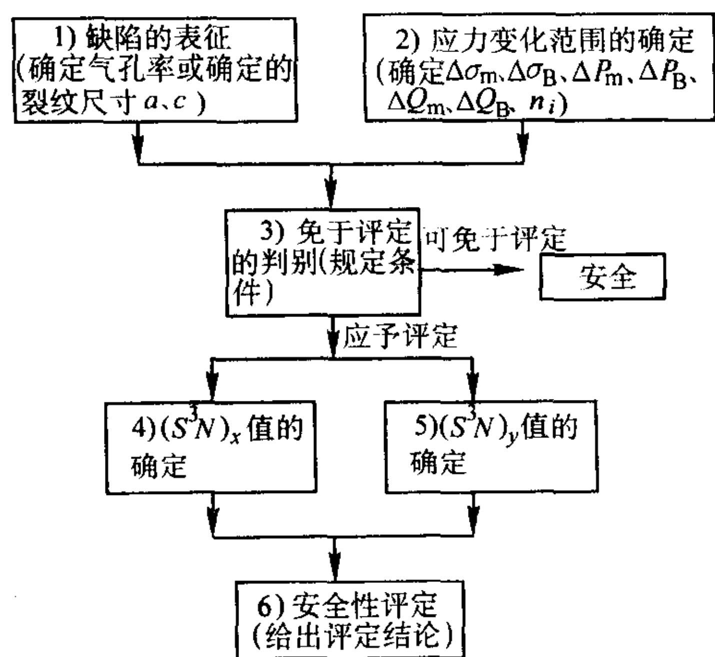


图 5.4-43　体积型缺陷的疲劳安全评定程序示意图

2) 应力变化范围及循环次数的确定　按本章 3.3.1 节中的规定确定应力的变化范围。但应注意，体积缺陷承受的循环次数 N 及 n_j 应为从该设备投入运行的时间开始计算。

(3) 免于疲劳评定的判别

与断裂体积型缺陷的评定类似，在疲劳评定过程中也有免于评定的判别，若通过此判别，不需要进一步的评定，即可认为是安全的，若不能通过此判别，则需要做进一步的疲劳评定。免于评定的判别有以下两点，符合以下条件之一者，可免于进行疲劳评定，并认为该缺陷是可以接受的。

1) 缺陷所在截面的工作应力变化范围低于 23 MPa;

2) 仅承受与焊缝方向一致的疲劳载荷的咬边缺陷。

(4) $(S^3N)_x$ 值的计算

由体积型缺陷的疲劳安全评定的步骤可以看出，与前面所有评定步骤有所不同，它需要进行 $(S^3N)_x$ 值和 $(S^3N)_y$ 值的计算和确定。针对不同的疲劳情况其 $(S^3N)_x$ 值具体计算方法如下。

1) 对于恒幅疲劳，可根据缺陷所在截面的应力变化范围 $\Delta\sigma$ 和在整个寿命期内的总循环次数 N，按式 (5.4-181) 计算 $(S^3N)_x$：

$$(S^3N)_x = (\Delta\sigma)^3 N \qquad (5.4\text{-}181)$$

2) 对于非恒幅疲劳，如有 d 种应力变化范围 $\Delta\sigma_i$ ($i=1, 2, \cdots, d$)，它们所承受的循环次数分别为 n_i ($i=1, 2, \cdots, d$)，则按式 (5.4-182) 计算 $(S^3N)_x$：

$$(S^3N)_x = \sum_{i=1}^{d}[(\Delta\sigma_i)^3 n_i] \qquad (5.4\text{-}182)$$

式中，$\Delta\sigma_i = (\Delta\sigma_m)_i + (\Delta\sigma_i)_i$。

3) 计算非恒幅疲劳的 $(S^3N)_x$ 时，所有小于表 5.4-12 规定的最小应力变化范围 $(\Delta\sigma)_{min}$，可以忽略不计。

表 5.4-12　计算非恒疲劳的 $(S^3N)_x$ 时可忽略的最小应力变化范围 $(\Delta\sigma)_{min}$

$(S^3N)_x$ 值	$(\Delta\sigma)_{min}$/MPa
1.52×10^{12}	42
1.04×10^{12}	37
6.33×10^{11}	32
4.31×10^{11}	28
2.50×10^{11}	23

4) 材料（或容器）受双向应力疲劳作用时，其疲劳评定按单向应力疲劳评定方法进行。

(5) $(S^3N)_y$ 值的确定

1) 对于气孔缺陷，根据气孔率可由表 5.4-13 来确定所容许承受的疲劳强度参数 $(S^3N)_y$ 值;

表 5.4-13　含气孔焊接接头容许承受的疲劳强度参量 $(S^3N)_y$

气孔在射线底片上所占的面积	$(S^3N)_y$
3%	4.980×10^6E
5%	1.196×10^6E

注：表 5.4-13～表 5.4-17 中的 E 为材料评定温度下该金属材料的弹性模量。

2) 对于夹渣缺陷，根据夹渣长度以及焊缝是否进行焊后消氢热处理的情况，按表 5.4-14 或表 5.4-15 确定所容许承受的疲劳强度参见 $(S^3N)_y$ 值;

3) 对于壁厚 $B=10\sim25$ mm、深度 <1 mm 材料（或容器）的咬边缺陷，根据咬边深度和壁厚 B 的比值，按表 5.4-16 或表 5.4-17 确定所容许承受的疲劳强度参量 $(S^3N)_y$ 的值。

表 5.4-14　含夹渣的焊态焊接接头容许承受的疲劳强度参量 $(S^3N)_y$

最大夹渣长度/mm	$(S^3N)_y$
2.5	7.270×10^6E
4.0	4.980×10^6E
10	3.029×10^6E
35	2.062×10^6E
>35	1.196×10^6E

注：E 的含量见表 5.4-13 注。

表 5.4-15　含夹渣的经消氢热处理的焊接接头容许承受的疲劳强度参量 $(S^3N)_y$

最大夹渣长度/mm	$(S^3N)_y$
19	7.270×10^6E
58	4.980×10^6E
>58	1.196×10^6E

注：E 的含量见表 5.4-13 注。

表 5.4-16　含咬边的对接焊接接头容许承受的疲劳强度参量 $(S^3N)_y$

最大咬边深度/壁厚	$(S^3N)_y$
0.025	7.270×10^6E
0.050	4.980×10^6E
0.075	3.029×10^6E
0.100	1.196×10^6E

注：E 的含量见表 5.4-13 注。

表 5.4-17　含咬边的角接焊接接头容许承受的疲劳强度参量 $(S^3N)_y$

最大咬边深度/壁厚	$(S^3N)_y$
0.050	3.029×10^6E
0.075	2.062×10^6E
0.100	1.196×10^6E

注：E 的含量见表 5.4-13 注。

(6) 安全性评价

根据前面对 $(S^3N)_x$ 值和 $(S^3N)_y$ 值的计算，如果体积缺陷经评定满足式 (5.4-183)，则该体积型缺陷是容许的或可以接受的；否则，是不能容许或不可接受的。

$$(S^3N)_y \geqslant (S^3N)_x \qquad (5.4\text{-}183)$$

(7) 英国标准 PD－6493 的质量带理论

用于含体积缺陷焊接结构的疲劳评定方法还有英国标准 PD－6493 的质量带理论。我国在劳动部门主持下，“八五”

期间对质量带理论是否适用于我国钢种，进行了大量的工作，以便消化、吸收国外标准，与国外先进规程接轨，赶上国际先进水平。英国标准 PD－6493 中对焊缝非平面缺陷，根据不同初始质量（即含不同密度的气孔和不同长短的夹渣），将焊接结构质量分成十个等级。这种方法建立在大量试验基础上，使用方便，较适合工程应用。因此这里对质量带理论进行一下简单的介绍。

英国标准 BSI　PD－6493 认为含体积型缺陷（如气孔和夹渣）焊接结构的疲劳寿命 N_f 与承受的交变应力幅之间存在如下关系：

$$S^3 N_f = C \tag{5.4-184}$$

式中，S 为结构承应力幅；N_f 为结构的疲劳寿命；C 与结构缺陷大小有关的常数。

根据 BSIPD-6493 的第 20 款表格计算得各质量等级的常数 C 列于表 5.4-18 中。

表 5.4-18　$S-N$ 曲线的质量等级 C 值

质量等级	C 值
Q1	1.48×10^{12}
Q2	1.01×10^{12}
Q3	6.55×10^{11}
Q4	4.39×10^{11}
Q5	2.43×10^{11}
Q6	1.60×10^{11}
Q7	9.83×10^{10}
Q8	6.75×10^{10}
Q9	3.46×10^{10}
Q10	2.66×10^{10}

具体评定方法如下：

1）用无损探伤法检测体积缺陷（如气孔或夹渣）的性质与几何尺寸，得到对结构的实际质量等级；

2）由该结构所要承受载荷大小和所要求寿命，得到结构要求的质量等级；

3）如果实际质量等级小于或等于所要求质量等级，则结构是安全的。

恒载作用时，如果已知结构实际质量等级 C 值和所承受恒应力幅 S，可以预测疲劳寿命 N_f，如图 5.4-44 所示。

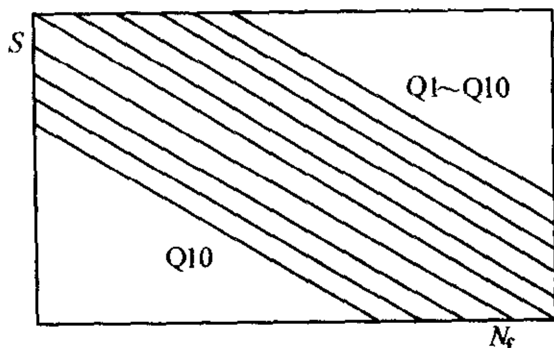


图 5.4-44　质量等级图

目前，从对 16 MnR 钢所作有限的实验结果来看，PD－6493 工程评定方法对我国 16 MnR 钢来说偏于保守。

4　其他失效模式的确定性安全评定技术和方法

除断裂和疲劳评定外，还有一些有关其他失效模式的安全评定技术和方法，例如塑性失稳、蠕变、腐蚀等。

4.1　剩余截面超载屈服

在实际工况中，有许多小结构截面需要进行剩余截面超载屈服的失效模式进行临界评定。在此种临界评定中需要进行详细的应力分析，包括考虑载荷重新分布的影响。否则不能将该种评定用于设备（压力容器）的缺陷或其他承受复杂载荷截面上的缺陷。因此首先要对评定部位进行详细的分析。这里将介绍一些有关净截面应力 σ_n 的求取方法。

4.1.1　平板净截面应力公式

（1）对于穿透裂纹

$$\sigma_n = \frac{P_b + \sqrt{P_b^2 + 9P_m^2}}{3[1-(2a/W)]} \tag{5.4-185}$$

（2）对于表面裂纹

1）在正常的弯曲约束下

$$\sigma_n = \frac{P_b + \sqrt{P_b^2 + 9P_m^2\ (1-\alpha)^2}}{3\ (1-\alpha)^2} \tag{5.4-186}$$

2）在可忽略的弯曲约束（如销钉连接）

$$\sigma_n = \frac{P_b + 3P_m\alpha + \sqrt{(P_b + 3P_m\alpha)^2 + 9P_m^2\ (1-\alpha)^2}}{3\ (1-\alpha)^2} \tag{5.4-187}$$

式中　$\alpha = (a/B)\ /\ [1+B/c]$　　对 $W \geqslant 2\ (c/B)$

　　　$\alpha = (2a/B)\ (c/W]$　　对 $W \leqslant 2\ (c/B)$

（3）对于埋藏裂纹

$$\sigma_n = \frac{P_b + 3P_m\alpha + \sqrt{(P_b + 3P_m\alpha)^2 + 9P_m^2[(1-\alpha)^2 + 4p\alpha/B]}}{3[(1-\alpha)^2 + 4p\alpha/B]} \tag{5.4-188}$$

式中　$\alpha = (2a/B)/[1+B/c]$　对 $W \geqslant 2(c/B)$

　　　$\alpha = (4a/B)/c/W]$　　对 $W \leqslant 2(c/B)$

4.1.2　内压作用下的弯曲壳体应力公式

对于球壳的所有缺陷（包括不同类型的缺陷）和管道中的轴向缺陷可用以下几式来计算应力。

（1）对于穿透裂纹

$$\sigma_n = 1.2 M_T P_m \tag{5.4-189}$$

式中，$M_T = [1+3.2\ (a^2/DB)]^{0.5}$；$P_m$ 为周向膜应力。

（2）对于表面裂纹

$$\sigma_n = 1.2 M_n P_m \tag{5.4-190}$$

式中，$M_n = \dfrac{1+\ (a/BM_T)}{1-\ (a/B)}$；$M_T = [1+3.2\ (c^2/DB)]^{0.5}$。

若计算所得 σ_n 的最大值不超过材料的屈服强度，即可对受拉伸载荷的情况进行详细的评定。对剪切载荷，应该计算净截面上的最大剪切应力，并且此值不能超过 0.48 倍的材料屈服强度。

当满足上述情况时，按照本章第 2 节断裂和塑性失效评定技术和方法的相关内容进行下一步评定即可完成整个评定过程。

4.2　破裂前的泄漏分析（LBB）

有一些情况，虽然缺陷可能贯穿壁厚，但不能导致灾难性的断裂失效，这在本章第 2 节中，是有可能通过安全评定的。但是对于承压、储存液体或真空等设备，凡是已形成内外通道的缺陷，或在使用中有可能形成这种通道和容器内壁缺陷都是不允许的。因此需要进行检测和计算分析，此类检测分析被称为破裂前泄漏分析（LBB）。

破裂前泄漏分析一般按下列步骤进行：

1）首先要采用本章第2节的相关内容证明构件的最大初始缺陷尺寸不会导致断裂。

2）仍采用本章第二节的相关内容确定一个完全贯穿裂纹的最大（临界）裂纹长度，当裂纹长度小于此长度时，所有的载荷工况都不会引起灾难性的破裂。

3）计算1）、2）步骤所确定的裂纹面积相应的泄漏面积。

在计算泄漏流速时需要知道一个可能的贯穿壁厚缺陷的裂纹张开面积（COA）。COA与裂纹的几何尺寸（有效长度、形状、方位等）、构件的几何尺寸、材料性质以及有关载荷。COA的计算方法可以分为三类：线弹性模型、带有小范围塑性修正的线弹性模型、弹塑性模型。它们的精度随几何尺寸、裂纹尺寸以及载荷的类型和幅度的变化而变化。下面介绍几种情况的COA的计算方法。

① 对于贯穿裂纹长为 $2c$ 的平板，裂纹长度小于板宽和板长。

$$COA_p = \frac{2\pi\sigma c^2}{E'} \tag{5.4-191}$$

式中，COA_p 为平板的裂纹张开面积；c 为贯穿裂纹的半长；E' 为弹性模量，平面应力时取 E，平面应变取 $E/(1-\nu^2)$；σ 为拉应力；ν 为泊松比。

② 对于有贯穿壁厚裂纹的圆柱壳或球壳

$$COA_s = \alpha\ (COA_p) \tag{5.4-192}$$

式中，COA_s 为壳的张开面积；α 为曲率因子。

③ 对于简单几何形状，曲率修正因子是壳参数 λ 的函数。

Ⅰ）圆柱壳中的纵向裂纹（适用于 $\lambda \leqslant 8$）

$$\alpha = 1 + 0.1\lambda + 0.16\lambda^2;$$

Ⅱ）圆柱壳中的周向裂纹（适用于 $\lambda \leqslant 5$）

$$\alpha = (1 + 0.17\lambda^2)^{0.5};$$

Ⅲ）球壳中的周向裂纹（适用于 $\lambda \leqslant 5$）

$$\alpha = 1 + 0.02\lambda + 0.22\lambda^2。$$

4）计算确定与3）步骤裂纹面积所对应的泄漏速率，并证明相应的泄漏可以通过所选择的泄漏检测系统进行检测。

流体流速或泄漏速率的计算是一个与裂纹的几何形状、流经路径的长度、流体摩擦的影响以及热力学有关的问题。一些文献提供了等温或多相流体或气体泄漏速率的近似计算方法，而某些文献则提供了两相气/水混合流体泄漏速率的计算方法。

对于用无损探伤法难以显示的泄漏通道，可以通过水压试验来进行检测。然后可以采取各方认可修补措施进行修补或改造。对于在操作期间产生的复杂泄漏通道，必需采取特殊的修补办法，如采用堆焊方法。

另外还要采用一些检测手段和方法。典型的检测方法，包括了下列的工具或方法：腐蚀传感器、氢传感器、UT测量和扫描、射线照片检查、对流体的化学分析、红外成像法、热电偶。

4.3 环境的影响

每当有腐蚀环境存在，在工程临界的评定中都必须考虑其影响。因为环境的影响往往是各有其特征的，所以只可能给出一般性的指导和说明。环境的影响可能对材料或操作条件的微小变化非常敏感，因此，任何评定或实际测量必须考虑在给定的材料标准范围内的热处理、化学成分等的差异。同样地，由于使用条件可能会改变材料对环境的敏感性，如回火脆性或材料的退化，因此，必须考虑其相应的使用条件，对材料进行评定或测量。

4.3.1 腐蚀与侵蚀

为了防止腐蚀与侵蚀等失效模式的出现首先要按具体情况（介质、环境等）选取合适的材料，同时还应注意下面几个方面的问题。

对于在两个表面上都没有露出的缺陷（如裂纹、未熔合、未焊透、气孔和凹坑等）应当去掉腐蚀/侵蚀裕量，然后按4.1节中的说明和建议进行评定。而对于腐蚀或侵蚀环境中在表面上的任何一种缝隙，如裂纹等时，应该估算其最大的腐蚀速率和扩展到认为不可接受的尺寸所需要的时间。

评定多点腐蚀（及其相互影响）的保守做法是取这组凹坑的包络面，并当作具有同样深度和表面长度的平面缺陷处理，视具体情况按2.4节的建议进行评定。

除此以外，腐蚀评定中还应考虑到由于缝隙腐蚀引起的缺陷尺寸增大而产生的屈服泄漏、断裂或其他方式造成失效的可能性。在缺乏足够的与缝隙腐蚀有关的资料时，任何表面的缝隙都是不可接受的。

为了确定构件上金属损失需要确定材料（如容器等）壁厚值。壁厚值的检测可以使用直梁式声波壁厚检测法（UT）。该法可以提供高精度的壁厚读数值。（有时测量表面需要进行清理，如使用钢丝刷，或使用砂轮等，若构件处在高温情况下，则需要使用温度补偿和专用的UT连接器）。

另外也可以用X射线检测（RT）技术。这些技术对环境的腐蚀和侵蚀会引起金属的损失，可能造成均匀减薄或局部减薄等都能够进行检测，并确定构件要求的最小壁厚 t_{min}。下面介绍有关腐蚀与侵蚀等失效模式安全评定的方法和步骤。

1）通过上面的方法确定构件金属损失的位置，以及测量的最小壁厚 t_{min}。

2）确定平均壁厚的长度 L 和计算剩余壁厚比 R_t

$$L = Q\sqrt{Dt_{min}}$$

式中，D 为圆筒壳、锥壳、球壳等的内径；Q 为剩余强度系数。

$$R_t = \frac{t_{min} - FCA}{t_{min}}$$

式中，FCA为未来的腐蚀裕量；t_{min} 为测量最小壁厚。

3）根据壁厚断面的数据，确定轴向和环向金属损失的长度 S 和 C，确定临界壁厚断面图。

4）若 $S<L$，则认为环向和径向上金属的损失程度就可以接受；$S>L$，则需要进一步的详细评定。

5）若金属的局部损失按此方法进行评定，则结果偏于保守，可以进行详细的评定。

更为详细的评定方法这里暂不作进一步的介绍，读者可查阅相关的文献。

4.3.2 环境促使开裂

当评定焊接结构的完整性时，还必需考虑到在使用条件下，环境促使开裂是否会引起亚临界裂纹扩展阶段。环境促使开裂是对应力腐蚀、腐蚀疲劳以及这两者的组合的一个比较综合的描述。这些失效机理间的相互影响和相互作用可能非常复杂，例如小循环载荷或非常缓慢单调递增的载荷会显著影响表面应力腐蚀门槛值和裂纹扩展速率。在所有场合中，腐蚀这个术语未必意味着表面整体氧化或溶蚀，而常常可能是指伴随着惰性保护氧化膜形成的轻微的一般性腐蚀。

由于这些相互作用的潜在复杂性以及各种环境组合的特殊性，所以在评定环境促使开裂时，只能使用与实际情况尽可能接近的条件下获得的数据。确定应力腐蚀的应力强度因

子门槛值 K_{ISCC}，裂纹扩展速率试验方法的一般性指导原则可查阅相关的文献和资料。在试验时，应该采用所关心材料的典型焊接件进行必要的试验，并应注意模拟环境的化学性质和温度；应采用合适的加载条件，包括适用于实际结构上的动载荷可能产生的任何影响。不应该考虑到化学瞬变现象的程度和持续时间、裂纹和缝隙化学介质变化的高概率、及氢脆过程可能影响埋藏缺陷的概率。

4.3.3　应力腐蚀

在应力（外加应力或内应力）和腐蚀介质共同作用下的腐蚀称为应力腐蚀。因此可以说应力腐蚀是由于腐蚀环境和外加静态应力或残余应力联合作用的结果，它包括金属溶蚀和氢脆引起的开裂等。应力腐蚀破裂一旦开始，只要保持驱动力并且局部环境维持不变，它通常会达到一个基本上均匀的速度。该速度的大小取决于具体的金属材料和环境。由于化学物质的集中，一个表面上没有任何危害的环境可能会引起表面发生局部腐蚀。工程上一般可采用以下几个参量来描述材料抵抗应力腐蚀断裂的能力：应力腐蚀断裂应力 σ_c，断裂时间 t_f，裂纹扩展速率 da/dt，应力腐蚀断裂的临界应力场强度因子 K_{ISCC}。

图5.4-45中给出了裂纹扩展速率 da/dt 与应力强度因子 K_{IC} 之间的关系。

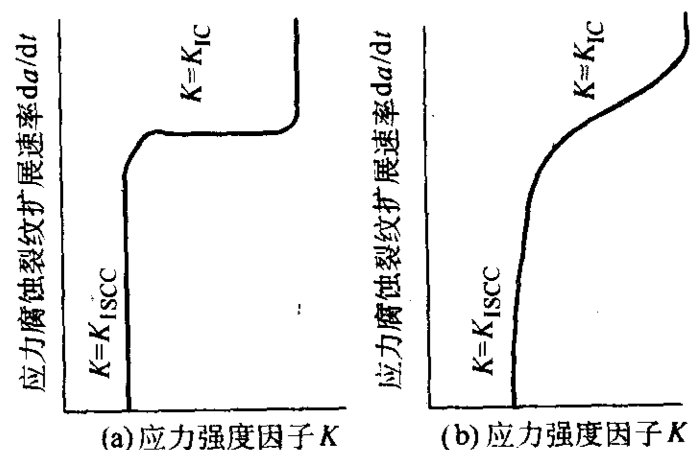


图5.4-45　裂纹扩展速率 da/dt 与应力强度因子 K_{IC} 关系曲线

而在评定应力腐蚀断裂时，临界应力场强度因子 K_{ISCC} 是一个非常重要的指标，因此这里对其值的确定方法加以简单介绍。

(1) 应力腐蚀断裂的临界应力场强度因子 K_{ISCC} 的确定

断裂通常是从表面开始的，对于长为 $2c$，深为 a 的半椭圆裂纹，可以通过式（5.4-193）估算 K_{ISCC} 的值。

$$K_{ISCC}^2=\frac{1.21\pi\sigma_f a}{\Phi^2-0.212\ (\sigma_f/\sigma_s)^2} \tag{5.4-193}$$

式中，σ_f 为断裂应力；σ_s 为材料的屈服强度；Φ 为第二类完全椭圆积分。

为了评定已知缺陷的严重性，需先根据静态载荷情况下安全评定中应力计算的方法（2.4.2节）估算外加应力强度因子 K，其中残余应力可能是最重要的。然后将经过计算所得的外加应力强度因子对应力腐蚀破裂敏感的门槛应力强度因子 K_{ISCC} 比较。可以初步认为

$$K_I<\left(\frac{1}{f}\right)K_{ISCC}\qquad 缺陷可接受$$

$$K_I>\left(\frac{1}{f}\right)K_{ISCC}\qquad 缺陷不可接受$$

式中，f 为安全系数（$f>1$），其值应该取得各方的同意。

在这些情况下，所需数据限于 K_{ISCC} 值。测定 K_{ISCC} 时应根据一些规定来进行。如果外加应力强度因子 $K_I>\left(\frac{1}{f}\right)K_{ISCC}$，应该认识到可能发生了应力腐蚀破裂扩展。这时，应该谨慎考虑是否需要采取适当的补救措施（如改变环境、除去缺陷或消除应力等）。同时还可以进行评定，决定在结构的设计寿命内或适当的检查周期内缺陷是否会扩展到不可接受的尺寸。这个评定应基于已有的应力腐蚀裂纹扩展速率数据，如果有必要，外推数据到低于应力强度因子 K_{ISCC}。外推的方法需征得各方的同意。这里将介绍一些相关的确定应力腐蚀裂纹扩展速率的建议。

在测定 K_{ISCC} 时应注意，试验条件的微小变化可能对腐蚀过程非常敏感。尤其是环境本身是一个重要的变量，必需格外小心以确保充分模拟服役条件。按常规应在静载进行 K_{ISCC} 的测定。然而还需注意到，结构很少承受纯粹的静载，众所周知，如果有一个即使是非常小的循环载荷叠加在静载荷上，K_{ISCC} 的值也可能要降低很多。因此采用在所研究的环境和载荷条件的疲劳裂纹扩展时的应力强度因子门槛值 ΔK_C，来评定缺陷的严重性常常更为合适。

如果肯定应用 K_{ISCC} 来评定是合适的，除了通过式（5.4-193）进行计算外，也可以使用相关的标准或规范来测定所研究焊接件的 K_{ISCC} 值。

(2) 确定应力腐蚀裂纹扩展速率

采用相关标准或规范中的开裂监测方法，在预制裂纹断裂力学试件的应力腐蚀试验中监测裂纹扩展速率，可获得所研究的焊接件的数据。这些数据使得应力腐蚀裂纹扩展速率 da/dt 的确定能以应力强度因子 K_I 的函数的形式来表示，如图5.4-45所示。

然后可用一个或多个形式来表示拟合所获得的数据，

$$da/dt=\alpha K^n$$

式中，α 和 K 均为常数。

另一方面，当设备使用中为应力腐蚀控制时，可以按合适的周期对设备进行适当的检测，以保守地估计扩展速率。这涉及到用从设备上取下的试样或在现场使用合适的金相技术，以认真地确定服役中的缺陷尺寸。显然，进行这样的测定必须相当严格和小心。如果决定必须用应力腐蚀破裂机理来评定缺陷的扩展，可将上述这些方程式进行积分，以估算在结构设计寿命期内，以及适当的检测周期内预期的裂纹扩展量。如果裂纹扩展后，缺陷没有达到其他失效机理（脆性断裂或韧性断裂）的最大容许尺寸，则缺陷就是可接受的。相反，如果计算发现在设计寿命终了或检测周期内，裂纹尺寸超过了其他失效机理的可接受尺寸，缺陷就是不可接受的。

4.3.4　腐蚀疲劳

腐蚀环境下的疲劳裂纹扩展要比在惰性环境中大。如图5.4-46所示。在不同的金属环境系统中已经观察到若干类型的腐蚀疲劳关系。这种复杂的行为意味着，如果没有所研究的材料、环境和载荷条件的可靠数据，必需采用下述的方法来确定腐蚀疲劳开裂的扩展数据。只有为数不多的情况下的描述腐蚀环境中疲劳裂纹扩展的方程是现成的（如一些海水环境下铁素体钢的裂纹扩展参数）。

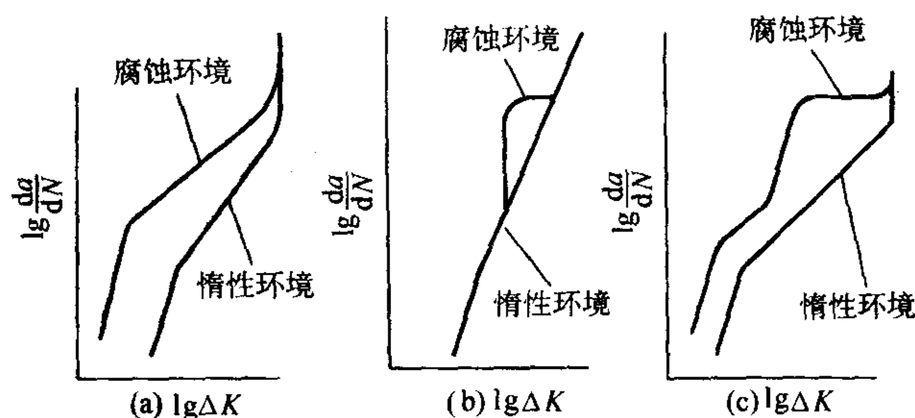


图5.4-46　腐蚀疲劳裂纹扩展行为的类型

在预制裂纹试样测定腐蚀疲劳数据时，可根据相关的标准或规范中的有关规定来进行测定。其中非常重要的是一定要采用能反映服役中实际环境状况、载荷循环和试验频率。除了温度和化学组成（包括充气程度）外，诸如流速、裂纹长度和形状、与异种金属接触或电化学极化等因素，均可以影响扩展速率，故应在试验中予以体现。在没有腐蚀疲劳试验的国家标准时，推荐用下述方法。

为了确定门槛值 ΔK_C（在特定试验条件下，高于 ΔK_C 时，腐蚀疲劳裂纹将起裂和扩展），需要非常近似地模拟合适的环境条件，这样就可以采用象疲劳门槛值试验一样的类似预制裂纹试件的几何尺寸和实验方案法。

类似的试件和方法还可以用来确定腐蚀疲劳的 da/dt 和 ΔK_C 的关系。在实际应用中若涉及随机加载循环载荷的周期变化已很清楚，最好是模拟工作条件，特别是应力比、循环频率和循环波的形式；如果试验时施加的循环频率太高，以至于无足够的时间发生腐蚀/扩展过程，或者频率太低而可能有再钝化效应，对环境影响的估计可能就不正确。在疲劳循环周期中施加的上升阶段时间的长短也可能会强烈影响疲劳腐蚀过程，所以在试验中应该采用恰当的循环波形。常用正弦波、三角波、锯齿波和方波的波形来模拟服役载荷条件，在适当的时候，还应在循环中施加有代表性的握执时间。

确定了合适的裂纹扩展关系后，应采用本章 2.4 节的建议进行评定。

4.3.5　点蚀

点蚀是金属局部腐蚀的一种。它的特征可以归结为：点蚀的大小与板壁厚度的数量级相同，或小到一个点直径的腐蚀；点蚀的深度小于板的壁厚。

（1）评定时所需数据

在对点蚀进行评定时，所需要的数据除了本章第 2 节中要求的数据外，还应包括点蚀的深度和直径，以及点蚀对的测量。在对每个点蚀进行测量时，把每个点蚀的金属损失看作一个当量的圆柱体。要想定义一个点蚀对，则需测量其每个点蚀坑的直径、深度、长度或点蚀中心之间的间距等。由于点蚀的出现一般是随机的，因此需要用户进行离散分析，以便选择点蚀对数量，使得该数量能恰当的代表构件中所有点蚀危害。在一个广布区内至少选取十个点蚀对进行测量。测量时建议使用扫描技术或射线照相技术（RT），且要对表面有锈皮、污物和已损坏的涂层进行清理（例如，采用喷砂处理），以便获得点蚀测量的精确结果。

（2）简化评定方法

如果所有点蚀的深度小于某一规定的腐蚀或冲蚀裕量，且壁厚较大，则不需要作进一步的工作。否则要进行评价。这里介绍一种简化的评定方法，其具体步骤如下。

1）评定所需的一些参数　D 为圆柱壳、锥壳、球壳等的直径（mm）；FCA 为估计的腐蚀裕量（mm）；RSF_a 许用剩余强度系数；t 为在点蚀危害区附近的当前壁厚（mm）。

2）确定点蚀对的相关参数　d_{ik} 为点蚀对 k 中点蚀 i 的直径（mm）；d_{ij} 为点蚀对 k 中点蚀 j 的直径（mm）；P_K 为点蚀对 k 中点蚀对的间隔距离（mm）；W_{ik} 为点蚀对 k 中点蚀 i 的深度（mm）；W_{jk} 为点蚀对 k 中点蚀 j 的深度（mm）。

3）所需的最小壁厚 t_{min}（mm）。

4）确定在所有点蚀对中深度小于 t_{min} 的点蚀，以及所有点蚀对 i 的深度均值 $\overline{W}_{i,k}$ 和所有点蚀对 j 的深度均值 $W_{j,k}$（如图 5.4-47 所示）。

在考虑所有读数值的情况下，计算平均点蚀深度 $\overline{W}_a$（如上所述，至少使用十个以上的点蚀对）。

$$\overline{W}_{j,k} = W_{j,k} - (t - \text{FCA} - t_{min}) \quad (5.4\text{-}194)$$

$$\overline{W}_{i,k} = W_{i,k} - (t - \text{FCA} - t_{min}) \quad (5.4\text{-}195)$$

$$\overline{W}_{a,k} = \frac{W_{j,k} + W_{i,k}}{2} \quad (5.4\text{-}196)$$

$$\overline{W}_a = \frac{1}{n}\sum_{k=1}^{n} W_{a,k} \quad (5.4\text{-}197)$$

式中，K 为第 K 个点蚀对；n 为检查中所记录的点蚀对总数。

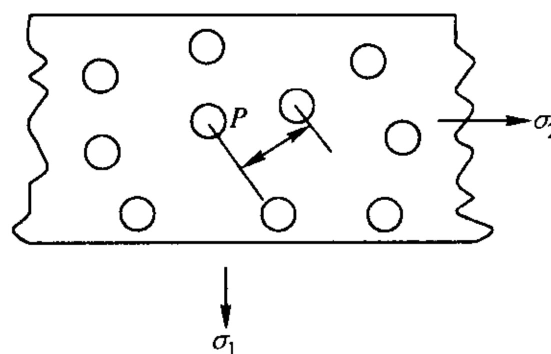


图 5.4-47　点蚀分析的参数

5）确定点蚀的平均直径 d_a、点蚀对之间的距离 P_a。其中 d_a 为检查记录中所有点蚀对每个点蚀直径的平均值，而 P_a 为所有点蚀对之间，即最近的邻居之间的平均距离。

$$d_{a,k} = \frac{(d_{ik} + d_{jk})}{2}$$

$$d_a = \frac{1}{n}\sum_{k=1}^{n} d_{a,k}$$

$$P_a = \frac{1}{n}\sum_{k=1}^{n} P_k$$

6）计算剩余强度系数 RSF。若 $\overline{W}_a \leqslant 0.00$，取 RSF = 1.0，即满足简化评定准则，进入步骤 8，否则使用下面的公式进行计算 RSF 值。

$$\text{RSF} = \min\left[1.0 - \frac{\overline{W}_a}{t_{min}} + \frac{E_a\ (t - \text{FCA} + \overline{W}_a - t_{min})}{t_{min}},\ 1.0\right] \quad (5.4\text{-}198)$$

式中，$E_a = \frac{\sqrt{3}}{2}\mu_a$；$\mu_a = \frac{P_a - d_a}{P_a}$。

7）按点蚀类型进行评定

① 广布型点蚀　对于出现在构件上的一个广泛区域内的广布型点蚀若有 $RSF \geqslant RSF_a$（RSF_a 为许用剩余强度系数），则点蚀可以接受，否则不可接受。

② 局部点蚀　若点蚀的危害是在一个局部区域内则首先确定危害面积。它是金属局部损失的一个当量区域（LTA）。当量 LTA 轴向和环向尺寸都应当按照所观察得到的，点蚀的实际边界加以确定。对于该当量 LTA 的当量厚度 t_{eq} 可由下式来确定。

$$t_{eq} = \text{RSF} \times t_{min}$$

③ 位于一个广布点蚀面积内的金属损失区，一个局部金属损失区（LAI）位于一个广布点蚀的面积内。

在该情况下要确定组合的剩余强度系数 RSF_{com}，其值可由下式确定

$$RSF_{com} = RSF_p \times RSF_e$$

式中，RSF_p 为点蚀危害后的剩余强度系数；RSF_e 为一个局部减薄区的剩余强度系数。RSF_e 若有 $RSF_{com} \geqslant RSF_a$ 则认为点蚀是可以接受的。

④ 在局部金属损失区内确认的点蚀，若点蚀的危害是出现在一个局部金属损失区，可以采用上面③的方法进行评定。

8）单个点蚀尺寸校核的推荐界限值。

① 点蚀的直径 d　对单个的点蚀，若它不满足下面的方程，则可把它看作是局部减薄区域加以评定。该局部减薄区域的尺寸即为该点蚀的直径 D 和下面所定义的剩余厚度比。尤其对大的点蚀更需要这一校核，以保证不在点蚀的底部出现局部的穿透带。

$$d \leqslant Q\sqrt{Dt_{\min}}$$

式中，d 为点蚀的直径；D 为圆柱壳、锥壳、球壳等的直径，mm；Q 为需要剩余强度系数（可以查阅有关的表）。

② 点蚀深度 R_t　建议采用下式的剩余厚度比的界限值，以防通过针孔型泄漏而发生的局部失效。

$$R_t \geqslant 0.20$$

构件不满足以上的评定要求，应对其进行详细评定（这里不加以介绍），或进行修理、补救或替换、退役等。

4.4　失稳（或变形过大，压曲）

对一些承受压应力的部件，需要对部件或结构产生压曲或变形过大等方面来进行评定。从这方面来讲，需要考虑的缺陷情况为以下三种：

1）可能减少有效局部横界面的缺陷；

2）平行于零部件表面的分层或层状缺陷；

3）与设计尺寸不符的缺陷（如加工问题，或一些错边、角变形、翘曲等缺陷）。

总的说来，在考虑失稳的失效的情况下，对以上缺陷进行评定时，应该综合考虑到疲劳、腐蚀、应力腐蚀（以上情况的分析和评定方法在前面章节中已经介绍过）、蠕变等情况。同时，评定还需要按设计寿命完结时的缺陷尺寸进行失稳分析。若结构已经出现翘曲，但仍在运行，尽管一些缺陷不会对初始的抗压强度有影响，但有可能对翘曲后构件的变形能力有影响，因此一定要根据现有的实际情况进行评定。下面对以上三种可能促使失稳失效的缺陷的评定给予分别的介绍。

（1）可能减少有效局部横界面的缺陷

在计算时需要保证任何一个横界面上，构件的抗压曲强度（考虑了各种缺陷的总集合面积）要大于最大的外载荷（在计算时为确保安全，通常对最大外载乘上一定的安全系数）。同时既要考虑缺陷对局部横截面减小的影响、缺陷在部件局部的位置情况，以及缺陷的存在对构件造成的偏心，且在计算时还要考虑横截面的二次矩（注意：在计算二次矩时应不计投影在垂直于压应力 P_m+P_b+Q 平面上的缺陷面积）。

如果缺陷平行于承受压应力的构件表面，且在构件表面与加强肋的连接焊缝下，则该缺陷即会减少加强肋与构件的有效连接长度。通常此类缺陷可能具有分层、层状撕裂或焊道下裂纹的一些性质。因此，若发现了此类缺陷，在对其进行评定时，应假定加强肋是间断焊接在构件上，缺陷在两端焊缝中间形成了“间隙”。有关此“间隙”的允许范围可以查阅有关的规定。

（2）平行于零部件表面的分层或层状缺陷

在对该类缺陷进行评定时，首先要对缺陷和构件表面材料每一层的压曲强度校核，然后再对压应力方向平行于构件表面的平面缺陷（分层或层状撕裂等）进行评定。在评定过程中，应把每一层材料看成是和缺陷面积统一的一个一个的单独的“板”，把缺陷和表面间的距离作为单独板的有效厚度，而且还应假定“板”是沿着缺陷边缘为边界，是沿着边界简支的。

（3）与设计尺寸不符的缺陷

此类情况又可分为两种，一种是加工尺寸与设计不符，这时应重新进行校核；另一种是与设计不符的一些缺陷，包括错边、角变形、翘曲等。这类缺陷可能对承受压应力的构件的压曲强度有很大的影响。有关此类缺陷的规定极限尺寸根据不同的构件材料和工况按相关的规定来进行评定和判定。

4.5　蠕变

对于在温度较高的条件下运行的设备，必须考虑其可能发生的蠕变引起的焊接缺陷的扩展，或者由于焊缝缺陷发生蠕变而引起周边材料韧带的抗蠕变能力。

对以上情况进行分析和评定时，首先要判定焊接构件是否在蠕变的范围内工作。若焊接缺陷是在非蠕变区域内工作，可以免于蠕变评定，则无需进一步考虑有关蠕变的问题；若其不能满足排除蠕变的准则（排除准则，即构件/温度/服役寿命的组合满足一定的条件），同时参考以往的经验，表明材料对蠕变裂纹的扩展不敏感，则可以根据基于剩余净截面不变观点的蠕变破坏计算来评定缺陷（其中可以借鉴一些相关的设计标准），同时还需按前面章节的有关说明进行断裂和塑性破坏的评定和分析。

而对于根据以往经验认为缺陷有可能发生裂纹扩展时，目前还没有非常具体的标准或规定，这里建议应用断裂力学方法来进行评定。

下面分别介绍与蠕变评定直接相关的蠕变排除准则和焊接微观结构方面的分析方法。

（1）蠕变排除准则

在进行以下所介绍的蠕变排除准则判定时需要注意，这些指导原则依赖于实际研究材料的蠕变数据，或某类合金的最低性能数据，且在只有平均数据的情况下，需要采用一定的安全系数。

对一个构件，若其满足下列条件之一时，可认为蠕变现象缓解。

1）在整个运行周期内，最高温度低于温度 T_0，其中 T_0 为结构材料和运行周期的特性值。对于在所讨论的时间/温度范围内，单轴蠕变断裂伸长率大于10%的材料，T_0 是指构件运行周期内，在相当于“试验强度”的应力水平下产生0.2%的累计蠕变所对应的温度，如图5.4-48所示。

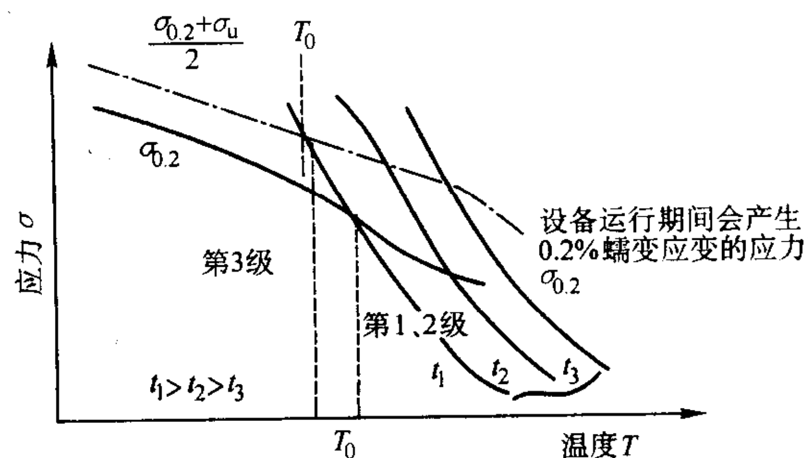


图5.4-48　温度 T_0 的确定

当 $L_r<1$ 时，“试验强度”等于工作温度下的 $\sigma_{0.2}$，而对于 L_r 可以大于1的情况下，“试验强度”等于工作温度下的 $(\sigma_{0.2}+\sigma_u)/2$。

表5.4-19给出了一些材料的 T_0 值，这些 T_0 值是基于一些已经公开发表的平均数据和 2×10^5 h的寿命。对于其他钢材或更长的寿命，则需从有关材料的蠕变试验数据来得出。

表 5.4-19 寿命为 2×10^5，可以忽略蠕变情况的最高温度

钢种	$T_0/℃$	
	简单评定和常规评定	分析评定
C-Mn 钢	345	310
铁素体	440	420
0.5Cr0.5Mo0.25V、1Cr0.5Mo、0.25Cr1Mo、9Cr1Mo、12CrMoV（W）、316	485	①

① 目前还缺乏足够的数据。

对于在前面所述的时间和温度范围内单轴蠕变断裂伸长率小于 10% 的材料，应该按实际的蠕变伸长率 1/50 的蠕变相对应的应力来确定 T_0。

2）构件的总体运行周期/温度历史满足式（5.4-199）。

$$\int 1/t(T)\mathrm{d}t < 1 \qquad (5.4\text{-}199)$$

式中，$t(T)$是对于蠕变断裂伸长率大于 10% 的时间。

在恒温 T 下，在等于"试验强度"的应力水平下，达到 0.2% 累计蠕变应变所需的时间，如图 5.4-49 所示。这里需要注意，应采用有关材料的蠕变断裂性能来获得时间 $t(T)$。

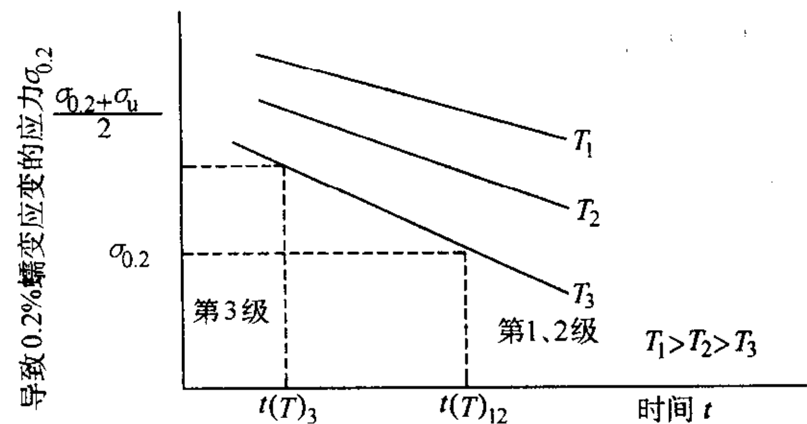


图 5.4-49 在应力水平下等于"试验强度"下达到 0.2% 累计蠕变应变所需时间 $t(T)$ 的确定

对于伸长率小于 10% 的材料，$t(T)$需按等于蠕变断裂伸长率 1/50 的蠕变应变来确定。

(2) 焊接构件的微观结构

前面部分的分析需要知道被评定缺陷所在部位材料的蠕变断裂性能。对于实际的热影响区，很难测得这些材料数据，但已经有了一些情况的微观结构性能的资料数据。若可以说明母材的抗蠕变性能不比热影响区好，则可以根据母材的性能来确定 T_0 和 $t(T)$。

除了蠕变本身以外，蠕变和疲劳的相互作用也不容忽视。但由于相关的数据不充分，无法给出统一的方案，但可以根据具体情况和专家的经验来考虑相关事宜。

5 概率安全评定方法

通常的基于材料力学、弹塑性断裂力学等安全评定方法都属于确定性评定方法，如前面部分介绍的安全评定方法即属于确定性的评定方法。按确定性方法评定过程中，所采用的各参数为确定值，然后通过取一定的安全系数，来保证判断的安全性，给出要么安全、要么不安全的评定。它一般不能用于定量比较构件的危险程度和安全程度。

而在实际当中，材料性能的实测数据具有一定的分散性，服役过程中有时由于损伤材料性能还要发生变化，构件的几何尺寸在公差范围内也存在波动；外载荷也同样具有不同程度的随机性，因而危险点的应力实际上是随机的；另外该处材料的强度也是随机的；同时缺陷的检测也具有一定程度的不确定性。由此可见，安全评定中应考虑上述的随机性，评定作出的判断也应该是不确定的，即评定作出的判断应该是危险程度和安全程度的概率。

为了进行概率安全评定的研究，首先要对评定中参数的不确定性进行分析和研究。

5.1 安全评定分析中不确定性影响因素

材料的基本性能参数是具有不确定性的，它们服从一定的统计分布规律。例如强度、断裂韧度以及疲劳裂纹扩展时裂纹尖端材料特性等都具有一定的波动性。

由于构件的形状和尺寸在形状和尺寸公差允许范围内波动，引起不确定性，使得应力和强度等产生一定的统计分布规律。

外载荷的随机变化，也会引起危险点应力的不确定性，使得应力产生一定的统计分布规律。

材料初始损伤程度的无损检测水平，造成初始损伤程度估计的不确定性。

上述四种不确定性影响因素，即材料缺陷尺寸的不确定性、材料基本性能的不确定性、构件形状尺寸的不确定性以及外载荷的不确定性，一般是服从一定的统计分布规律的，可以通过概率统计的方法描述，它们的不确定性也可称为统计不确定性。由此，即可分析它产生的统计分布特性，从而实现对不确定参数进行定量的描述。

由于损伤物理（或力学）模型过于简化而引起的不确定性，这类不确定性与失效分析中所作假设有关，是非统计不确定性，例如采用 Miner 线性累积损伤理论描述疲劳损伤过程所带来的误差、谱载荷作用时没有考虑载荷序影响、实验数据拟合和外推所引起的误差等。

下面就具有统计不确定性的参数进行较为详细的随机性分析。

5.1.1 材料静强度指标的随机性

材料成分和组织结构的不均匀性、使用工况条件的差异以及测量误差造成了材料性能参数的不确定性。屈服强度 σ_s 和抗拉强度 σ_b 等常规力学性能参数通常认为能较好地符合或近似符合正态分布。多数材料的延伸率 δ 符合正态分布。抗剪强度 τ_b 与 σ_b 有近似线性关系，故近似服从正态分布。

多数材料的硬度接近正态分布，也常能较好地符合威布尔分布。

金属材料的弹性模量 E、切变弹性模量 G 及泊松比 ν 也具有离散性，可认为近似服从正态分布，但其分散性比较小，可视为确定性参数。

一般情况下，变异系数 $C_X<0.2$，若 $C_X>0.2$，则说明材料的生产工艺不稳定，应当进行调整或改进以保证产品质量的稳定性。

5.1.2 材料断裂韧度的随机性

一般材料的疲劳寿命、疲劳强度和断裂韧度，比材料静强度指标的分散性大。

在一般条件下，材料断裂韧度 K_{IC} 可以认为服从正态分布、对数正态分布或威布尔分布。

有时从韧性很高的钢材直接测定 K_{IC} 值比较困难，于是借助于夏氏 V 型缺口冲击能量 C_V 测得。C_V 和 K_{IC} 之间的关系式为

$$K_{IC}=\sqrt{5\ (C_V\sigma_s-0.05\sigma_s^2)}$$

式中，各参数均为英制单位，K_{IC} 的单位为 ksi $\sqrt{\text{in}}$，屈服极限 σ_s 的单位为 ksi，而 C_V 的单位为 ft·lbf。

这时的断裂韧度可以由 C_V 和 σ_s 的分布确定 K_{IC} 分布。

上述各英制单位与常用单位的换算关系是：

1ksi = 6.894 8 MPa;　　1in = 25.40 mm;

1lbf = 4.448 2 N;　　1ft = 0.304 8 m;

1ft·lbf = 1.355 8 N·m。

为了简化计算可以认为其服从正态分布。

5.1.3　几何尺寸的随机性

生产中，构件的容许偏差为公差，如果构件尺寸偏差范围为 $\bar{x}\pm\Delta x$，这时均值为 $\bar{x}$，标准差为

$$\sigma=\frac{(\bar{x}+\Delta x)-(\bar{x}-\Delta x)}{6}=\frac{\Delta x}{3}$$

如果其他性能指标在手册中给出容许偏差范围时，都可按上述“3σ 规则”近似估计均值和标准差。

对于由 k 个构件组成的组件，若第 i 个构件容许偏差范围为 $x_i+\Delta x_i$，则组件的尺寸均值及标准差分别为：

$$x_\Sigma=x_1+x_2+\cdots+x_k$$

$$\sigma_{x_\Sigma}=\sqrt{\sigma_{x_1}^2+\sigma_{x_2}^2+\cdots+\sigma_{x_k}^2}=\sqrt{\left(\frac{\Delta x_1}{3}\right)^2+\left(\frac{\Delta x_2}{3}\right)^2+\cdots+\left(\frac{\Delta x_k}{3}\right)^2}$$

构件的几何尺寸在尺寸公差范围内波动。若它们的尺寸公差和测量误差相对较小，在评定中可以忽略其随机性，可以认为它们是确定性参数。

5.1.4　构件危险点的疲劳应力随机性

恒幅载荷的作用下，由于构件的形状和尺寸在公差允许范围内波动以及危险点材料性能的随机波动等，危险点的应力服从一定的统计分布，称为疲劳应力的横向分布。所有使疲劳应力产生横向分布的因素称为横向影响因素。所以横幅循环载荷作用时，疲劳应力的随机性是由横向影响因素造成的。

复杂的载荷-时间历程作用时，构件危险点的应力将是随机变量或随机过程，而峰值和均值应力服从一定的统计分布，称这种疲劳应力分布为疲劳应力的纵向分布。使疲劳应力产生纵向分布的因素称为纵向影响因素。所以构件承受随机载荷或变幅载荷时，疲劳应力除了受横向影响因素作用外，还受纵向影响因素影响。

5.1.5　缺陷参数的不确定性分析

通常根据缺陷的形态可分为平面缺陷和体积缺陷。如前所述，在评定中需将表面平面缺陷规则化为长轴 $2c$、短轴为 a 的半椭圆缺陷，将埋藏平面缺陷规则化为长轴为 $2c$、短轴为 $2a$、与最近表面的距离为 p_1 的椭圆状缺陷。而这些参数都是由无损检测的方法来确定的。虽然无损检测技术近年来有了很大发展，但由于构件材料、形状、位置、检测方法、检测手段、缺陷取向等诸多因素造成了检测结果具有不确定性和分散性。

对其不确定性和分散性一般可以通过对多次检测结果进行数理统计分析，得到其均值、变异系数和分布形式。但由于时间和资源等的限制，有时无法完成多样本的检测，只能获得参数的均值，可以通过以往评定的经验来确定其变异系数和分布形式。

5.1.6　初始裂纹尺寸 a 的随机性

由于材料初始损伤程序不同，无损检测技术水平，即裂纹检出概率等影响，以及小尺寸裂纹检测的困难（$a<0.5$ mm），使得构件在制造和使用过程中初始裂纹分布估计困难。为了便于损伤容限设计，常用当量初始裂纹尺寸 EIFS（equivalent initial flaw size）表征结构原始质量。利用它作为计算起点，将结构裂纹形成寿命和裂纹扩展寿命之和仅表示为扩展寿命。所以当量初始裂纹尺寸分布并不是构件初始裂纹的真实分布，而是为了寿命估算设定的裂纹分布。

值得注意的是：铆接件产生可观察裂纹（例如 $a>0.5$ mm），经历的时间较长，当量初始裂纹尺寸 EIFS 用于寿命估算，将带来很大方便。焊接焊缝中有大量气孔和夹渣等体积型缺陷，这些缺陷呈现一定分布，产生可观察裂纹（例如 $a>0.5$ mm），经历时间较短，存在实际缺陷扩展产生的真实裂纹分布。

5.1.7　失效物理（失效力学）模型过于简化造成的不确定性

失效物理模型的建立借助于工程力学、材料学、数理统计方法和可靠性工程学。失效物理模型过于简化也会造成一定的偏差，例如对含宏观裂纹构件仍用材料力学方法进行安全评估，显然会得出不安全的结论，因为含裂纹构件安全评估需要采用断裂力学方法。

5.2　含裂纹结构缺陷的统计方法

使用各种无损检测手段对结构关键部位进行检测，以发现存在的裂纹和缺陷。实践证明：无论哪一种无损检测手段都不能百分之百地将所有不同尺寸的裂纹检测出来。为了定量描述无损检测手段可靠性的高低，人们采用裂纹检出概率曲线。

在一定检测条件下（检测手段、仪器、被检结构材料、外形、环境、检验员技术水平及责任心等条件），尺寸为 a 的裂纹在一次检查中被检出的概率称为裂纹检出概率，记为 $P(D|a)$。$P(D|a)\sim a$ 曲线被称为裂纹检出概率曲线。

5.2.1　裂纹检出概率曲线测定方法

裂纹检出概率曲线是采用实验室模拟测定法作出。这种方法要求：

1）在试件材料、裂纹形式、含裂纹部分的局部外形等方面模拟真实情况；

2）制作总数不少于 30 个试件，其中一半左右含有不同长度的裂纹，其余不含裂纹；

3）用指定仪器，按规定条件对试件反复进行无损检测，记录检测总数和检出裂纹次数；

4）检测人员不少于 5 人，全部检测工作必须独立进行，不得相互交流；

5）一个检验员对同一试件的两次检测，在时间上至少要间隔一天。

具体测定方法是：

1）将裂纹长度范围分成若干区间。

2）对同一裂纹长度区间内裂纹的 n 次检测，若试验中实际检出此区间的裂纹数为 S_n，则该区间裂纹检出概率 p 的点估计为

$$\hat{p}=S_n/n$$

3）为了可靠起见，规定置信水平 $1-\alpha$ 的裂纹检出概率 p 的置信下限 P_L，作为该区间裂纹检出概率，近似计算公式为：

$$P_L=\frac{n_2}{n_2+n_1F_\alpha(n_1,\ n_2)}\tag{5.4-200}$$

式中，$n_1=2(n-S_n+1)$ 为 F 分布的上自由度；$n_2=2S_n$ 为 F 分布的下自由度；$F_\alpha(n_1,\ n_2)$ 为 F 分布关于 α 的上侧分位数，即满足 $P[F(n_1,\ n_2)>F_\alpha(n_1,\ n_2)]=\alpha$。

4）对每个裂纹长度区间，根据裂纹检出数据，作图 $\hat{P}-a$ 和 P_L-a 曲线，为裂纹检出概率曲线，如图 5.4-50 所示。

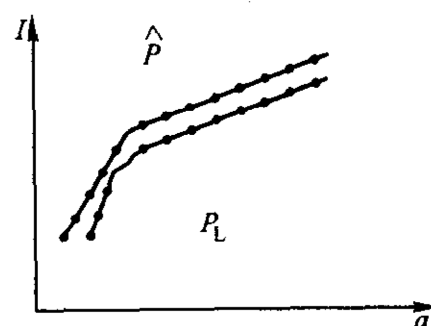


图 5.4-50　裂纹检出概率曲线

美国空军损伤容限标准 USAF MIL-A-83444 要求在规定在 95%的置信水平下裂纹检出概率不小于 90%。这一要求（记为 90/95）是比较严格的要求，通过试验确定裂纹检出概率的置信下限 P_L 时，独立检测次数几乎总是大于 30 的。例如，满足 90/95 要求时，对同一长度区间内裂纹的独立检测次数 n 和允许的最大漏检次数必须满足下述条件：

独立检测次数	最大漏检次数
29	0
46	1
61	2
75	3
89	4
103	5

5.2.2　裂纹检出概率曲线的拟合

人们根据现有裂纹检测资料，提出了多种裂纹检出概率曲线的函数形式，供拟合检测结果使用。下面介绍其中四种。

（1）幂函数型

$$P(D|a)=\begin{cases}\left(\dfrac{a-a_1}{a_2-a_1}\right)^m & (a_1\leqslant a\leqslant a_2)\\ 0 & (a<a_1)\\ 1 & (a>a_2)\end{cases}\tag{5.4-201}$$

式中，a_1、a_2 和 m 为由试验确定的常数。

（2）指数函数型

$$P(D|a)=\begin{cases}c_1\ \{1-\exp[-c_2(a-a_0)]\} & (a>a_0)\\ 0 & (a\leqslant a_0)\end{cases}\tag{5.4-202}$$

式中，c_1、c_2 和 a_0 为由试验确定的常数。

（3）双参数威布尔函数型

$$P(D|a)=\begin{cases}1-\exp[-(\alpha/\beta)^{\alpha}] & (a>0)\\ 0 & (a\leqslant 0)\end{cases}\tag{5.4-203}$$

式中，α 和 β 为由试验确定的常数。

（4）三参数威布尔函数型

$$P(D|a)=\begin{cases}1-\exp\{-[a-\varepsilon]/\beta)^{\alpha}] & (a>\varepsilon)\\ 0 & (a\leqslant\varepsilon)\end{cases}\tag{5.4-204}$$

式中，α、β 和 ε 为由试验确定的常数。

5.2.3　确定初始裂纹尺寸的概率分布

如果 $f_a(x)$ 表示构件中原始裂纹的概率密度函数，构件中存在尺寸为 x 的缺陷的概率为

$$P(x)=f_a(x)\,\mathrm{d}x\tag{5.4-205}$$

尺寸为 x 的缺陷存在并且被检出的概率为：

$$P(D|x)\,f_a(x)\,\mathrm{d}x\tag{5.4-206}$$

所以结构中各种尺寸裂纹被检出的概率为：

$$P(D)=\int_0^{\infty}P(D\mid x)f_a(x)\mathrm{d}x\tag{5.4-207}$$

如果 $F(x)$ 表示检测出的裂纹尺寸的概率密度函数，$P(x|D)$ 表示检测出尺寸为 x 的裂纹的概率，则由 Bayes 公式可知：

$$P(x\mid D)=\frac{P(D\mid x)P(x)}{P(D)}=\frac{P(D\mid x)f_a(x)\mathrm{d}x}{\int_0^{\infty}P(D\mid x)f_a(x)\mathrm{d}x}\tag{5.4-208}$$

如果拟认为 $P(x|D)=f_D(x)\,\mathrm{d}x$，则

$$f_D(x)=\frac{P(D\mid x)f_a(x)}{\int_0^{\infty}P(D\mid x)f_a(x)\mathrm{d}x}\tag{5.4-209}$$

式中，$P(D|x)$ 为由裂纹检出概率曲线得到；$f_D(x)$ 为无损检测出的裂纹概率密度函数，由无损探伤检测数据分析得到。

所以结构中初始裂纹的概率密度函数为

$$f_a(x)=\frac{f_D(x)}{P(D\mid x)}=\int_0^{\infty}P(D\mid x)f_a(x)\mathrm{d}x\tag{5.4-210}$$

$f_D(x)$ 可以根据裂纹初始分布规律，以及裂纹扩展规律和以往经验，通过式（5.4-210）迭代求出。

如果检查出的裂纹全部进行了修理清除，则修理清除后存在的是漏检裂纹，漏检裂纹的概率密度函数为

$$f(x)=[1-P(D|x)]\,f_a(x)\tag{5.4-211}$$

5.2.4　当量初始裂纹尺寸概率分布的确定

这种方法将结构的全寿命当量用裂纹扩展寿命表示，即裂纹形成寿命和裂纹扩展寿命之和，仅仅表示为裂纹扩展寿命。当量初始裂纹尺寸 EIFS（equivalent initial flaw size）不是结构中真实裂纹尺寸。

（1）用断口观测反推法估计 EIFS 的概率分布

试件在规定裂纹载荷谱下进行试验，对疲劳断口上形成的疲劳条带进行观测，得到一系列表示裂纹扩展的数据。然后用解析的裂纹扩展公式拟合观测结果，反推到时间为零处，得到当量初始裂纹尺寸。反复重复此过程，即可得到当量初始裂纹尺寸的概率分布。

TTCI（time to crack initiation）是指裂纹尺寸达到给定值时的疲劳寿命，如图 5.4-51 所示。通常通过 TTCI 分布和一个确定性裂纹扩展方程相结合确定 EIFS 分布。

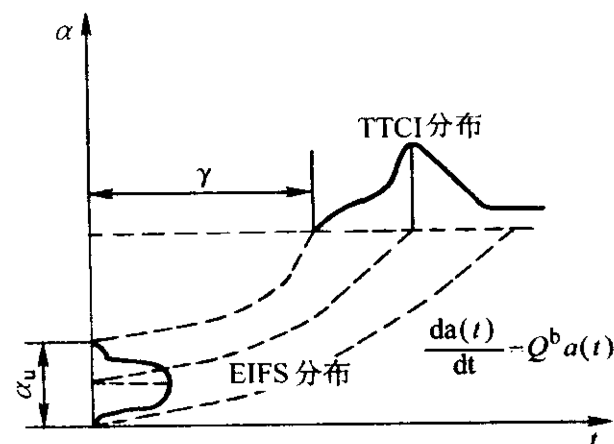


图 5.4-51　TTCI 分布和 EIFS 分布

现在举例说明如何确定 TTCI 分布。假设在某应力作用下通过试验，获得了 l 个有效断口（l 个试件）。第 k 个断口（第 k 个试件）有 m 对（a，t）数据，设短裂纹扩展公式为

$$\frac{\mathrm{d}a(t)}{\mathrm{d}t}=Qa^b(t)\tag{5.4-212}$$

式中，Q、b 为实验测定常数。此处为了讨论方便，设 $b=1$。

设第 k 个断口的 Q 值为 Q_k（$k=1, 2, \cdots, l$），将式（5.4-212）分离变量，得

$$\mathrm{d}t=\frac{1}{Q_k}a^{-1}\mathrm{d}a\tag{5.4-213}$$

式（5.4-213）由 t_i 到 t_j 积分（$t_i>t_j$），得：

$$a(t_j)=a(t_i)\exp[-Q_k(t_i-t_j)]$$

$$\ln[a(t_i)/a(t_j)]=Q_k(t_i-t_j)$$

式中，$i>j$，$j=1, 2, \cdots, i-1$；$i=2, 3, \cdots, m$。最小二乘法拟合确定 Q_k 为

$$Q_k = \frac{\sum (t_i - t_j)\ln[a(t_i)/a(t_j)]}{\sum[(t_i - t_j)^2]} \qquad (5.4\text{-}214)$$

通过内插到外推法，由式（5.4-213）或式（5.4-214），可以得到第 k 个断口任意给定裂纹尺寸 a_i 时的时间 $(T_k)_i$，由 l 个有效断口在任意给定裂纹尺寸 a_i 时的时间 $(T_1)_i$，$(T_2)_i$，…，$(T_l)_i$，可得到对应于裂纹尺寸 a_i 的 TTCI 分布。如果对每个断口由已知 Q_k 外推到 $p=0$ 时刻，可以得到零时刻当量初始裂纹值，通过统计分析这组当量初始裂纹值，得到当量初始裂纹分布尺寸 EIFS。

各断口的 Q_k 值是不相同的，并不描述给定应力水平下的裂纹扩展平均规律，在工程上可以用平均 TTCI 值 T_i 来估计给定应力作用下裂纹扩展曲线的 Q_i。

$$T_i = \frac{1}{l}\sum_{k=1}^{l}(T_k)_i \qquad (5.4\text{-}215)$$

式中，$(T_k)_i$ 为第 k 个断口与 a_i 相对应的 TTCI 值。

这样对每个 a_i 就有相应的 TTCI 平均值 T_i，即 (a_i, T_i) 是裂纹扩展平均曲线上的点，再代入式（5.4-215）就可以求出给定应力作用下裂纹扩展平均曲线的 Q_i。如果 $b\neq1$，同样可用最小二乘法确定 TTCI 分布和裂纹扩展平均曲线的 b 值。

（2）已知 TTCI 分布反推 EIFS 的分布

用 T 表示被研究构件的 TTCI，A_i 表示 EIFS。给定裂纹尺寸 a_F 的寿命 T 与 A_i 之间的函数关系

$$A_i = \varphi(T, a_F) \qquad (5.4\text{-}216)$$

其反函数为 $T=\Psi(A_i, a_F)$ 或 $t=\Psi(A_i, a_F)$，其中，t 和 a_i 分别为随机变量 T 和 A_i 的取值。

当 a_F 一定时，t 和 a_i 之间具有单调递减函数关系。若设 T 的概率密度函数为 $f_T(T)$，可求函数 a_i 的概率密度函数为：

$$f_{A_i}(a_i) = f_T[\Psi(a_i, a_F)]\left|\frac{dt}{da_i}\right| \qquad (5.4\text{-}217)$$

一般规定的裂纹尺寸 a_F 较小，在短裂纹尺寸范围，如果裂纹扩展速率可表示为

$$\frac{da(t)}{dt} = Qa^b(t) \qquad (5.4\text{-}218)$$

式中，Q、b 为实验测定常数，对实验数据具有较好拟合作用。为了讨论问题方便，令 $b=1$。

对式（5.4-218）（$b=1$ 情形）两边积分，得

$$t = \ln\left(\frac{a}{a_i}\right)^{1/Q} \qquad (5.4\text{-}219)$$

$$a_i = a_F e^{-Qt} \qquad (5.4\text{-}220)$$

如果已知 T 服从三参数威布尔分布，分布函数为

$$F_T(t) = P(T\leqslant t) = 1-\exp\left[-\left(\frac{t-\gamma}{\beta}\right)^{\alpha}\right],\quad t\geqslant\gamma$$

根据式（5.4-220），$A_i = a_F e^{-QT}$，于是得到 A_i 的分布函数

$$\begin{aligned} F_{A_i}(a_i) &= P(A_i\leqslant a_i) = P(a_F e^{-QT}\leqslant a_i) \\ &= P\left(T\geqslant\frac{\ln a_F-\ln a_i}{Q}\right) = 1-F_T\left[\frac{\ln a_F-\ln a_i}{Q}\right] \\ &= \begin{cases} \exp\left[-\left(\dfrac{\ln a_F-\ln a_i-Q\gamma}{Q\beta}\right)^{\alpha}\right] & (0<a_i<a_u) \\ 1.0 & (a_i\geqslant a_u) \end{cases} \end{aligned}$$

式中，a_u 为 A_i 取值上界，根据 $\ln a_F-\ln a_i-Q\gamma\geqslant0$ 得到：

$$a_u = a_F e^{-Q\gamma}$$

如图 5.4-51 所示。

如果已知 TTCI 分布概率密度函数，即给定 a_F 时 $f_T(T)$ 已知，将式（5.4-219）代入式（5.4-217），可得 EIFS 概率密度函数。所以可以根据问题由 TTCI 分布求 EIFS 分布，或由 TTCI 概率密度函数求 EIFS 概率密度函数。

这种方法的缺点是，所得的 EIFS 的概率分布为一种诱导分布，一般比较复杂，使用起来不甚方便。

5.3 静载下（含平面缺陷）概率安全评定方法

根据本章第2节的有关内容，即设备平面缺陷静载下弹塑性断裂的安全评定技术为基础，选 L_r 和 K_r 为基本变量；根据本节前面对各参数的（不确定性和随机性）分析及其算法来确定 L_r 和 K_r 的均值、变异系数和分布形式以及相关系数，求出失效率，然后进行概率安全评定。失效率通常可以采用三种方法计算，即积分法、第二水平法和蒙特卡洛法。通过对这三种方法的归纳总结可以得知，积分法需要大量的先验信息，并且只能通过近似计算得到数值解；蒙特卡洛法需要信息量相对较少，计算方法简单，但往往需要进行上百万次的模拟，才能获得一次“命中”，所以运行成本很高；而第二水平法则兼具二者的优点，又避免了其二者的不足之处，即需要的信息量相对较少，计算复杂程度相对适中，解的精度能够满足工程的需要，运行成本低，因此这里介绍第二水平法对构件的失效率进行计算。计算的流程图如图 5.4-52 所示。

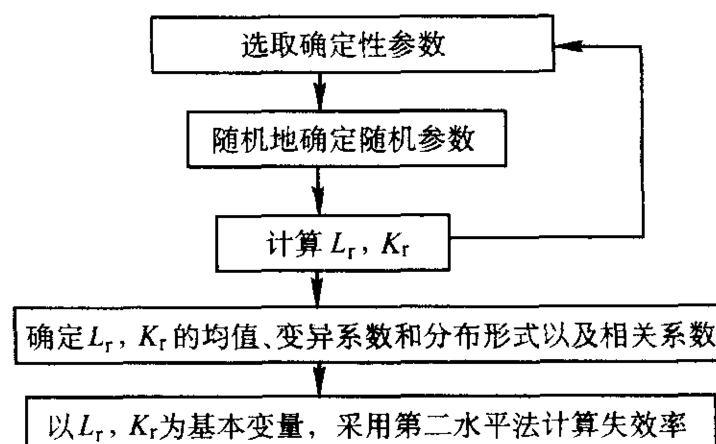


图 5.4-52 静载下含平面缺陷的概率安全评定流程图

5.3.1 L_r 和 K_r 的统计变量计算

如图 5.4-52 所示，首先选取确定性参数，并随机抽取随机变量的值，根据本章第2节的内容中确定性安全评定方法来计算 K_r，L_r 值，并重复 n 次，得到 L_{r1}，L_{r2}，…，L_{rn} 以及 K_{r1}，K_{r2}，…，K_{rn} 的值，然后对其进行统计分析：

L_r 的均值为
$$u_{L_r} = \frac{1}{n}\sum_{i=1}^{n}L_{ri} \qquad (5.4\text{-}221)$$

L_r 的变异系数为
$$v_{L_r} = \frac{\sqrt{\dfrac{1}{n-1}\sum_{i=1}^{n}(L_{ri}-u_{L_r})^2}}{u_{L_r}} \qquad (5.4\text{-}222)$$

K_r 的均值为
$$u_{K_r} = \frac{1}{n}\sum_{i=1}^{n}K_{ri} \qquad (5.4\text{-}223)$$

K_r 的变异系数为
$$v_{K_r} = \frac{\sqrt{\dfrac{1}{n-1}\sum_{i=1}^{n}(K_{ri}-u_{K_r})^2}}{u_{K_r}} \qquad (5.4\text{-}224)$$

L_r 和 K_r 的相关系数 ρ_{L_r,K_r} 为

$$\rho_{L_r,K_r} = \frac{\sum_{i=1}^{n}(K_{ri}-u_{K_r})(L_{ri}-u_{L_r})}{\sqrt{\sum_{i=1}^{n}(K_{ri}-u_{K_r})^2\sum_{i=1}^{n}(L_{ri}-u_{L_r})^2}} \qquad (5.4\text{-}225)$$

用柯尔莫哥洛夫检验法，即可确定 L_r 和 K_r 的分布形式。经检验发现它们基本上服从正态分布，这与众多因素的随机变量一般服从正态分布的数理统计经验是一致的。

5.3.2 失效率的计算

以 L_r 和 K_r 为基本变量，广义安全裕度 $p^*(L_r, K_r)$ 为

$$p^*(L_r,K_r)=\begin{cases}(1-0.14L_r^2)[0.3+0.7\exp(-0.65L_r^6)]-K_r & L_r<L_r^{\max}\\ 0 & L_r\geqslant L_r^{\max}\end{cases} \tag{5.4-226}$$

由于 L_r 和 K_r 服从正态分布，所以有 L_r 和 K_r 的联合分布概率密度函数 $f_{L_r,K_r}(L_r, K_r)$ 为

$$f_{L_r,K_r}(L_r,K_r)=\frac{\exp\left[-\dfrac{\left(\dfrac{L_r-u_{L_r}}{u_{L_r}v_{L_r}}\right)^2-2\rho_{L_r,K_r}\dfrac{L_r-u_{L_r}}{u_{L_r}v_{L_r}}\dfrac{K_r-u_{K_r}}{u_{K_r}v_{K_r}}+\left(\dfrac{K_r-u_{K_r}}{u_{K_r}v_{K_r}}\right)^2}{2(1-\rho_{L_r,K_r}^2)}\right]}{2\pi u_{L_r}v_{L_r}u_{K_r}v_{K_r}\sqrt{1-\rho_{L_r,K_r}^2}} \tag{5.4-227}$$

由于式（5.4-227）没有解析解，因此采用第二水平法计算失效率：

首先计算 $L_r<L_r^{\max}$ 的概率 q_1。由上可知，L_r 服从正态分布，因此有

$$q_1=\Phi\left(\frac{L_r^{\max}-u_{L_r}}{u_{L_r}v_{L_r}}\right) \tag{5.4-228}$$

式中，Φ 为标准正态分布的分布函数，它是增函数，所以有当 L_r 的均值或变异系数增加时，q_1 减小；

然后计算 $L_r<L_r^{\max}$ 时的广义失效率 q_2。

$$q_2=1-\Phi(\beta) \tag{5.4-229}$$

设 C_{L_r,K_r} 为 L_r 和 K_r 的协方差矩阵

$$C_{L_r,K_r}=\begin{bmatrix}(u_{L_r}v_{L_r})^2 & u_{L_r}v_{L_r}u_{K_r}v_{K_r}\rho_{L_r,K_r}\\ u_{L_r}v_{L_r}u_{K_r}v_{K_r}\rho_{L_r,K_r} & (u_{K_r}v_{K_r})^2\end{bmatrix}$$

C_{L_r,K_r}^{-1} 为其逆矩阵，可靠性指标 β 为

$$\beta=\min_{p^*(L_r,K_r)=(1-0.14L_r^2)[0.3+0.7\exp(-0.65L_r^6)-K_r]=0}\left\{\begin{bmatrix}L_r-u_{L_r}\\ K_r-u_{K_r}\end{bmatrix}C_{L_r,K_r}^{-1}\left[L_r-u_{L_r},\ K_r-u_{K_r}\right]\right\}^{1/2} \tag{5.4-230}$$

通常可以采用一阶二次矩迭代法求解上式，但由于其表达式较复杂，常用的迭代法都不收敛。由分析可知，式（5.4-230）是有约束的双变量数学规划问题。因此可以采用数学规划法进行求解，最后计算失效率 q_f。由全概率公式有，广义失效率 q_f^* 为

$$q_f^*=q_1q_2+(1-q_1) \tag{5.4-231}$$

5.4 疲劳载荷下概率安全评定方法

有关疲劳寿命与应力幅度之间的关系及 Miner 疲劳累计损伤理论和变幅载荷作用下疲劳寿命预测的问题在前面几章中已经讨论过，下面就基于 Paris 公式的疲劳寿命预测和断裂失效的概率评估方法加以介绍。

本节首先对疲劳裂纹的确定性模型以及影响疲劳裂纹扩展的因素加以分析，然后在此基础上对基于 Paris 公式的不确定性进行分析，进而对疲劳概率评定方法进行讨论。

（1）疲劳裂纹扩展确定性模型

如前面几章的介绍可知，疲劳裂纹扩展，常用一简单关系 Paris 公式表示

$$\frac{da}{dN}=C(\Delta K)^m \tag{5.4-232}$$

式中，$\Delta K=K_{\max}-K_{\min}$ 为应力强度因子幅度；C，m 为材料常数；a 为裂纹半长；N 为循环次数或循环时间。

为了描述 $K_{\max}$ 接近断裂韧度 K_C 时裂纹扩展明显加快的现象，Forman 等人提出：

$$\frac{da}{dN}=\frac{C(\Delta K)^m}{(1-R)K_C-\Delta K} \tag{5.4-233}$$

式中，$R=\sigma_{\min}/\sigma_{\max}$ 是应力比。

为了考虑门槛值特性的影响，Donalhue 等人提出了如下的关系式。

$$\frac{da}{dN}=C(\Delta K-\Delta K_{th})^m \tag{5.4-234}$$

式中，ΔK_{th} 为应力强度因子门槛值。

把在高 ΔK 及低 ΔK 值时对幂规律所发生的偏离特点进行综合，Priddle 等推导出关系式。

$$\frac{da}{dN}=C\left[\frac{\Delta K-\Delta K_{th}}{K_C-K_{\max}}\right]^m+C' \tag{5.4-235}$$

式中，应力强度因子门槛值并不是材料的常数，而是随应力比 R 变化；$\Delta K_{th}=A(1-R)^\gamma$，其中 C，C'，γ 为实验常数。

为了进一步表明裂纹扩展速率与应力比 R 的密切关系，Walker 等人提出下列关系式。

$$\frac{da}{dN}=C\left[\frac{\Delta K}{(1-R)^n}\right]^m \tag{5.4-236}$$

Erdogan 提出

$$\frac{da}{dN}=\frac{C(1+\beta)^n[\Delta K-(\Delta K)_{th}]^m}{K_C-(1+\beta)\Delta K} \tag{5.4-237}$$

式中 $\beta=\dfrac{K_{\max}+K_{\min}}{K_{\max}-K_{\min}}$。

注意在这个关系式中含五个常数：m，n，C，K_C 和 $(\Delta K)_{th}$。事实上，正如 Erdogan 指出的，尽管在验证疲劳幂规律关系式有效性方面作出了不少成果，这些关系式主要是属于经验性的，没有一个是从基本力学观点引出的。因此门槛值 $(\Delta K)_{th}$、应力比 R、断裂韧度 K_C 等的影响如何拟合，没有定论，在公式中如何体现，也没有解决。

必须指出 Paris 公式中只引入了主要参量 ΔK，影响疲劳裂纹扩展速率的其他因素，如应力比 R、环境因素（温度、湿度、介质等）以及加载频率都隐含在系数 C 和 m 之中，只适用于一定的 ΔK 范围，在 ΔK 很大或很小时，Paris 公式都不能正确表达 da/dN 的变化规律。所以，有时裂纹扩展速率方程分三段分别表示裂纹扩展初期、裂纹扩展中期、裂纹扩展后期的扩展规律。

一般认为影响裂纹扩展速率的因素有：

1）应力比 R（或平均应力 σ_m）；
2）加载频率（加载频率减小，da/dN 增大）；
3）腐蚀性介质（加载频率低时，作用更明显）；
4）材料厚度（裂纹尖端应力状态）；
5）载荷相互作用（载荷顺序）；
6）裂纹闭合；
7）材料组织结构均匀性、第二相粒子。

有效应力强度因子幅值为

$$\Delta K=K_{\max}-K_{op} \tag{5.4-238}$$

式中，K_{op} 对应裂纹张开应力 σ_{op}，此时有

$$\frac{da}{dN}=A(\Delta K_{eff})^m \tag{5.4-239}$$

式（5.4-239）的优点在于，张开应力由实验得到，把应力比 R 等诸多因素作用包含在材料常数 C，m 和张开应力 σ_{op} 中，用较少的实验测定常数描述疲劳裂纹扩展，工程应用方便。

为了修正 Paris 公式提出了立方裂纹扩展速率，提出以下公式。

$$\frac{da}{dN}=\exp\left[C_1+C_2\ln\Delta K+C_3\ (\ln\Delta K)^2+C_4\ (\ln\Delta K)^3\right]$$

或
$$\frac{da}{dN}=10^{[C_1+C_2\ln\Delta K+C_3(\ln\Delta K)^2+C_4(\ln\Delta K)^3]}$$

即 $\lg(da/dN)=C_1+C_2\ln\Delta K+C_3(\ln\Delta K)^2+C_4(\ln\Delta K)^3$

也就是认为 ΔK 是裂纹扩展驱动力，在 Paris 公式基础上考虑了 $\lg\Delta K$ 高阶项影响。通过逐步回归实现，从而提高了拟合能力。

(2) 疲劳裂纹扩展不确定性方程

前面讨论了疲劳裂纹扩展确定性关系。但是疲劳裂纹扩展时，由于几何结构、材料常数和外载的偏差，发生疲劳裂纹扩展的随机性，即使上述因素是确定的，由于疲劳裂纹尖端材料微观、细观组织结构的随机变化，偏离均匀连续状态，也会发生疲劳扩展的随机性。所以，确定性疲劳裂纹扩展速率方程可以认为是由实验得到的集平均意义上的确定性裂纹扩展速率方程。

1) 疲劳裂纹扩展的统计分析必要性　由于疲劳裂纹扩展的复杂性，涉及各种不确定性影响因素，目前描述裂纹扩展的确定性方程，基本上是实验总结，而不是由力学原理推导出来。

如前所述，Paris 和 Erdogan 提出了方程式（5.4-232）。对大多数工程常用的金属材料，$m=2\sim4$；Forman 等提出考虑平均应力和裂纹扩展后期 $K_{max}\rightarrow K_C$ 时，快速扩展的方程式（5.4-233），$m=3$。

这些公式仅仅是定性反映了应力比 R、最大应力强度因子 K_{max} 接近断裂韧度时裂纹快速扩展的影响，只是实验数据拟合。

Priddle 考虑门槛值影响，提出式（5.4-235），式中，$C'=2.4\times10^{-7}$mm，接近金属的微观晶粒尺寸，也是定性反映 ΔK_{th} 影响。

1989 年 Short 和 Hoeppner 由能量原理提出，裂纹扩展速率与塑性功增加速率 $W^{(P)}$ 相关。Dolinski 认为塑性功增加速率

$$W^{(P)}=C_2\Delta K^2+C_4HK^4$$

裂纹尖端平均塑性功与 ΔK^4 成比例，裂纹尖端断裂面上的塑性功与 ΔK^2 成比例，到裂纹快速扩展阶段，裂纹较大，ΔK^4 所起作用增大。这在某种程度上解释了 $m=2\sim4$ 之间的原因。

所以由实验数据拟合的实验参数具有一定的不确定性。

从力学角度考虑，由于连续性假设，表征裂纹尖端弹塑性应力、应变场的量为 ΔK（或 ΔK_{eff}），应力比 R 及板厚 B，用这三个有限的理想化参量描述裂纹尖端复杂状态自然使方程中参数产生一定的不确定性，并且连续性假设是对真实情况的理想假设。

从材料学角度看，起始小裂纹状态，中间稳定扩展状态，最后快速扩展状态，裂纹塑性区大小及参与的晶粒数目不同，沿晶、穿晶或混合断裂等断裂方式影响，沿板厚参与裂纹扩展晶粒数目等影响，都有一定的起伏，所以反映出裂纹扩展的不确定性。

从疲劳实验可观察到二个现象：①疲劳裂纹扩展具有一定的不确定性，并且 a-N 曲线缠结在一起，如图 5.4-53 所示；②表征一次应力循环的裂纹扩展量-疲劳条带宽度，可能时宽时窄，但总体趋势是逐渐变宽。所以我们描述疲劳裂纹扩展规律的方程要能够反映上述两个特点。

总之，从实验数据、力学和材料学方面综合分析，由于各种不确定性影响因素影响，实验数据不充分，还缺乏对基本规律完整的了解，只有采用统计分析方法才能较完整反映疲劳裂纹扩展的规律。

2) 疲劳裂纹扩展方程随机化方法　疲劳裂纹扩展方程随机化方法可分为随机变量方法（Random Variables）和随机过程方法（Random Process）。

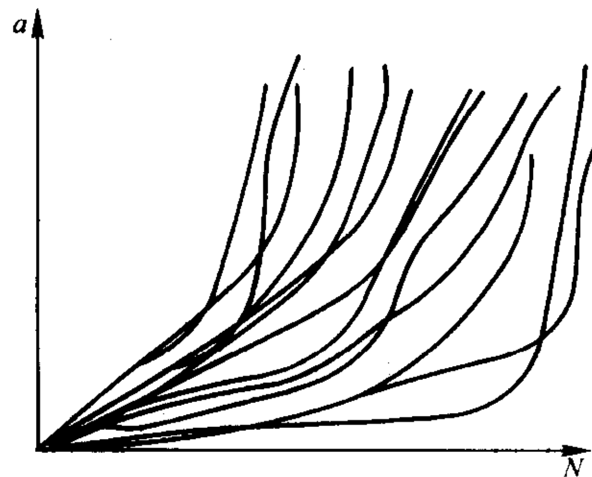


图 5.4-53　疲劳裂纹扩展时 a-N 曲线

① 随机变量方法　由式（5.4-232）知 Paris 公式为

$$\frac{da}{dN}=A\ (\Delta K)^m$$

a) 单随机变量法　m 为常数；C 看作是随机变量。把疲劳裂纹扩展随机性用随机变量 C 表征。

这种方法的优点是进行断裂失效分析时，方程求解简单。其缺点是 C 抽样值大时，裂纹扩展快；C 抽样值小时，裂纹扩展慢，不会产生实验观察到的 a-N 曲线缠结交叉现象。

b) 双随机变量法　把 C，m 看作是随机变量，此时方程求解比单随机变量法复杂，优点是 C，m 随机抽样时，能够反映 a-N 曲线缠结交叉现象，但是还不能反映疲劳条带时宽时窄现象。

② 随机过程方法　Paris 公式写成

$$\frac{da}{dN}=CX\ (N)\ (\Delta K)^m$$

式中，$X\ (N)$ 为均值为 1 的循环次数的随机序列（随机过程）。

由于随机过程 $X\ (N)$ 随循环次数时大时小，可以反映疲劳条带时大时小的规律，即可以反映给定试样疲劳条带变化的规律。但没有反映不同试样之间 C，m 值有时差异较大的特点。所以当某种材料不同试样之间 C，m 的线性回归差异不大时，可以采用此不确定性方程。

为了综合反映不同试样 C，m 值有时变化较大，以及疲劳条带时大时小的规律，可以采用以下 Paris 公式

$$\frac{da}{dN}=CX\ (N)\ (\Delta K)^m \tag{5.4-240}$$

式中，C、m 为相互独立随机变量；$X\ (N)$ 为与 C，m 相互独立的随机过程。用 C，m 表征不同试样之间裂纹扩展速率差异，而用 $X\ (N)$ 表征给定试样在 C，m 值确定情况下微小扰动的影响。

有时随机过程表示为裂纹半长的随机序列 $X\ (a)$，此时 Paris 公式可写作

$$X\ (a)\ \frac{da}{dN}=C\ (\Delta K)^m \tag{5.4-241}$$

用循环次数随机序列 $X\ (N)$，还是裂纹半长随机序列 $X\ (a)$，取决于实验结果。

式（5.4-240）是普遍表达式，当 $X\ (N)=1$ 时退化为双随机变量方法。进一步当 m 给定时，退化为单一随机变量方法。

3) 疲劳裂纹扩展不确定性方程的实验测定　本节讨论疲劳裂纹扩展不确定性方程中随机变量 C，m 以及随机过程 $X\ (N)$［或 $X\ (a)$］的实验确定方法。

① 单随机变量裂纹扩展方程的实验测定　设试样数目为 m 个，第 i 个试样（$i=1, 2, \cdots, m$）的采样点总数目

为 n_i，j 表示第 j 个采样点（$j=1, 2, \cdots, n_i$）。n_1，n_2，$\cdots$，n_m 分别表示第一个试样，第二个试样，$\cdots$，第 m 个试样总采样点数目。

由 $(\mathrm{d}a/\mathrm{d}N)_{ij}$ 和 Δk_{ij}（$i=1, 2, \cdots, m$；$j=1, 2, \cdots, n_i$）可求随机变量 C

$$C=\frac{(\mathrm{d}a/\mathrm{d}N)_{ij}}{(\Delta K_{ij})^n}$$

采用优化法，选取 m，使得 μ_c 趋近直线和 σ_c/μ_c 取极小值，求得单变量法的参数 m，以及随机变量 C 的均值 μ_c 和标准差 σ_c，如图 5.4-54 所示。

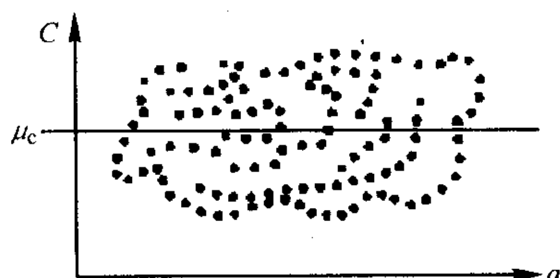


图 5.4-54　双随机变量

② 双随机变量法裂纹扩展方程的实验测定　设 C、m 为随机变量，对第 i 个试样有：

$$\left(\ln\frac{\mathrm{d}a}{\mathrm{d}N}\right)_{ij}=\ln C_i+m_i\ln(\Delta K)_{ij}\ (j=1, 2, \cdots, n_i)$$

最小二乘法线性回归得 C_i，m_i。由每个 C_i，m_i（$i=1, 2, \cdots, n$）确定随机变量 C，m 的均值和标准差，以及所服从的分布类型。

③ 随机过程法裂纹扩展方程的实验测定　设疲劳裂纹扩展方程为

$$\frac{\mathrm{d}a}{\mathrm{d}N}=CX(N)(\Delta K)^m$$

设对每个试件已经求得了随机变量 C，m（或 C_i，m_i（$i=1, 2, \cdots, m$），对第 i 个试样有

$$[\ln X(N)]_{ij}=\left(\ln\frac{\mathrm{d}a}{\mathrm{d}N}\right)_{ij}-\ln C_i+m_i\ln(\Delta K)_{ij}(j=1,2,\cdots,n_i)$$

由 $[\ln X(N)]_{ij}$ 可估计随机过程第 i 个试样 $[\ln X(N)]_{ij}$ 的自协方差函数的参数，综合考虑每个试样 $i=1, 2, \cdots, m$ 的自协方差函数，取平均作为 $X(N)$ 的自协方差的参数。

当 Paris 方程表示为

$$X(a)\frac{\mathrm{d}a}{\mathrm{d}N}=C(\Delta K)^m=\ (j=1, 2, \cdots n_i)$$

此时：$[\ln X(a)]_{ij}=\ln C_i+m_i\ln(\Delta K)_{ij}-\left(\ln\frac{\mathrm{d}a}{\mathrm{d}N}\right)_{ij}(j=1,2,\cdots,n_i)$

同样，可以估计 $X(a)$ 的自协方差参数。

(3) 基于 Paris 公式的含平面缺陷的疲劳概率安全评定

含平面缺陷的疲劳概率安全评定流程图如图 5.4-55 所示。具体的计算评定方法详见下文。

1) 统计变量的计算　如图 5.4-55 所示，首先选取确定性参数，并随机抽取随机变量的值，根据前面确定性安全评定方法计算裂纹的最终尺寸。设裂纹的初始长半轴为 c_0，短半轴为 a_0，埋藏裂纹与最近表面的距离的初始值为 p_{10}。将裂纹的扩展过程分为 N（N 为循环次数）个阶段。在第 $i+1$段，可用第 i 段的裂纹尺寸计算裂纹尖端的应力强度因子的变化幅度 $\Delta K_{c_{i+1}}$ 和 $\Delta K_{a_{i+1}}$，设 ΔK_{th} 为裂纹扩展的门槛值，则若有

$$\Delta K_{c_{i+1}}>\Delta K_{th} \qquad (5.4\text{-}242)$$

则有

$$c_{i+1}=c_i+C_1(\Delta K_{i+1})^m \qquad (5.4\text{-}243)$$

若

$$\Delta K_{a_{i+1}}>\Delta K_{th} \qquad (5.4\text{-}244)$$

对表面裂纹有

$$a_{i+1}=a_i+C_1(\Delta K_{i+1})^m \qquad (5.4\text{-}245)$$

对埋藏裂纹有

$$\begin{cases}a_{i+1}=a_i+C_1(\Delta K_{i+1})^m\\ p_{1i+1}=p_{1i}+C_1(\Delta K_{i+1})^m\end{cases} \qquad (5.4\text{-}246)$$

且有若 $p_{1i+1}>0.8a_{i+1}$则埋藏裂纹需重新表征为表面裂纹，并在以后按表面裂纹计算。

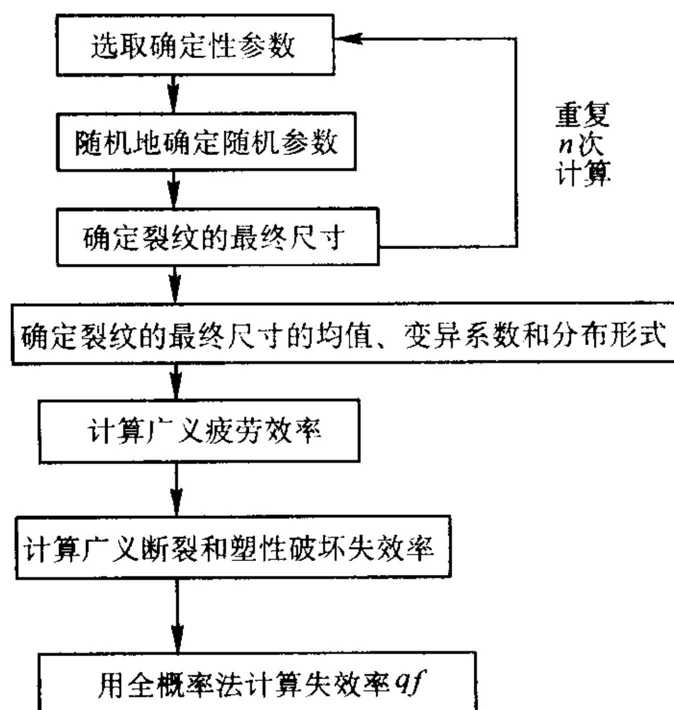


图 5.4-55　平面缺陷的疲劳概率安全评定流程图

根据上节给出的随机变量方法（Random Variables）和随机过程方法（Random Process）并按图 5.4-55 给出的步骤计算出裂纹最终长半轴 c_f 和最终短半轴 a_f 和埋藏裂纹与最近表面的距离 p_{1f}。并重复 n（抽样次数）次，得到 c_{f_1}，$c_{f_2}\cdots c_{f_i}\cdots c_{f_n}$ 和 $a_{f_1}\cdots a_{f_i}\cdots a_{f_n}$，（及 $p_{1f_1}p_{1f_2}\cdots p_{1f_i}\cdots p_{1f_n}$），然后进行统计分析，求出 c_f 的均值 u_{c_f}，变异系数 v_{c_f}，a_f 的均值 u_{a_f}，变异系数 v_{a_f} 及 p_{1f}的均值 u_{p1f}，变异系数 v_{p1f}，并用柯尔莫哥洛夫检验法确定 c_f、a_f 及 p_{1f}的分布形式。通常 c_f、a_f 及 p_{1f}服从正态分布或对数正态分布。

由此可以看出，与确定性安全评定不同的是，概率安全评定不是按确定性评定那样由 c_f、a_f 及 p_f（或 $c_{破}$、$a_{破}$ 及 $p_{破}$）来比较判断此时是否会发生静载（或疲劳）下的断裂，而这里需要由 c_f、a_f 及 p_{1f}的统计量进行概率评定。

2) 广义断裂和塑性破坏失效率的计算　由前面的计算可知，L_r 和 K_r 基本上服从正态分布，然后按全概率式(5.4-231)，计算疲劳载荷作用下的广义失效率 q_f^*。

以上为平面缺陷的疲劳概率安全评定的方法和步骤。

6　工业风险评估及其主要方法和应用

6.1　概述

近年来，风险评估的概念越来越为更多的生产者和经营者所接受，安全管理正逐步向风险管理的方向过渡。能够造成人身伤害和财产损失的一切事件可看作是危险，它代表着产生潜在损失的特征。从绝对意义上，在人类的生存环境中，危险是无处不在和无处不有的。但是，危险的发生是有条件的。有的危险发生十分迅速，如油气产品燃烧或爆炸；而有的危险发生地十分缓慢，如腐蚀造成的油气管道穿孔破裂。因此，“事物具有危险性”是可以被普遍接受的，而“危险的事物必定会造成损失”则是不正确的。

风险可以定义为危险转变为现实的概率和损失程度的综合。通常风险 R 可以定义为概率 P（Probability）与后果 C（Consequence）的乘积。即

$$R = PC \tag{5.4-247}$$

如果一个事件由几个独立的事件组合起来，则

$$R(x) = \sum_{i=1}^{n} P(x_i)C(x_i) \tag{5.4-248}$$

如果该事件是连续作用的，则其发生概率随 x 的变化而变化。

$$R(x) = \int_0^{\infty} P(x)C(x)\mathrm{d}x \tag{5.4-249}$$

风险是描述某一事件过程的量，可随时发生变化。它和危险不同。危险是固有的、静止不变的，而风险是可以通过人们的主客观努力在事件过程中降低的。

燃气管道风险评估与经济成本的关系如图 5.4-56 所示。

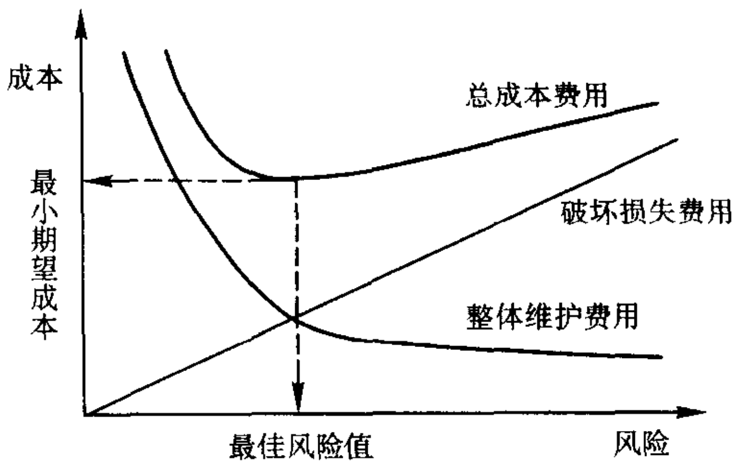


图 5.4-56　经济成本与风险高低的关系

对某项工程活动进行风险评估，可以帮助辨识引起风险的危险源，度量、量化和评价风险以及风险后果，从而将生产活动带来的风险控制在人们可以接受的范围内，因此可以更经济、有效地进行经营管理，合理分配资源，以便于更有效地进行预防、检测，提高安全性，减少灾难性事故的发生。

根据墨菲定律，要降低一个系统的风险就必须向该系统不断地注入能量，亦即风险的降低总是要以成本的增加为代价的。从理论上讲，风险与成本的关系可以用图 5.4-56 所示的曲线来描述。从该曲线可见，增加投资可以明显降低风险，直至到达某一风险值。之后，在大大提高成本的条件下，才有进一步的风险值减低，其利润回报是很小的。由此可见，对系统要依靠增加投入来消除风险不仅是得不偿失的，而且是大可不必的。

如果把图 5.4-57 的曲线划分为 3 个区域，则在第一区内实际成本的增加会明显降低风险值，在第二个区域是一个转变带，在第三区内要想获得任何风险的降低要付出很大的开支。所以，从经营者的角度出发，在第二区进行运行是最合理的，因为在第一区内降低风险值的措施可能还不够，而在第三区的运行显然不是明智的经济决策。

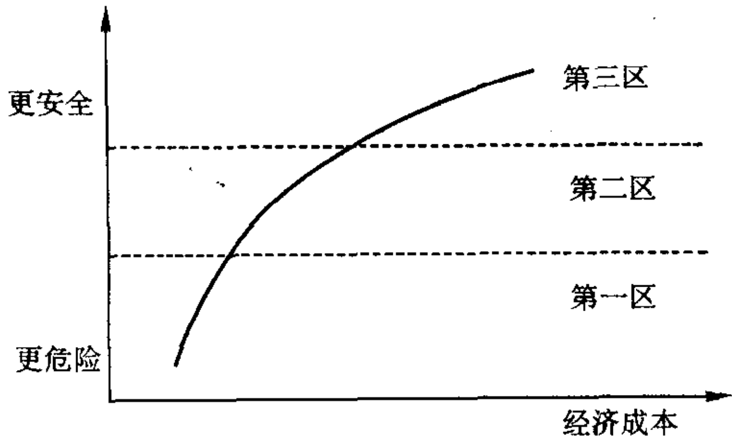


图 5.4-57　理论成本优化关系

如果把系统的安全维护成本曲线与其事故成本曲线放在一个坐标系中进行考虑，我们可很方便地得到一条系统总成本与风险的关系曲线（如图 5.4-56）。利用该曲线可以通过数学方法在最少期望成本前提下的一个最佳风险控制值。这就是最佳风险管理的思想。

风险工程学包括风险设计、风险评价、风险预测和风险管理。后者主要指控制危险及对危险采取相应措施的决策。风险工程学是一门新兴学科，具有跨学科的特点。它包括的内容很广泛，举凡可靠性工程学、失效分析、失效预测和预防、结构完整性评价和工业经济预测与决策等尽在其中。不仅如此，由于工业领域危险源各式各样，危险事故发生机理千差万别，防范措施也因不同对象而异，经济投资和决策方式也不尽一致，风险工程学的研究内容和方法也随不同工业类别或工艺过程、装置的不同而迥异。但是，作为一门学科，除了有针对性研究各个工业领域风险的个性问题外，在共性问题或方法上，近年来成为研究的热点。

6.2　风险评估的主要任务

对工业装置进行风险评估，主要应解决三个方面的问题。识别危险源、分析危险概率和后果是风险的三要素：①危险一般由于能量或毒物释放失去控制而引起。在进行风险分析时，首先要确定危险源种类，如毒物释放、爆炸、火灾等，其次要确定系统中哪一部分是危险的来源，如压力容器、压力管道、储罐、动力装置等；②危险源发生破坏的概率，即危险转化为现实事件的概率，以及环境、人员或其他生态系统、建筑物或构筑物暴露于危险区域的程度；③危险一旦发生，对暴露目标的有害作用或可能造成的损失。在进行风险分析时，要对这三个要素中的每个环节作具体分析和评价。

6.2.1　风险识别与失效概率的确定

正确识别系统的风险因素，找出系统中存在的可能发生危险破坏的薄弱环节（危险源）是进行有效的风险评估的第一步。引起工业生产装置失效破坏的因素是多方面的。以燃气管道为例，燃气管道及其附属设施（包括阀门，调压器等）是一个十分复杂的系统，可能包含了复杂的管道自身结构情况和地面环境情况。美国的 PRCI（Pipeline Research Committee International）针对美国和欧洲的输气管道事故数据进行了分析和分类，归纳总结出 22 种引起管道失效的“基本因素”，其中只有一种因素的本质原因是“未知的”，即不能确定它的本质特性。其余的 21 种失效若按照它们的本质和发展特性分为 9 种，若按照与时间的关系分为 3 类，如表 5.4-20 所示。

6.2.2　失效概率的确定

如果对一个由多个部件（部件之间可能有复杂的串、并联关系）构成的系统进行风险评估，必须首先明确各个部件的失效概率，并按照各个部件之间的逻辑关系，计算该系统的失效概率。燃气管道的风险的定义应是各管段各自的失效概率和失效后果的乘积的总和，用公式表示就是

$$Risk_{\mathrm{SYS}} = \sum Risk_i = \sum P_i C_i$$

在表 5.4-20 所示的各种失效因素中，它们引起管道失效的概率是不一样的。表 5.4-21 列出了近年来世界各国燃气管道的一些失效因素的统计数据。可以看出，在不同地区受不同环境条件（包括地理环境、运行环境、管理环境等）的影响，在不同地区同一失效因素引起燃气管道失效的概率是不一样的。以腐蚀和外部干扰为例，加拿大和俄罗斯由于地广人稀，野外的输气管道由于没有人员维护而经常受到腐蚀破坏，而欧美国家则因为人口相对稠密，外部的施工和人为挖掘比较频繁，管道受到外力干扰而破坏的情况比较多。

表 5.4-20 PRCI 针对输气管道总结出的失效“基本因素”

发生变化且变化规律与时间关系密切的因素	外腐蚀	
	内腐蚀	
	应力腐蚀	
稳定存在的失效因素	与生产制造有关的缺陷	焊缝缺陷
		管材缺陷
	与焊接、装配有关的缺陷	管道周向焊接有缺陷
		管道焊接装配有缺陷
		起皱或翘曲
		螺栓掉落/管道破裂/接头失效
	与设备有关的失效	衬垫或密封圈失效
		控制/排放设备故障
		密封/捆扎失效
		装备混杂
发生变化但变化规律与时间无关的失效因素	第三方/机械破坏	由第一方，第二方或第三方引起的破坏（瞬间/延时失效）
		管道先前曾遭受破坏（延时失效模式）
		蓄意破坏
	误操作	错误的操作程序
	与自然和外力有关的失效因素	严寒
		闪电
		暴雨或洪水
		地层运动

表 5.4-21 部分国家和地区输气管道事故原因统计数据 %

地区 \ 失效因素	腐蚀	外部干扰	施工或材料缺陷	地基位移	误操作	其他
欧洲	13.91	52.17	19.13	5.22	4.35	5.22
加拿大	42	0	23	19	0	16
美国	23.4	39.7	14.4	0	0	22.5
前苏联	39.9	16.9	35.0	0	2.9	5.3

注：数据的统计年份分别为：欧洲：1970~1992年；加拿大：1985~1995年；美国：1990~2001年；前苏联：1981~1990年。

历史数据中的失效率并不能直接用来当作失效概率来使用。失效概率计算方法主要包括统计法和解析法。统计法是通过收集相应的管道历史失效数据进行分析，得出该危险过去的发生频率，并以此预测现在或将来的发生频率。对于某些失效原因由于还没有建立相应的失效模式，不能利用解析法计算其失效概率，因此这类情况只能利用统计法进行分析。当历史数据不能获取或不充分时，则采用故障树法、Monte-Carlo法、概率断裂机制方法、结构可靠性方法等解析方法来计算失效概率。国外在研究管道的失效风险过程中，建立了相应的管道失效数据库。

由管道失效数据库可推得各种管道的平均失效概率。由于历史失效数据只能确定所有管道的平均失效概率，在20世纪80年代中期，英国 British Gas R&D Division 研究出一种计算方法，将管道的断裂机理和操作破坏数据相结合计算管道的失效概率。该方法的优点是管道的断裂机理考虑了管道的几何形状、材料性能和操作条件，从而预测单条管道的失效概率。一些文献根据英国健康与安全委员会（HSE）推荐的数学模型，基于R6双判据法的通用失效评定曲线，把材料的力学性能、缺陷尺寸分布及外加载荷当作随机变量，采用一阶可靠性方法/二阶可靠性方法（FORM/SORM）计算了陆地气体输送管道由于受到外部冲击引起的轻微泄漏或断裂的失效概率。

6.2.3 失效后果的判断

估计事故造成的后果是管道风险评估中重要的环节。一般可从人身安全、财产损失、环境影响和信誉损失几方面来考虑。其中财产损失是最容易估计和有明确标准的，包括计算损坏的建筑物、车辆和其他财产，管输介质的损失，环境损失以及进行管道修复的费用等。人身伤亡损失和环境污染损失应与各国社会的价值观取向关系密切。包括按国家有关工业事故伤亡赔偿标准支付的抚恤金、救治伤员的开支、支付政府环保罚金和恢复环境生态所需要的各项费用。信誉损失是最难估计的损失。它关系到众多的社会因素和政治因素。必须通过专家和业主对包括社会的、政治的、经济的多方面因素进行综合分析才能评定。

一般来说燃气管道释放的可燃气体迅速扩散到大气中，对环境的影响很小，主要是着火或爆炸对人体和财产造成很

大影响。最终结果以死亡率和总的经济损失表示。英国健康与安全委员会（HSE）研制出估算来自管道危险的个体、社会风险估算模型，用于计算高压输气管道失效时火球的持续时间、喷射火焰长度等，以及由此引起的热辐射强度和危险剂量等。在某文献中，作者根据数学模型计算了管道失效时的着火、爆炸等对人体带来的影响，见表5.4-22。

表5.4-22 API581关于财产损失面积和人员伤亡面积的计算

环境低于燃点		
泄漏类型	财产损失面积/m^2	人员伤亡面积/m^2
连续性泄漏	$A=93x^{0.98}$，x为泄漏速度，kg/s	$A=235x^{0.96}$
即时性泄漏	$A=70x^{0.67}$，x为总泄漏量，kg/s	$A=134x^{0.67}$

6.3 风险评估的主要方法

风险评价方法有定性评价方法、定量评价方法和介于两者之间的半定量方法。定性评价方法相对简单，容易使用，但主观性较强，难以说服民众和相关管理部门相信最终的评估结果。而定量风险评价技术则利用概率断裂力学、有限元方法、可靠性与维修技术和各种强度理论，对系统的风险进行定量评价和决策。它与定性风险分析技术的不同之处在于必须在大量设计、施工和运行资料的基础上，建立完善的数据库管理系统，并掌握裂纹缺陷的扩展规律和管材的腐蚀速率，由此运用确定性或不确定性方法来建立评估的数学模型，然后进行分析求解。其结果精确性取决于原始数据的完整性、数学模型的精确性和评估方法的合理性。定量风险评估包含了风险识别、风险分析、风险计算和风险评估。目前在工程中实际应用的风险评估方法有以下几种。

6.3.1 失效模式、后果与严重度分析

失效模式和后果分析（failure modes and effects analysis，FMEA）在风险分析中占重要位置，是一种非常有用的方法，主要用于预防失效。但在试验、测试和使用中又是一种有效的诊断工具。欧洲联合体ISO 9004质量标准中，将它作为保证产品设计和制造质量的有效工具。它如果与失效后果严重程度分析联合起来（failure modes，effects and criticality analysis，FMECA），应用范围更广泛。

FMEA是一种归纳法。对于一个系统内部每个部件的每一种可能的失效模式或不正常运行模式都要进行详细分析，并推断它对于整个系统的影响、可能产生的后果以及如何才能避免或减少损失。

进行FMEA工作所涉及的主要问题如下。

(1) 失效

针对系统的具体情况，以设计文件或相关标准、规范为依据，从功能、工况条件、工作时间、结构等确定本系统失效的定义，并确定表征失效的主要参数。Henley和Kumamoto对于过程装置如塔器、压力容器、压力管道和储器等，提出构造FMEA需要考察、校核的项目。

变量：流量、温度、压力、浓度、pH值、饱和度等；

功能：加热、冷却、供电、供水、供空气、供N_2、控制等；

状态：维修、开车、停车、更换催化剂等；

异常：很不正常、略有一些不正常、无不正常、位移、振荡、未混合、沉淀、着火、腐蚀、断裂、泄漏、爆炸、磨损、液体溢出、超压等。

仪表：灵敏度、安放位置、响应时间等。

(2) 失效模式

考虑系统中各部件可能存在的隐患，依据具体内容确定失效模式。如：

1) 功能不符合技术条件要求；

2) 应力分析中发现的可能失效模式；

3) 动力学分析、结构分析或机构分析中发现可能失效的模式；

4) 试验中发生的失效，检验中发现的偏差；

5) 完整性评价、安全性分析确定的失效模式。

(3) 失效机理

根据所确定的失效模式，进行失效机理分析，并确定失效或危险发生的主要控制因素。

(4) 失效后果

在进行失效后果分析时，应考虑任务目标，维修要求以及人员和设备的安全性等。要考虑原始失效（一次失效）可能造成的从属失效（二次失效）；要考虑局部失效可能造成的整体失效，要考虑对全系统工作、功能、状态产生的总后果。

在进行失效模式和后果分析时，应按照上述内容编制FMEA表格，逐项填写。有些场合，也需要进行半定量分析。设定失效发生频率程度、失效后果严重程度、失效原因被检出程度3个指标，根据经验或与所考察对象相似系统的失效记录，用1到10数字标定。各指标标定值的乘积称为风险乘数（risk product number）。风险乘数的大小表示不同失效模式的相对重要度。

表5.4-23为某文献给出的FMEA半定量分析各项指标参考值。

表5.4-23 FMEA半定量分析各项指标参考值

等级	失效发生频率程度	失效后果严重程度	失效原因被检出程度
不可能检出	—	—	10
微小	1	1	8~9
小（少）	2~3	2~3	6~7
中等	4~6	4~6	4~5
高	7~9	7~9	2~3
很高	10	10	1

6.3.2 失效树分析法

失效树分析法是一种非常重要的安全分析方法。它的有关基本概念和性质等请参见第1章的相关内容，这里不再赘述。这里简单介绍该方法在风险评估中的应用情况。

(1) 失效树的定性分析

如前面章节所述，失效树中的逻辑函数主要有与门和或门两种。失效树寻找失效原因的过程就是通过逻辑函数的运算找到造成顶事件的最小割集。所谓最小割集是指一个包含了若干基本事件的集合，这个集合内全部基本事件的发生是造成顶事件发生的充分必要条件，每一个最小割集都是使顶事件发生的一种的组合方式。

失效树分析就是要找出系统（燃气管道）失效树的最小割集，得到顶事件的全部失效模式，以发现系统中的最薄弱环节，集中力量对关键部位进行强化。

按照逻辑关系从顶事件逐级向下分析确定最小割集的基本原则是与门使割集容量增大，或门使割集数量增加。利用布尔代数的等幂律（$A\cdot A=A$）去掉重复事件，得到全部的最小割集。

(2) 基本事件发生概率的确定

在得到失效树的最小割集后，如果能得到各基本事件的发生概率，便可按照计算规则对顶事件的发生概率进行定量的计算。如何确定基本事件的发生概率是得到正确计算结果的必要过程。

1）后验概率的计算　可通过直接统计得到基本事件的发生概率。有些事件虽无先例（或先例很少），但与有些多发性事件类似，可以通过类比分析、相似归纳与经验修订的方法，估算出事故发生概率。这些概率数值，建立在大量资料归纳、统计基础上，虽有相当大的可靠性，但也有很大的局限性。由于资料是在一定历史条件（相对一定的科学技术及管理水平）下取得的，不同的使用下差异很大。如果加入主观订正则必然引起数据可靠性的下降。

2）先验概率的计算　先验概率的计算建立在各种事件概率分布形式已知（一般为正态分布）的基础上，先计算得出各基本事件发生概率，用演绎推理的方法从各基本事件发生概率估算出发逐步向顶上事件推算，最后得出危险源事故发生概率。

应当指出，如果需要计算的事件发生概率具有一定的模糊性，应引入隶属度等模糊数学的概念，采取模糊数学的逻辑计算方法进行计算，才能准确地描述事件，得到精确的计算结果。

(3) 失效树的定量分析

失效树中的逻辑运算函数有如下的定义。

设 $x_i(t)$ 为基本事件 i 在时刻 t 所处的状态，并有定义如下的状态函数：

$$x_i(t)=\begin{cases}0 & \text{事件 } i \text{ 在 } t \text{ 时刻不发生}\\1 & \text{事件 } i \text{ 在 } t \text{ 时刻发生}\end{cases}\qquad(5.4\text{-}250)$$

基本事件 i 在时刻 t 发生的概率等于 $x_i(t)$ 的期望值，因而有

$$P_i(t)=P\{x_i(t)=1\}\qquad(5.4\text{-}251)$$

顶事件的状态必然是由底事件状态组成的 n 维向量 $X(t)=\{x_1(t), x_2(t), x_3(t), \cdots, x_n(t)\}$ 的函数，设 Y 为描述顶事件的随机变量，则顶事件在时刻 t 发生的概率 PY 为：

$$PY=P\{Y[x_1(t), x_2(t), x_3(t), \cdots, x_n(t)]=1\}\qquad(5.4\text{-}252)$$

结构函数与门、或门的运算规则可分别由下式表示：

$$\gamma^{\text{and}}=\prod_{i=1}^{n}x_i\qquad(5.4\text{-}253)$$

$$\gamma^{\text{or}}=1-\prod_{i=1}^{n}(1-x_i)\qquad(5.4\text{-}254)$$

将基本事件的发生概率代入便可进行顶事件失效概率的计算。

6.3.3 基于可信性的风险评估

可信性（Dependability）是一个非定量的集合性术语。根据 1994 年 ISO-9000 的定义，可信性是可靠性、维修性、保障性和测试性内容的综合。

可靠性是系统或装置在规定条件下和规定时间内完成规定功能的能力。

维修性是在规定条件下和规定时间内，按照规定程序和方法对系统或装置进行维修时，保持或恢复系统或装置达到规定状态的能力。

保障性是系统或装置的设计特性和计划的保障资源能满足使用要求的任务。或者，系统或装置的安装、运行及维修，需要一定的保障，如资金、技术、维修方针、程序、工具、仪器设备、相关规范、文件、技巧熟练的技术人员等。这些都是支持保障性不可或缺的条件。

测试性是系统或装置能及时并准确地确定其状态（运行正常、故障或性能下降）的特性。利用监控、检测等手段可以确定系统或装置内部的危险源以及性能蜕化的影响，要求具备测试功能和测试精度。

基于可信性的危险可能有两类：

1）系统或装置研制生产过程，预计进度计划受到干扰，预计的资金被突破或没有到位，可信性达不到预期水平。

2）系统或装置运行时发生失效，其后果可能导致人身伤亡，建筑物破坏，环境污染，造成经济损失。

基于可信性的风险分析就是按照这两类危险，按照可能发生的概率大小和发生后造成的后果来度量。

风险度（r）取决于危险发生的概率 P_f 及危险发生后果严重程度 C_f，它们之间的关系按事件和表示：

$$r=P_f+C_f-P_fC_f\qquad(5.4\text{-}255)$$

现代系统、装置中，除了硬件本身外，一般都使用了大量软件，许多功能由软件执行，进行风险分析时，对软件可能的风险也必须计及。

系统、装置完善程序不同、结构和操作复杂程度不同，引起失效和可能性也不相同。前者（硬件完善程度）的失效概率用 P_{hp} 表示；后者（硬件复杂程度）的失效概率用 P_{hc} 表示。软件执行功能的完善程度不同、复杂程度不同，引起失效的可能性也不相同。前者（软件完善程度）的概率用 P_{sp} 表示；后者（复杂程度）失效概率用 P_{sc} 表示。故有

$$P_f=aP_{hp}+bP_{sp}+cP_{hc}+dP_{sc}\qquad(5.4\text{-}256)$$

式中，a、b、c、d 是加权系数，它们的和为 1。

在设计、研制过程中，由于技术水平或其他技术因素，使系统、装置技术性能降低，因而可能对失效后果严重程度造成不同影响（用 C_t 表示）；由于资金超过预算而使系统、装置研制计划调整，因而可能对失效后果严重程度造成不同影响（用 C_c 表示）；工程进度与原计划不符，为了调整进度可能影响整个工作程序、甚至工程质量，因而造成对失效后果严重程度的不同影响（用 C_p 表示）。故有

$$C_f=eC_t+fC_c+gC_p\qquad(5.4\text{-}257)$$

式中，e、f、g 是加权系数，它们的总和为 1。

某文献中给出了美国国防系统管理学院《系统工程管理》的 P_{hp}，P_{sp}，P_{hc}，P_{sc}，C_t，C_c 和 C_p 的取值（对我国的产品，这些取值仅供参考）和计算示例。

6.3.4 基于模糊数学的风险评估

风险评估的关键问题之一就是在众多的影响系统安全性的各类因素与系统安全性之间建立量化的计算关系。建立风险评估体系后，应逐步实现系统失效率与风险值的定量化计算，对各失效因素之间的关系要定量描述并建立数学模型。数学量化的深入不仅可以揭示这些失效因素发展的内在规律，还可以考察因素间相互影响的性质。

然而，影响工业装置（系统）失效率的众多因素来自于物的因素、人的因素和综合作用与二者的系统管理因素，是一个十分复杂的客观体系。许多影响因素具有不确定性，或无法进行量化。随着指标量化的发展，研究者发现经典数学方法在描述这种体系的时候显得能力不足。为了使该体系数量模型化，往往需要假定一些约束条件，只对重要的因素进行分析。这使得评估体系的假定性过强，因此，即使建立了公式化计算的大型评估体系，其结果并不能完全符合实际的客观规律，这直接影响了系统风险评估的发展。

风险评估的许多方面具有模糊性。如工业装置（系统）失效的可能性，失效引起的后果的严重性，都依据不同的评估而有不同的变化。引入模糊数学的概念，进行燃气管道的半定量风险评估，并充分承认失效因素量化和相互之间关系的模糊性，是进行装置（系统）风险的合理评估的有效方法。模糊数学的方法特别适合于高非线性系统或系统的数学模型不准确甚至很难得到的情况。

使用模糊数学的方法进行风险评估的方法如下：

1）确定因素集 $U=\{u_1, u_2, \cdots, u_n\}$。此处 u_i（$i=1, 2, \cdots, n$）表示对工业装置（系统）的风险有影响的第

i 个因素；

2）选择评价集 $V=\{v_1, v_2, \cdots, v_m\}$，此处 v_j（$j=1, 2, \cdots, m$）表示评价的第 j 的等级；

3）单因素评判。定出每个因素对于各评判等级的隶属度。可采取专家评判法、德尔菲法等。例如，可采取如下方法：成立一个由 l 人组成的评判组，每位组员给每一个 u_i（$i=1, 2, \cdots, n$）给出 v_j（$j=1, 2, \cdots, m$）中的一个且仅一个等级。若 l 位组员中对 u_i 的评定为等级 v_j 的有 l_{ij} 个人，则 $\sum_{i=1}^{m} l_{ij}=l$，并且对 u_i 的评判结果为一模糊子集

$$R_i=\left(\frac{l_{ij}}{l}, \cdots, \frac{l_{im}}{l}\right)=(r_{il}, \cdots, r_{im})=\frac{r_{i1}}{v_1}+\cdots+\frac{r_{im}}{v_m} \tag{5.4-258}$$

由此可得到模糊关系矩阵

$$R=(R_1, R_2, \cdots, R_n)^{\mathrm{T}}=(r_{ij})_{n\times m}=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix} \tag{5.4-259}$$

必须明确：R 中第 i 行 R_i 反映的是被评判的对象的第 i 个因素对于评价集中各等级的隶属度；第 j 列反映的是被评对象的各因素分别去评价集中第 j 个等级的程度。

4）确定各因素的权重 $A=(a_1, a_2, \cdots, a_n)$；此处 a_i（$i=1, 2, ..., n$）是由评判组的因素 u_i（$i=1, 2, \cdots, n$）对此事物的影响相对于其他各因素的重要程度，一般 $0\leqslant a_i\leqslant 1$ 且 $\sum_{i=1}^{k} a_i=1$，对同一事物而言，不同的评判观点可能不同，因而，不同评判组确定的权重 A，可能是不同的。

5）综合评判。做矩阵乘法

$$B=AR=(a_1, a_2, \cdots, a_n)\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix} \tag{5.4-260}$$

得到模糊评判向量 B（如果关系矩阵有多层结构则需进行多级评判），然后通过最大隶属度法、加权平均法、模糊分布法等，得到评价结果，对装置（系统）的风险等级做出评价。如果装置（系统）的结构复杂，则需通过二级或多级综合评判才能得到结果。具体做法可参阅关于模糊评判的有关专著。

6.4　风险评估的主要应用

6.4.1　化工装置的定性风险评估

定性风险评价可以对燃气管道系统的各管段快速地进行风险排序，虽然比较粗略，但可以为进一步的定量风险评估提供基础。图5.4-58是定性风险评估的一种表示方法，评估结果可放入5×5的矩阵中，是按高、较高、中等及低风险来分级的，垂直轴代表失效可能性的大小，水平轴代表失效后果的严重程度。

将各装置的风险评价结果列入风险矩阵中，就会得到风险水平的指示。风险矩阵的评估结果可以用于确定有潜在危险最大的区域，确定需要加强检测的管段，以及相应减少风险的方法，同时可用于是否要进行全面的定量评价的决策。

对化工、石化工艺过程装置进行风险评价的方法很多，但大多数都是以化学物质对火灾、爆炸、毒物泄漏的敏感性为基准，适当考虑装置的操作方式、工况条件和化学反应的危险性。根据火灾、爆炸的可能性，以及推断的事故发生可能造成的损害程度，确定对付风险的对策。

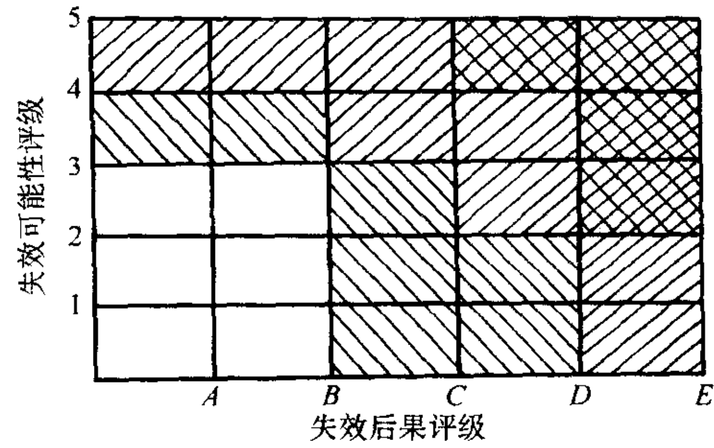


图5.4-58　定性风险矩阵

高风险　中高风险　中等风险　低风险

日本高压气体保安协会提出的设备重要度分类法，有别于上述方法。它将装置进行分类，按照危险度的不同，给出重要度系数。不仅做到对装置风险的评价，而且据此确定适宜的检查周期，使装置维修管理科学化，加强了计划性。

这种方法将危险性很高的装置与危险性很低的装置，用重要度系数加以区分。将危险度分为四级，按危险性从高到低给以重要度系数为5、3、2、0.5。

在判定装置危险度时，除了考虑介质性质、工况条件（温度、压力）等外，还要考虑装置操作的历史记录和使用环境。由于装置在生产过程中发挥的作用不一，有的是关键装置，有的是辅助装置。为了区分，对它们给以1、2、3、4的乘数。其中乘数数值最大的是对生产过程影响最大的装置。

将重要度系数与乘数相乘，它们的乘积就是风险系数，用以确定装置等级。根据装置等级制订检查、维修方案和防范事故的措施。若干化工装置的等级列于表5.4-24。

表5.4-24　装置等级

级别	（重要度系数×乘数）累积和	典型装置
Ⅰ级	35以上	例如高温高压临氢的加氢反应器及其周围的装置；在苛刻工况条件下运行的装置
Ⅱ级	20～25	一般的高压装置
Ⅲ级	20以下	一般的工艺装置

6.4.2　燃气管道的定量风险评估

定量风险评估是风险评估发展的最高阶段，是一门新型管理技术，只有将风险评估建立在定量分析的水平上，管道的风险评估才真正具有系统性、精确性和预见性，其显著的社会经济效益已在国外一些国家得到证实。燃气管道的风险评价结果可最大限度地减少管道发生突发性事故所造成的综合经济损失，提高管道的安全性。同时，管道风险分析过程中得到的一些有用结论和最终结果，还可为管道的安全设计、选择合适的运行线路提供依据。如何建立实用有效的定量风险评估的方法是研究者们的共同课题。在燃气管道的定量风险评估中，也存在许多关键的技术问题。

（1）确定管道的失效模式

在初步进行风险识别，找到构成管道风险的主要失效因素后，需要确定管道失效的主要模式。按照管道事故的严重程度，管道失效模式可分为三种，即轻微泄漏、管道穿孔和管道断裂。具体划分标准与管道本身特征（如直径、壁厚）有关。根据美国和欧洲20世纪70～80年代统计资料，在所有输气管道事故中，泄漏占40%～80%，穿孔占10%～40%，断裂占1%～5%。管道缺陷的评定技术可为管道的失

效模式提供很好的数学模型以供计算。而管道缺陷的评定对于不同的管道材料有着不同的研究。对于韧性很好的奥氏体钢管道的母材，潜在的失效模式为塑性极限载荷失效，因此，可直接以塑性极限载荷作为管道的极限承载能力进行缺陷评定。对于奥氏体钢管道的焊缝，当含有裂纹类缺陷时，一般也是塑性极限载荷控制，偶然也有弹塑性断裂失效的情况发生。

对于铁素体钢管道，无论裂纹缺陷位于母材还是焊缝，脆性断裂失效、弹塑性断裂失效及塑性极限载荷失效三种模式均有可能发生。为此，美国标准规范ASMEIWB-3650给出了判别三种潜在失效模式的筛选准则，相应的判定方法和具体的判别依据如图5.4-59和图5.4-60所示。三种失效模式的判别依据为 S_c，S_c 也是图5.4-60所示的失效评定点与坐标连线的斜率（K_r/L_r），其判别值为0.2和1.8。当 $S_c>1.8$ 时，管道的失效模式将是塑性极限载荷失效。

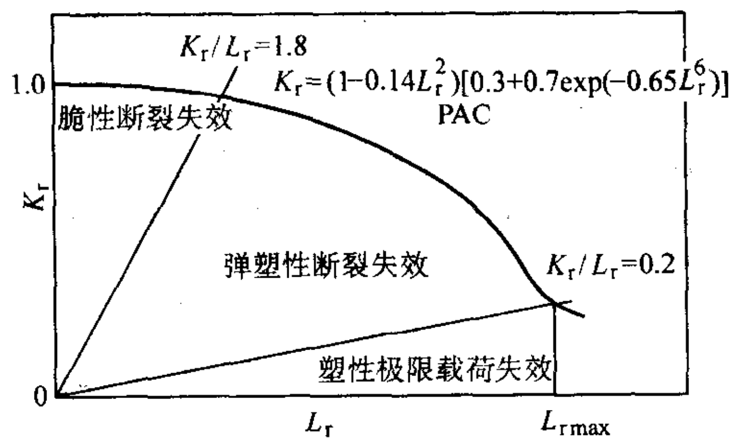


图5.4-59 含缺陷压力管道的三种失效模式

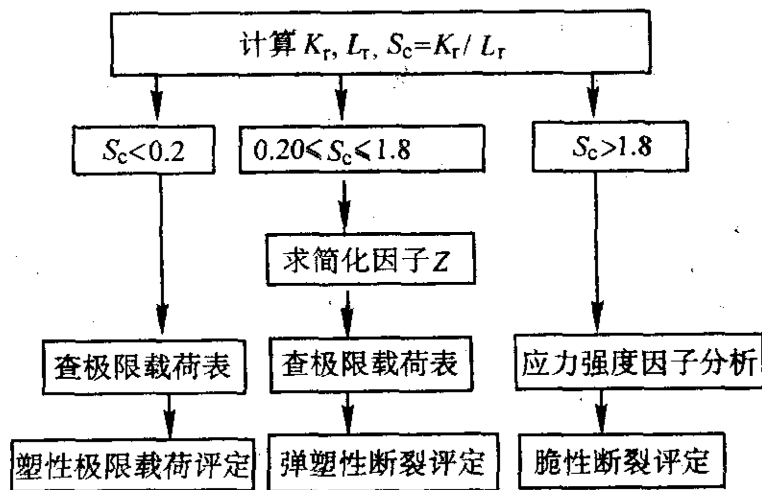


图5.4-60 ASME规范判别含缺陷铁素体管道潜在失效模式的筛选原则

失效模式的确定不仅对于失效概率的计算有着重要的参考价值，而且还直接影响到失效后果的计算。不同的泄漏孔直径直接导致了管输介质泄漏量的多少，从而决定了泄漏后的各种表现形式（安全泄放、蒸气云爆炸、闪火、喷射火焰、火球等）的影响面积。API581中对于失效后果的计算即是基于以上的问题进行考虑的，并且对不同的管输介质和管道运行环境都提供了经验公式以供计算失效后果。

(2) 失效概率的计算方法

管道失效概率计算方法包括统计法和解析法。统计法是通过收集相应的管道历史失效数据进行分析，得出该危险过去的发生频率，并以此预测现在或将来的发生频率。对于某些失效原因（如地表运动等）由于还没有建立相应的失效模式，不能利用解析法计算其失效概率，因此这类情况只能利用统计法进行分析。当历史数据不能获取或不充分时，则采用故障树法、Monte-Carlo法、概率断裂机制方法、结构可靠性方法等解析方法来计算失效概率。国外在研究管道的失效风险过程中，建立了相应的管道失效数据库。

由管道失效数据库可推得各种管道的平均失效概率。由于历史失效数据只能确定所有管道的平均失效概率，在20世纪80年代中期，英国British Gas R&D Division研制出一种计算方法，将管道的断裂机理和操作破坏数据相结合计算管道的失效概率。该方法的优点是管道的断裂机理考虑了管道的几何形状、材料性能和操作条件，从而预测单条管道的失效概率。根据英国健康与安全委员会（HSE）推荐的数学模型，基于R6双判据法的通用失效评定曲线，可把材料的机械性能、缺陷尺寸分布及外加载荷当作随机变量，采用一阶可靠性方法/二阶可靠性方法（FORM/SORM）计算了陆地气体输送管道由于受到外部冲击引起的轻微泄漏或断裂的失效概率。

(3) 失效后果计算

管道失效后果计算主要考虑管道失效所造成的对人身生命安全的影响以及对环境和财产的影响。其中，对人身安全的影响可分为长期影响和短期影响、死亡和伤害以及不同的伤害类型（烧伤程度、呼吸问题等）。对环境的影响主要考虑失效事故对空气的污染以及对生态系统的影响（如对失效地点附近植物群和动物群的破坏等）。对财产的影响主要考虑公司的各种经济损失，包括因产品损失、管道修补、财产破坏等造成的不同的经济损失。

为了准确计算失效后果，风险分析者就必须根据特定管道的失效形式采用特定的释放模型、扩散模型、着火模型、爆炸模型等模拟有关现象，并要考虑危险发生时的天气情况以及失效点附近的地质形貌，计算人体所受到的热辐射强度、有毒物质危险剂量以及压力的影响等。计算原理示意图见图5.4-61。

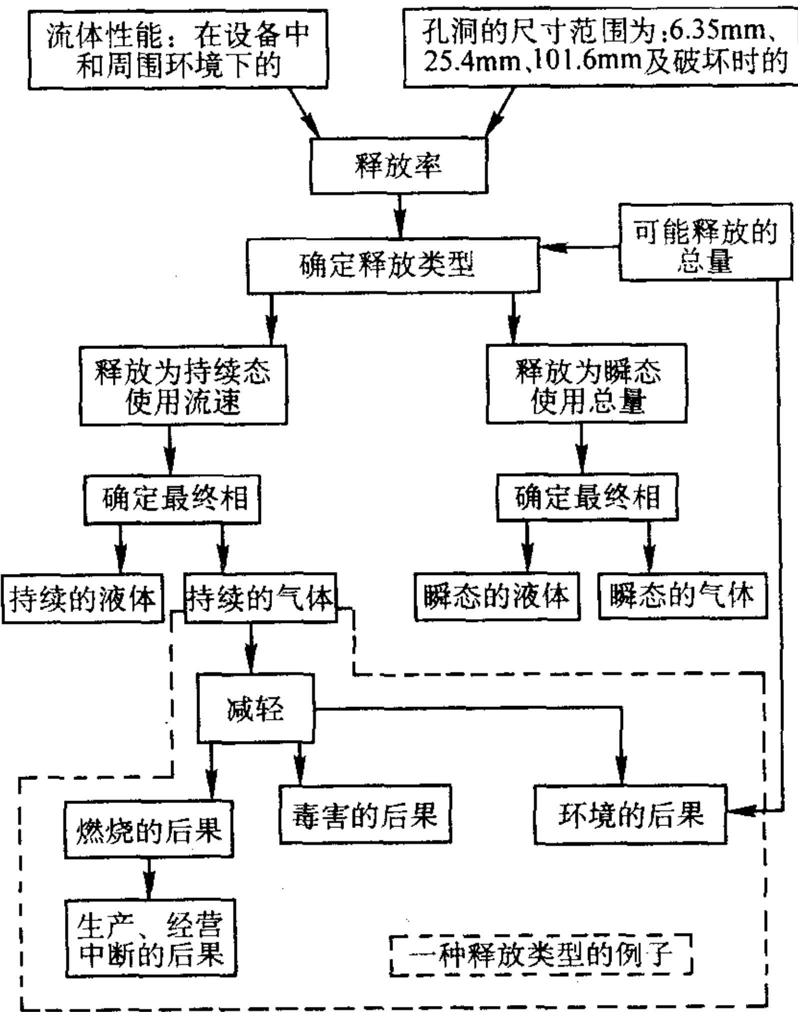


图5.4-61 API581中关于管道失效后果的计算原理示意图

一般来说燃气管道释放的可燃气体迅速扩散到大气中，对环境的影响很小，主要是着火或爆炸对人体和财产造成很大影响。最终结果以死亡率和总的经济损失表示。英国健康与安全委员会（HSE）研制出估算来自管道危险的个体、社会风险估算模型，用于计算高压输气管道失效时火球的持续时间、喷射火焰长度等，以及由此引起的热辐射强度和危险剂量等。在一些文献中，作者根据数学模型计算了管道失效时的着火、爆炸等对人体带来的影响。

(4) 风险计算及评价燃气管道的失效风险值

定量风险评估的结果用管道的失效概率与失效后果的乘积表示，考虑周期一般为一年。燃气管道定量风险评价的基

本原则是必须考虑管道失效时所有危险的总概率，既不能漏掉某一危险的失效概率，又不能重复计算，即分别考虑在各种失效模式情况下，不同的天气和风向条件下，计算各种危险的失效概率，然后将各种概率相加，得到危险总失效概率。定量风险分析重复性大，一般要借助于计算机程序进行计算。

风险的表现形式有个体风险和社会风险两种。个体风险表示的是当危险发生时所引起的位于管道附近某一点的个体死亡率，该风险依赖于危险发生点的地理位置，与人是否存在于该点无关，因此个体风险给出的是管道周围的不同地点的风险分布。图5.4-62为某一管道失效时的个体风险图，随着与管道失效点的距离 d 的增加，个体风险降低。社会风险表示的是在某一给定地区，当危险发生时所造成的死亡人数（N）和累积失效概率（F）之间的关系，以 F-N 形式表示。社会风险示意图见图5.4-63。社会风险与危险发生地点无关，与管道周围的人口密度有关。因此，如果当危险发生时附近没有人群存在，则社会风险为零，而个体风险可能很高。

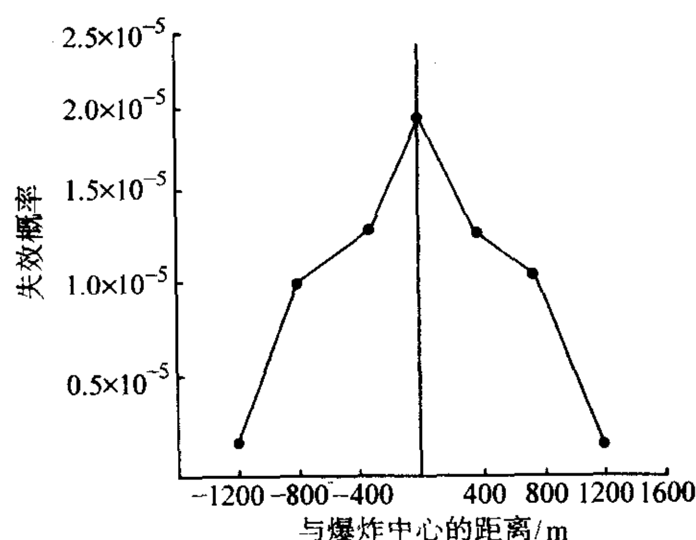


图5.4-62　个体风险示意图

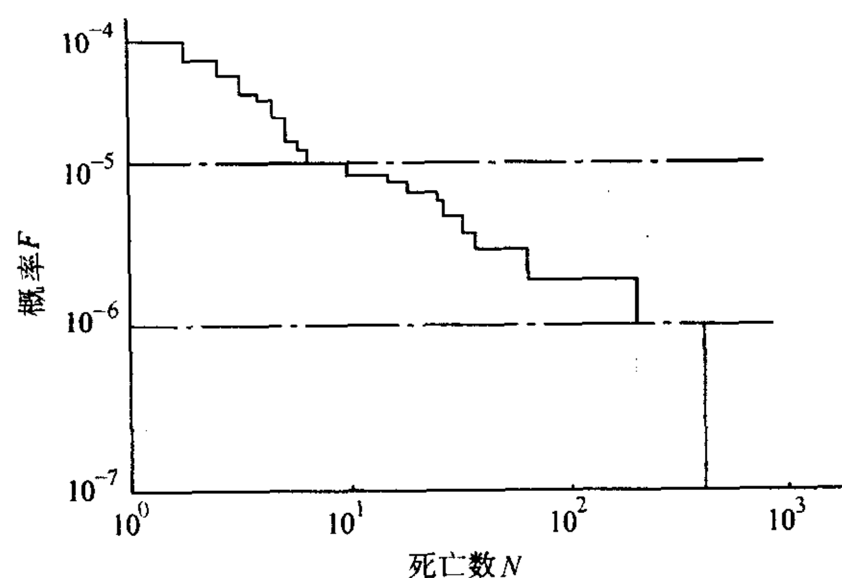


图5.4-63　社会风险示意图

当已知管道各段的风险后，就可以得到沿整条管道的失效风险分布情况。关于管道的定量风险评价，目前还没有国际统一标准。英国和荷兰的管道管理机关建立了相应的标准。英国提出了“最低合理使用原则（ALARP）”。当风险水平超过上限值（大于 10^{-4}/a）时，该风险不可接受；当风险值低于下限（小于 10^{-6}/a）时，不需要采取任何安全改进措施；当风险值处于上下限之间时，需要采取预防措施或缓和措施将风险降低到可接受的水平。

6.5　风险评估近年来的主要进展

6.5.1　失效概率的专家评估法

对于失效概率比较低（例如燃气管道的失效率一般数量级都在 10^{-6}左右）、事故案例较少的工业装置（系统），通过经典的统计学很难使用这些较少的数据建立风险评估模型。为了解决缺乏失效数据这一难题，可以请专家对所研究的管道进行合理有效的评估，记录专家的合理意见并对专家的知识进行提取，建立风险评估模型并对其不确定性进行合理的表示。专家评估的方法不是随意的意见调查，是需要科学有效的方法，如德尔菲法等。德尔菲法的基本步骤为：①选择咨询专家，一般以30~50人为宜，意见咨询以书信的形式进行，避免相互间的意见干扰；②设计调查表，将调查表及资料一并寄给专家，进行问卷调查；③回收调查表并进行统计处理，将处理结果作为下轮调查的背景资料和调查表的设计依据。若认为调查结果满意则进行下一步，否则转向第二步；④整理最终的调查报告，给出说明性的意见。

一个切合实际的专家评估的方法必须解决两个难题，一是如何将专家的意见准确地体现，使对燃气管道失效概率可能性的评估切实代表的专家的意见；二是如何将不同专家的不同意见合理地组织在一起。有研究者通过贝叶斯统计法解决这一问题，将专家的意见看作是贝叶斯统计的问题，变量的不确定性可用对变量的观察和专家估计基础上的概率分布来进行描述。下式为通用的失效概率计算公式。

$$p(x,\theta \mid y)=\frac{p(y \mid x,\theta)p(x,\theta)}{\int_{x,\theta}p(y \mid x,\theta)p(x,\theta)\mathrm{d}x\mathrm{d}\theta} \qquad (5.4\text{-}261)$$

式中，p 为失效概率；y 为专家意见；θ 为关系参数。

可对该式进行进一步假设，如进行层次分析法，以利于求出最终的解析解。

6.5.2　对与时间相关的失效因素的研究

随着使用时间的延长，腐蚀一类的因素必然会引起工业装置特别是化工装置的穿孔或破裂的事故的发生，所以在各类风险评估中腐蚀因素是作为重点分析的。由于影响腐蚀进程的因素十分复杂，许多研究者都对由于腐蚀引起的装置可靠性与剩余寿命进行了研究。其中，对于腐蚀速率的研究是评定腐蚀因素诱发管道事故可能性的关键环节。目前解决这一课题的办法有现场检测技术、室内模拟实验分析技术和现场挂片对比分析技术。

近年来，研究者还对随着时间的推移而随机发生的失效因素对系统寿命的影响做出了一些研究，将各种失效因素的概率分布密度函数进行整合，并建立了加速寿命模型等。

$$\lambda_m(t)=\int_0^{\infty}\lambda(t,\theta)\pi(\theta \mid t)\mathrm{d}\theta \qquad (5.4\text{-}262)$$

式中，t 为时间；θ 为参数；$\lambda(t,\theta)$为系统寿命为分布函数；$\lambda_m(t)$ 为系统多重作用下的失效速率；$\pi(\theta \mid t)$为条件概率密度函数。

6.5.3　对人的可靠性的研究

对于一个工业生产装置（系统），其失效概率的分析不应仅局限于管道的硬件和环境，还应包括对其运营者——人的可靠性的研究。人在工程系统的可靠性中起着关键作用，因为装置（系统）都是通过人这个环节相互连接，在燃气管道系统的设计、制造、管理和应用等各个阶段，人都起到了关键的作用。对人的可靠性的研究分为对与时间相关的人为差错、人的行为特征和人的模型等，以及对人员可靠性的定量分析。如人员辨认能力可靠性的模型：

$$P(t)=\exp-\left\{\frac{t/T_{0.5}-C_{\gamma i}}{C_{\gamma i}}\right\}^{\beta_i} \qquad (5.4\text{-}263)$$

式中，t 为完成某种事故响应操作的时间；$T_{0.5}$为完成这项操作所需时间的估计中值；$C_{\gamma i}$、β_i 为各类（i 指类别，即技能，规则，知识）人员素质的修正系数；$P(t)$ 为在给定的反映时间窗口 t 条件下，不能响应的概率。

6.5.4 GIS在风险控制中的应用

地理信息系统（GIS）是集计算机科学、信息科学、现代地理学、遥感测绘学、环境科学、城市科学、管理科学和现代通讯技术于一体的一门新兴边缘学科。经过近40年的发展，GIS不仅可以像传统的数据库一样管理数字和文字信息，而且还可以管理空间（图形）信息。GIS已经成为一种成熟的空间数据处理技术和方法，在风险评估和风险控制中也已经开始逐步应用。

GIS在风险评估和风险控制中的应用主要有这样几个方面：①事故信息搜索：通过GIS系统能够立即准确地得到事故发生的地点以及周围的地理环境、人口分布等。在事故发生之后，管理者可以方便地找到所有必要的联系方式和信息资源以控制事故，包括工程负责人员、公安、消防等部门，周围的人员可迅速撤离危险区域；②事故情形分析：针对特定装置的事故，GIS结合一些事故模拟分析系统可以进行事故发展模式预测（包括烟羽扩散模型和爆炸伤害模型等），并据此确定人员撤离的区域和路线；③辅助风险决策：风险的控制一般要考虑多方面的因素，涉及到各方面的专家、领导的意见和建议。为了加快事故处理，可以开发并嵌入辅助的专家系统和决策支持系统，辅助风险分析和决策；④辅助风险管理：利用GIS作为用户界面，把分析模型、管道的数据库信息、人口和建筑物的分布信息以及专家系统等有机结合起来，采用多媒体技术进行紧急事故处理（如图5.4-64所示），代表着未来风险管理的方向。

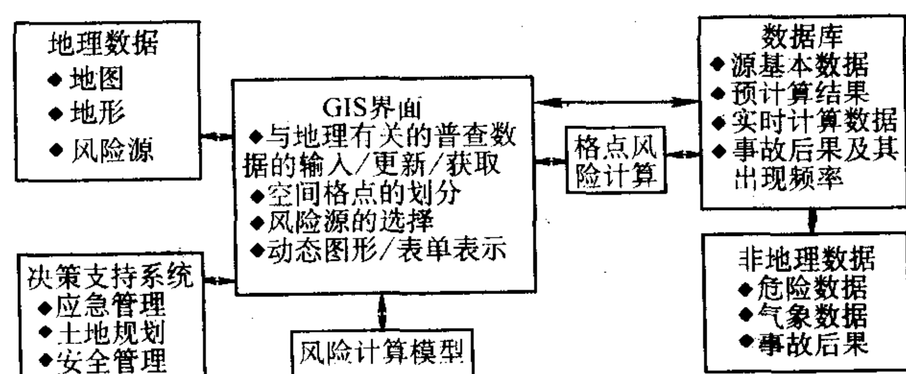


图 5.4-64 基于GIS的风险评估与风险评价

6.5.5 后果评估的研究

在风险评估中需要对人员伤亡带来的经济损失做出评价，这就需要对所谓“人的价值”做出评价。衡量“人的价值”在感情上往往很难接受，这经常被称为是不人道的，因为人的生命通常是被当作无价的。但事实上，评价人的价值是为了在有限的经济投入下优先对重大的危险源进行风险控制，从而更多地保护人身不受到伤害。在工程实践中，对人身安全的投资却不是无限的，因为经常需要衡量用于获得收益的投资和用于保证人身安全的投资。否则每一个十字路口都应该有红绿灯，每辆汽车都应该配备最先进的保护装置。此外不同地区对人自身价值的认同是不一样的，如发达地区的人们愿意花费更多的钱用于购买人身保险就是一个例子，因此需要认识到并定量地描述不同地区的这一差别。实际的工程决策中需要在保证人身安全的前提下，切实地计算人的价值，这比“零伤亡”、“零风险”之类的口号更有意义。

目前在工程中，通常将为避免一个人的伤亡（保护一个人不发生生命危险）而进行的资金投入定义为一个人的价值，计算公式为投入的资金总数除以可能发生伤亡的人数。主要应通过不确定性评价法（contingent valuation）进行。专家调研和意向调查的方式进行。也可定义为年人均国民生产总值乘以人均寿命。

当已知管道各段的风险后，就可以得到沿整条管道的失效风险分布情况。关于管道的定量风险评价，目前还没有国际统一标准。英国和荷兰的管道管理机关建立了相应的标准。英国提出了“最低合理使用原则（ALARP）”。当风险水平超过上限值（大于10^{-4}/a）时，该风险不可接受；当风险值低于下限（小于10^{-6}/a）时，不需要采取任何安全改进措施；当风险值处于上下限之间时，需要采取预防措施或缓和措施将风险降低到可接受的水平。

7 安全评定技术的其他发展趋势

7.1 安全评定技术的发展趋势

总体看来，安全评定技术经过几十年的发展，目前在以下几个方向发展。

1）扩大安全评定的对象。安全评定对象从最初的核压力容器，到压力管道，进而扩展到起重机、大型钻井平台、大型游乐机等大型机械装备机械（钢）结构等，评定对象不断增加和扩展。

2）扩大安全评定范围。从原来仅考虑内压载荷断裂疲劳来评定安全与否，深入到考虑环境介质的作用以及材料性能的退化和蠕变等情况的影响，使安全评定、剩余寿命和延寿技术相结合、形成更为实用的体系。

3）深入研究起裂控制的失效模式和断裂诊断技术，完善多模式、多级的断裂安全评定体系。包括深入研究与裂纹尖端三维约束程度有关的多（双）参量弹塑性断裂、非匹配焊缝断裂分析工程方法和裂纹的非J积分控制扩展的静载荷断裂评定等；深入研究各种情况下的体积缺陷的安全评定技术和方法体系，形成更为完整的体系缺陷的安全评定技术和方法体系；研究管系、容器的相互作用载荷及大应变疲劳扩展，高温蠕变、蠕变疲劳、应力腐蚀、腐蚀疲劳等情况下的裂纹扩展速率，完善特殊环境下的安全评定技术和方法等。

4）安全评定技术的工程化。随着安全评定工作在各个领域的开展，工程化、实用化和简单化的需求越来越强烈。这就需要研究人员尽量将最后的研究成果表达为简单易行的表格或简单公式，以满足工程需要。

5）从确定性的评定方法，向概率评定、模糊评定和智能评定的方向发展。前面已经介绍了有关概率评定的一些基本情况，这里不再赘述，在下文中将简单介绍有关模糊评定和智能评定的一些有关情况。

6）形成安全评定的在线监控体系。目前安全评定技术同新的传感技术、检测技术（特别是故障和缺陷的在线诊断技术）等同计算机技术相结合，逐步形成崭新的安全评定和检测控制技术体系。

7.2 模糊安全评定技术

工程中普遍存在大量的不确定性问题，主要可以分为两大类，一是随机不确定性，另一个是模糊不确定性。前者，如本章第5节所述，需要通过概率统计的方法和可靠性理论来进行分析和求解；而后者，则需要采用模糊集理论进行分析处理。

在设备的安全评定过程中，存在大量的模糊因素，如设备设计、制造、运行和维护期间的各种数据，测试过程中的各种数据，评定过程中的分析判断等都存在类似“可能”、“将要”、“大概”、“好像”等人的感觉、表达、思维、推理等不确定性及模糊性，它是事物的差异之间存在着中介过渡过程或亦此亦彼状态。事物在中介过渡过程中对一定质的隶属程度的不断变化正是模糊概念的基础。

在安全评定过程中如不同测试人员的结果，断裂韧度的选用等都存在模糊问题。另外，专家在作评定时，也习惯用

模糊语言变量来进行推理和判断。安全评定过程是非常复杂且难以用精确的数学变量来表示的过程。同时，人们不可能对每一个因素都进行精确的测试和考察，因此在一个压缩了的低维空间，利用模糊的概念来研究安全评定问题却能够得到更加准确的评定结果。

在安全评定的研究中采用模糊变换原理和最大隶属度原则，考虑与被评价事物相关的各种因素，所做出的综合评价即为模糊综合评定。

模糊安全评定技术的研究是随着计算机和人工神经网络技术的发展，近几年刚刚发展起来的，该项技术的发展必将大大提高安全评定的可靠性和水平。

7.3　智能安全评定技术

随着计算机技术的迅猛发展和计算机的广泛普及，人工智能（artificial intelligent）和专家系统（expert system）技术也日渐成熟，并且渗透到各个领域，如医学等领域已经建立了非常成功的专家系统。

专家系统的基本结构请见图 5.4-65。

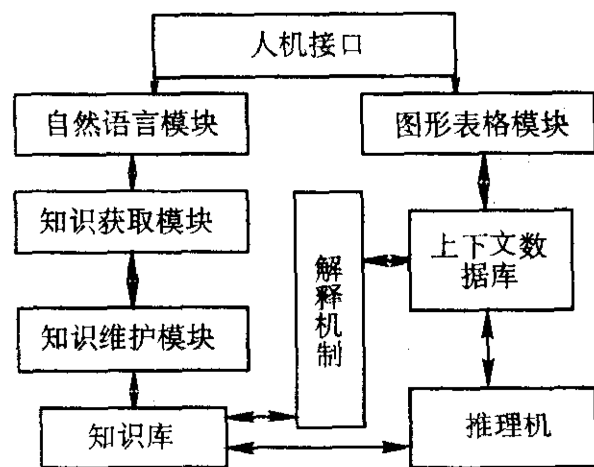


图 5.4-65　专家系统的基本结构

传统的专家系统一般是通过知识工程师将专家的知识和经验以某种形式化的方式整理出来以便获取知识，从而建立专家系统的知识库。但将专家的知识和经验转化为知识库当中的知识是非常困难的事情，主要原因是专家的经验知识是经过长期的积累而获得的，并不一定在专家的头脑中形成组织结构和逻辑性非常好的系统，而是带有很大的不确定性和经验性，同时若让专家来确定各种推理模型中使用的可信度、概率模糊隶属度等不确定性因子，也同样是非常困难的。

因此目前在专家系统中通常要引入人工神经网络技术（artificial neural network）。人工神经网络技术是受生物大脑的启发而构成的一类信息处理技术。该技术模拟人在思维过程中通过神经元之间同时相互作用的动态信息处理过程来完成。它是一个非线性动力学系统，其特色在于信息的分布式存储和并行协同处理。虽然单个神经元的结构极其简单，且功能也有限，但由大量神经元组成的网络系统却能实现非常丰富的功能，而且神经网络系统具有基体运算的能力和自适应的学习能力。另外，它还有非常强的容错性和鲁棒性，能够进行联想、综合和推广。

目前国内的一些研究者已建立了基于我国 CVDA-84 规范的外壳断裂疲劳的“断裂力学专家系统”，初步的“管道安全评定模糊专家系统”，国外 Jovanovic AS 等建立了基于“先漏后爆”（LBB）准则的压力容器完整性评定的专家系统，Battele Columbus Laboratory 编制了采用 J 积分的 5 种评定方法来计算含裂纹管道的起裂载荷和最大承载能力的专家系统 NRCPIPE，美国的 BMW（Boiler Maintenance Workstation）是用于锅炉管道的失效分析的专家系统，专家系统 ESR 可用于进行高温构件（管道、管系部件等）的剩余寿命预测等。这些工作都对安全评定向智能化发展有了推进的作用。当然，随着安全评定基础理论的进一步发展，应用范围的不断拓展，将会出现更加实用的适用于不同场合的安全评定专家系统。同时各种安全评定专家系统的出现也将推进安全评定技术的工程化和实用化。

编写：骆红云（北京航空航天大学）

第 5 章　失效预防技术和方法

1　失效的工程预防技术和方法

1.1　抗断裂失效设计技术

1.1.1　应力集中的影响及消除方法

（1）应力集中的影响

等截面杆受轴向拉压时，其截面上应力分布是均匀的，但实际的构件常因结构需要而制成阶梯形杆，或在杆上有油孔、切槽等。实验指出，在截面突变处（如孔、圆角、沟以及截面过渡处等），有应力骤然增大的现象。这种现象，称为应力集中。应力集中的程度以应力集中系数来说明，当静态应力不超过弹性极限时，缺口应力集中程度可用理论（或弹性）应力集中系数 K_t 来表示，即

$$K_t = \frac{\sigma_{max}}{\sigma_n} \tag{5.5-1}$$

式中，σ_{max}为应力集中区的最大应力；σ_n 为名义应力。

机械零件和材料性质急剧改变的地方，如零件的台阶、沟槽缺口、台肩圆角、开孔、转角、螺纹、键槽，材料表面缺陷——夹砂、斑疤、划痕、折叠、氧化、脱碳、粗晶环，材料内部缺陷——缩孔、气泡、疏松、偏折、夹杂物、白点、过热、过烧、发纹等处附近，都会出现应力集中。而且缺陷或沟槽、转角等的形状愈尖锐、材料的强度愈高、塑性愈低，应力集中现象或程度愈严重。当这种应力集中大于材料的抗拉强度时，机械零件就首先在应力集中部位或附近发生破断失效。

按照位错观点，对于延性或半脆性材料来说，形变和破断两个过程的共同存在和发展的必要条件是使金属发生屈服。当金属发生屈服以后，大量的位错在其运动过程中，由于受障碍物作用而堆积起来，就造成了应力集中。这个集中的应力若被形变过程所松弛，则破断过程被抑制，变形得以继续进行而不破断；反之，若以裂纹的发生与发展过程来松弛，则材料即会发生破断。

机械零件破断失效部位，往往发生在零件几何形状突然变化引起应力集中的部位。裂纹扩展的方向也是由零件应力状态和材料强度决定的。且裂纹扩展总是按照最小阻力路线前进的。在金属脆性破断、疲劳破断和应力腐蚀破断情况下，其裂纹的扩展方向一般都垂直于主拉伸应力的方向；而当韧性金属承受扭转载荷或金属在平面应力的情况下，其裂纹的扩展方向一般平行于切应力的方向。

除了机械材料强度和零件几何形状影响裂纹的生核位置外，零件的应力状态也对裂纹的起源位置产生影响。假如在机械材料质量合格、零件形状设计合理的情况下，裂纹将在应力最大处成核。例如，在单向弯曲疲劳的情况下，疲劳裂纹一般起源于受力一边的应力最大处，在双向弯曲疲劳的情况下，疲劳裂纹一般起源于受力两边的应力最大处；在齿面上的磨损裂纹，一般起源于齿轴节圆附近（该处的受力和相对运动速度最大，磨损也最严重）。

（2）降低机械零件应力集中系数的方法

为了保证应力集中部位的强度，单纯增加零件截面尺寸，是不合适的；主要应设法改善其局部状况，使应力集中系数降低。主要方法是采取局部强化和改善机械零部件应力集中部位的几何形状。

1）局部强化法　通过局部强化以提高应力集中处的材料强度，从而减少应力集中的危害。

① 表面热处理强化　包括表面感应加热淬火、渗碳、氮化和氰化等，可得到软的心部，硬的表层。在表层中还存在着残余压应力。这样降低了应力集中不良影响。

② 薄壳淬火　粗而短的且有截面变化的零件选用低淬透性钢，经强烈淬火后，工件表层获得一薄层淬硬层。其内部存在残余压应力。这种方法与表面感应淬火相比有其优点，即感应加热淬火，对这样的零件容易使截面变化的过渡区淬不上火而存在残余拉应力，反而促进了应力集中的有害作用。

③ 喷丸强化　喷丸处理能使金属表层强化且产生大的残余压应力，从而降低应力集中的有害作用，高强材料表面粗糙或有缺陷时，喷丸处理对降低应力集中的作用更明显。

④ 滚压强化　用滚轮对零件表面进行滚轧的强化方法称为滚压强化。它使零件表面形变强化并产生残余压应力，从而降低应力集中的有害作用。其效果与滚压参数及材料本身的组织性能有关。同样道理，对工件上的内孔，可采用涨孔的办法使孔边加工硬化，产生残余压应力。

2）改善机械零部件应力集中部位几何形状的方法

① 尽可能加大过渡部分的圆角　加大圆角时注意使其接近于流线型过渡，同时考虑其工艺性。如大型水轮机的空心主轴，在接近法兰处采用锥形过渡，并用大小双圆弧圆角，经试验和计算，其扭转应力系数只有 1.1 左右；而空心轴采用单圆弧过渡，其应力集中系数为 1.3～1.4。一般采用椭圆形过渡，也可使应力集中系数明显降低。

② 根据零件的受力位置和方向选择适当的开孔位置　一般开孔应选择在低应力区，如必须在高应力区开孔时，应采取补强措施。开椭圆孔时，其长轴方向应和主应力方向平行，如在一受拉平板上开一椭圆形孔，其长短轴之比应为 4:1。

③ 在应力集中区附近的低应力部位增开缺口或圆孔　这样可使力流能从低应力区平滑过渡到高应力区，从而降低应力峰值。如在受拉平板中间开一小圆孔，其应力集中系数为 3，而在受拉力方向离圆孔 $3r$（r 为圆孔半径）距离处再开一个同样大小的圆孔，其应力集中系数即降为 2.63。又如受拉力的板边开一个半圆形缺口，其应力集中系数为 3.06；开两个半圆形为 2.75；开五个半圆形时为 2.41。如将靠近两端的半圆形缺口半径逐渐减小，应力系数降低更明显。

1.1.2　抗断裂结构设计

断裂是一种重要的失效模式，它的发生常常引起灾难事故，因此，在重大机械产品设计中采用防断裂设计，将有助于机械产品设计水平的提高，保证机械产品的安全性和可靠性。在不可避免地存在缺陷或裂纹的构件中，可以防止事故的发生，减少不应有的损失。

防断裂设计国外常称为破损安全设计或损伤容限设计。它是从材料或结构在制造和运行时不可避免地产生缺陷的事实出发，认为一定条件下的缺陷是容许存在的，并能达到安全可靠的要求。通过防断裂设计，应使扩展的裂纹在达到保证安全所允许的尺寸界限前，及时查明其扩展情况并进行修补，以确保结构的正常工作。

断裂的常见模式分为脆性断裂、韧性断裂和疲劳断裂等，对于可能出现的不同的失效模式，在结构设计时应采取不同的设计思想和方法。

1.2　制造工艺及质量控制技术

1.2.1　制造工艺控制方法

在生产过程中，改变生产对象的形状、尺寸、相对位置和性质等使其成为成品或半成品的过程称为工艺过程。以工艺文件的形式确定下的工艺过程称为工艺规程。

由原材料经浇铸、锻造、冲压或焊接而成为铸件、锻件、冲压件或焊接件的过程，分别称为铸造、锻造、冲压或焊接工艺过程。将铸、锻件毛坯或钢材经机械加工方法，改变它们的形状、尺寸、表面质量，使其成为合格零件的过程，称为机械加工工艺过程。在热处理车间，对零件的半成品通过各种热处理方法，直接改变它们的材料性质的过程，称为热处理工艺过程。最后将合格的零件和外购件、标准件装配成组件、部件和机器的过程，称为装配工艺过程。

(1) 铸造工艺控制

铸件形成后，在最后凝固部件，由于收缩出现缩孔和缩松。为了控制缩孔和缩松应该采取以下控制工艺。

1) 控制铸件的凝固方向，凝固方向应符合定向凝固或同时凝固的原则。

2) 选择适当的浇注系统和浇注工艺，控制凝固顺序。

铸件在凝固后的冷却过程中将继续收缩，可能出现应力、变形或裂纹。

铸件的变形主要有扩口变形、翘曲变形和微量自变形等，可以分别采取以下控制工艺来控制变形。

1) 对于柔性结构铸件的扩口变形，可用拉肋消除开口的自由结构部分，使柔性结构铸件具有刚性，待铸件热处理后再去除拉肋；也可增加形成开口处的砂型或砂芯退让性，减小扩口变形。

2) 合理设计铸件结构，力求壁厚均匀，条形或板形铸件应具有刚性断面，有利于防止翘曲变形。合理的铸造工艺措施，实现同时凝固有利于防止翘曲变形。根据铸件翘曲变形的方向和大小，采用反挠度制作模样，也可消除铸件翘曲变形。采用干型、对柔性结构铸件应尽早卸除压箱铁或松开砂型的紧固装置等措施也可防止翘曲变形。

3) 选择硅/碳比值高的灰铸铁，改善铸造性能；使铸件各部分实现同时冷却和同时凝固，以减少铸件残余应力；采用分散引入内浇口，使铸件冷却时温度均匀，或在型砂或芯砂中加入木屑，增加铸型或砂芯的退让性，获得残余应力较低的灰铸铁件；在机械加工以前，对铸件进行时效热处理等措施都可以防止微量自变形。

铸件出现热裂纹时，在生产实际中应根据铸件结构的特点，正确采取各种铸造工艺措施，综合性地防止铸件热裂。

1) 应改善型砂和芯砂的退让性和溃散性。铸型、型芯紧实度不宜过高，芯骨尺寸不宜过大，砂箱箱带不宜过密，以保证型、芯有足够的吃砂量。

2) 采用分散的多道内浇；浇、冒口与铸件连接处要有适当圆角，浇冒口形状和位置不应阻碍铸件的正常收缩；浇注系统内浇道从铸件薄处分散引入型腔，实现同时凝固。

3) 控制浇注温度和浇注速度，避免因浇注温度过高使型砂烧结阻碍铸件收缩和凝固。

4) 尽量使铸件壁厚均匀，避免壁厚突变和多重交接。

铸件出现冷裂纹时，采取以下控制工艺措施。

1) 改善铸件结构，减少应力集中。

2) 控制有害杂质的含量，采用合理的熔炼工艺。

3) 采用正确的铸造工艺措施，包括铸件实现同时凝固、正确确定铸件在砂型中的停留时间、增加砂型、砂芯的退让性。

4) 及时进行时效热处理，避免过大的残余应力使铸件产生冷裂纹。

气孔是铸造过程中容易出现的另一种缺陷，根据气孔的形成机理不同可分为侵入气孔、卷入气孔、析出气孔和反应气孔。为了防止气孔产生，可采取的工艺措施包括：

1) 防止侵入气孔的措施　提高砂型和砂芯的透气性，控制砂型的紧实度；控制气体产生的速度；提高砂型和砂芯的排气能力；适当地提高浇注温度；适当地提高浇注温度；提高金属液的熔炼质量。

2) 防止卷入气孔的措施　使金属液平稳流动和充型，避免金属液直接进入直浇道；采用开放式浇注系统。

3) 防止析出气孔的措施　保证炉料清洁；合理地选择熔炼设备和熔炼工艺；控制炉料的配比；采用各种脱气方法；阻止气体析出；使金属液在压力下结晶。

4) 防止反应气孔的措施　避免在潮湿环境中浇注钢或铝铸件；对球墨铸铁要减少含气量、降低镁含量、控制砂型的水分；保持适当提高浇注温度，采用慢浇工艺。

铸件夹杂物是指铸件内部或表面存在的和基体金属成分不同的质点，包括渣、砂、氧化物等。防止这类缺陷的工艺控制措施为：

1) 熔炼时防止带入其他有害金属；

2) 造型操作中，保证砂型紧实度均匀，砂芯有足够的高温强度；保证砂型的完好；防止产生过烧现象。

3) 合理设计浇注系统，在浇注系统中加设挡渣装置。

铸造组织缺陷的防止措施：

1) 严格控制原材料；控制铁水的碳当量；提高煅烧温度和煅烧质量，防止铁水增碳。

2) 改进铸件结构，尽量使铸件壁厚均匀一致；造型时分型面要平整；提高铸型的冷却能力。

(2) 锻造工艺控制

锻造过程中容易出现氧化、脱碳、折叠和裂纹等缺陷，在锻造过程上应控制锻造工艺：

1) 防止氧化的措施　减少金属与氧的接触时间；在保护性介质中加热；采用先进加热技术。

2) 防止脱碳的措施　缩短加热总时间；控制加热气氛；加热时钢的表面用覆盖物或涂料加以保护。

3) 防止折叠的措施　合理选择毛坯尺寸；清除毛坯上毛刺和氧化皮；降低模具表面粗糙度；增大模具圆角半径；加强润滑；注意锻造时的送进量和操作方法等。

4) 防止裂纹的措施　①变形时尽量减小拉应力；②选择合适的变形温度；③控制应变速度；④中间退火；⑤提高材料的塑性。

(3) 焊接工艺控制

1) 热裂纹　焊接时，焊缝和热影响区金属冷却到固相线附近的高温区产生热裂纹，根据热裂纹的形态、产生机理以及产生的温度区间不同，可分为结晶裂纹、液化裂纹、多边化裂纹。

① 结晶裂纹的控制措施　a) 从冶金方面控制结晶裂纹；b) 制造工艺方面控制措施：防止熔池过热；适当增加焊缝成形系数；降低冷却速度；采用合理的焊接顺序。

② 液化裂纹的防止措施　a) 选择合适化学成分的母材；b) 控制焊缝的凹度；c) 采用小的焊接线能量。

③ 多边化裂纹的防止措施　a) 加入合金元素，阻碍位错移动，防止或减弱多边化晶界的出现；b) 在工艺上采用小的线能量、快的焊接速度，减少金属的过热。

2) 冷裂纹　防止焊接产生冷裂纹可以从两种方面入手。

① 冶金方面　选择高质量的钢材；选择低氢型的焊接材料；CO_2 焊接方法。

② 工艺方面 a）选择低氢型焊条；b）焊前清理焊丝、坡口及附近表面的油污和铁锈；c）焊前预热；d）适当加大线能量；e）焊后热处理。

3）气孔和夹杂 防止夹杂的措施是严格控制夹杂物的来源。限制焊丝、母材中杂质的含量。保证焊接工艺的正确。提高操作技能。

防止气孔时首先要控制冶金因素，在工艺上应适当增加熔池的存在时间。焊前对焊条进行烘干，对焊接区、焊丝充分考虑清理。

1.2.2 质量控制方法

产品质量控制是指为达到质量要求所采取的作业技术和活动。整个制造过程是由众多的工序组成的，工序状态的优劣，决定了产品质量的好坏，要想进行质量控制，必须进行工序质量控制。若工序发生异常，要迅速找出原因并消除异常，保持工序的稳定性。工序质量的稳定涉及到人、机器、材料、工艺方法、检测和环境（即5M1E）等因素。这些因素，特别是主导因素发生变化，就会直接影响工序质量的稳定，为此要管好这些因素和条件，做大量细致的工作。

工序质量控制是指为把工序质量的波动限制在规定的界限内所进行的活动。工序质量控制活动是制造过程的基本质量职能。其目的是使工序质量的波动处于规定的范围内，保证工序能稳定地制造出合格的产品。工序质量控制的对象是该工序的质量特性值的波动范围及质量特性值波动的中心。由于工序能力是受人、机器、材料、工艺方法、检测和环境等因素影响的，因此，工序质量控制就是控制这6大因素，通过控制这6大因素来实现对工序能力及分布中心位置的控制。这6大因素之所以影响产品质量发生波动，不外乎两类原因，即偶然性原因和系统性原因。

偶然性原因是对产品质量经常起作用的因素，如原材料成分、质量上的微小变化、机床的微小振动、刀具的正常磨损、夹具的微小松动、工人操作中的微小变化等。一般来说，这些微小变化的方向和大小都不一定对质量波动有多大影响，在技术上难以测量、消除和避免，在经济上也不值得消除。因此，由偶然性原因所引起的质量波动一般认为是正常的波动，不需要加以控制。

系统性原因是指原材料成分、规格的显著变化，机床与刀具的安装调整不当，夹具损坏，刀具的过度磨损，设备的故障，工人违反操作规程等。一般来说，这些变化对产品质量影响很大，在技术上易于识别，可以测量检查出来，所以是可避免的。因此，由于系统性原因所引起的质量波动，一般认为是不正常的波动，应严加控制。

由于工序种类多、工序因素复杂且所需的工具及方法是各式各样的，因此，现场工作人员应根据工序特点来选择合适的控制方法。

为了便于质量控制执行和检查，应当将“质量要求”转化为“质量特性”，并用定量或定性的规范来表示。

制造质量控制主要由三个重点子环节组成：原材料、毛坯、标准件和协作件质量控制；加工零件质量控制；部件和总装质量控制。

1）原材料的质量控制 不合格的原材料不准投料，这是把好产品制造质量的第一关。由生产厂家供应的原材料的质量多数都是合格的，但出厂的产品系抽样检查，难免出现漏检或误检，因此在下料前要对原材料进行严格地检验。另外，随着企业生产规模的扩大，需要购进大量的原材料。在原材料订货时应该提出合理的检验项目，到货后则按合同规定的检验基础上进行验收。原材料缺陷及检验方法一般有：非金属夹杂物的检验、偏析的检验、裂纹的检验。外购的材料还应对原材料进行化学成分和力学性能的验收检验，以便确信所订购的原材料符合质量要求。

2）毛坯的质量控制 毛坯是零件制造的第二个阶段，其质量直接影响产品最终的性能。因此在毛坯制成之后都应进行检验，发现缺陷及早剔除，并且要尽快查明缺陷产生的原因。毛坯质量检验主要有锻件缺陷检验、铸件缺陷检验和焊件缺陷检验等。

重要的锻件可用渗透法或磁粉探伤法检查表面缺陷，用超声波探伤法检查内部缺陷。铸件通常用金相检验或目检。对于焊件，其表面裂纹用液体渗透和磁粉探伤法检验；其内部焊缝缺陷采用超声波探伤法。

对承受压力的铸件或贮存流体的焊接压力容器等，除进行无损检验外，还要进行致密性试验，以确保安全。常用的检验方法有水压试验和气压试验等。

另外，锻件、铸件和焊件毛坯还应进行尺寸、形状和内部组织等检验，这些项目常用普通量具和分析仪器进行检测。

3）热处理缺陷及常用的检查方法 零件在热处理过程中，由于加热与冷却不均匀以及加热时与空气直接接触，因此在操作过程中很容易产生诸如过热与过烧、氧化与脱碳、硬度不足与软点、变形与开裂、组织不良等缺陷。其中脱碳、过热、过烧、组织不良等缺陷可以采用金相分析法进行检验；硬度不足与软点缺陷一般都采用硬度计进行检验；变形缺陷则可以采用目测或千分尺等量具进行检验；开裂缺陷可以采用目测、金相观察以及无损探伤等方法进行检验。

4）机械加工缺陷 零件的加工质量主要是指零件的加工精度能否满足设计要求，加工精度是指零件在几何量方面达到的精确程度。机械产品都是由许多机械零件组装而成的，这些零件的加工精度如何，将直接影响产品的使用性能和使用寿命，特别是对运动精度、定位精度、振动、噪声以及动力消耗等的影响更为突出。所以控制零件加工质量是控制产品质量的最基本的环节。

零件加工质量控制最直接有效的方法之一就是对加工精度进行跟踪，及时发现并削除产生废品的因素。跟踪检测有两种方式，分别称为主动检测和被动检测。主动检测多应用在现代化的自动生产线上，加工检测由检测装置自动完成。其检测过程与加工过程同时进行，相互储存，并以检测信息控制加工运动，因此很少出现废品；而被动检测过程与加工过程是分离的，分别单独进行，加工在前，检测在后。由于检测并不参与加工过程，所以不能制约加工精度，不能在加工过程中防止废品的产生。被动检测只能发现废品，剔除废品，通过分析研究废品产生的原因后再去调整加工过程。无论是主动检测还是被动检测，都是通过对零件的几何量检测确定零件的加工精度是否符合要求，达到控制零件加工质量的目的。

几何量检测的具体对象是各种机械零件的几何参数，抽象概括为长度、角度、长度与角度的复合量、形位误差和表面粗糙度。

长度是几何量中最基本的参数，不仅包括面与面、线与线、点与点之间的距离及点线面之间的距离，而且包括轴径和孔径等。角度也是几何量中最基本的参数，它不仅包括平面角，而且也包括立体角（锥度）和圆分度角。长度与角度的复合量是指螺纹和丝杠的螺旋线、分划板上的阿基米德螺旋线、凸轮的轮廓线、齿轮的渐开线齿形等，它们都是由长度量和角度量按不同规律变化时形成的轨迹。形位误差是指零件的形状误差（如圆度、圆柱度、直线度、平面度等）和位置误差（如平等度、垂直度、倾斜度、同轴度、对称度等）。表面粗糙度是指零件加工表面微观几何形状误差，是评定零件表面质量的重要参数。

5）装配质量控制 按规定的技术要求，将零件或部件

进行配合联接，使之成为半成品或成品的工艺过程，称为装配。装配工作是产品制造工艺过程中的最后一道工序，装配的好坏，对产品的质量起决定性作用。若装配不当，即使所有零件加工合格，也不一定能够装配出合格的高质量机械；反之，当零件的制造质量不十分良好时，只要装配中采用合适的工艺方案，也能使机械达到规定的要求。例如车床的主轴与床身导轨装配得不平等，车削出来的零件就会出现锥度；车出的零件端部就会不平。这样就会降低设备的使用寿命。装配得不好的设备，其生产能力就要降低，消耗的功率就会增加。

装配工作通常分为部件装配和总装配。部件装配是将两个以上的零件组合在一起或将几个零件组合在一起，成为一个装配单元的工作，都可以称为部件装配。总装配是将零件和部件组合成一台完整机械的工作过程。

1.3　表面防护与强化工艺技术

工程材料表面处理的目的是利用各种表面涂镀层及表面改性技术，赋予基体材料本身所不具备的特殊力学、物理和化学性能，从而满足工程上对材料及其制品的防腐蚀和耐磨等要求。

表面处理的方法可以分为表面强化处理、表面防腐及保护和表面装饰等几类。

1.3.1　表面强化方法

表面强化处理是材料表面处理的重要领域。它是通过材料表层的相变、改变表层的化学成分、改变表层的应力状态以及提高表层的冶金质量等途径来改变性能，从而达到强化表面的目的。目前常用的表面强化方法有以下几类。

1）表面的覆盖层强化法，它是在材料表面获得特殊性能的覆盖层，如气相沉积层、镀覆层、热喷涂层等，以达到提高强度、耐磨、耐蚀、耐疲劳等性能的目的。

2）表面形变强化法，它是通过喷丸、滚压、挤压等产生形变强化层，从而获得较高的疲劳强度的工艺方法。

3）表面热处理强化法，它是通过化学热处理，如渗碳、氮、硼及铬等金属和非金属元素，表面淬火，如感应加热、火焰加热、激光、电子束表面淬火等方法使表面强化。

4）表面复合处理强化法，是将两种以上的表面强化工艺复合用于同一工件上，在性能上可以发挥各自优点的处理方法，如渗氮后进行高频感应表面淬火、镀覆后进行热扩散等。

下面分别介绍几种主要的表面强化技术。

（1）表面覆盖层强化

1）气相沉积技术　气相沉积技术是利用气相之间的反应，在各种材料或制品表面沉积单层或多层薄膜，从而使材料或制品获得所需的各种优异性能。薄膜的制备方法很多，根据制备环境压力可分为真空、常压和高压三类。制备方法根据压力由低到高的顺序来排列有：真空蒸镀、离子镀膜、溅射镀膜、低压化学气相沉积、等离子体化学气相沉积、常压化学气相沉积、氧化法、电镀、涂敷沉淀法、高压氧化法等。

真空蒸镀处理的工件材料可以是金属、半导体、绝缘体乃至塑料、纸张、织物等，镀膜可以是金属、合金、化合物、半导体和一些有机聚合物等。溅镀膜的致密性和结合强度较好，基片温度较低，但成本较高。离子镀的镀膜致密，结合牢固，可在工件温度低于550℃时得良好的镀层。上述三种方法属于物理气相沉积的范围。

化学气相沉积是将工件放密封室，加热到一定温度，同时通入反应气体，利用室内气相化学反应在工件表面沉积成膜。源物质除气态外，也可以是液态和固态。所采用的化学反应有多种类型，如热分解、氢还原、金属还原、化学输运反应、等离子体激发反应、光激发反应等等。工件加热方式有电阻、高频感应、红外线加热等。主要设备有气体发生、净化、混合、输运装置以及工件加热、反应室、排气装置。主要方法有热化学气相沉积、低压化学气相沉积、等离子体化学气相沉积、金属有机化合物气相沉积、激光诱导化学气相沉积等。

2）热喷涂　热喷涂是将金属、合金、金属陶瓷材料加热到熔融或部分熔融，以高的动能使其雾化成微粒并喷至工件表面，形成牢固的涂覆层。热喷涂的方法有多种，按热源可分为火焰喷涂、电弧喷涂、等离子喷涂和爆炸喷涂等。经热喷涂的工件具有耐磨、耐热、耐蚀等功能。热喷涂技术具有适用范围广、工艺灵活、喷涂层的厚度可调范围大、工件受热程度可以控制、生产率高的特点，因此，热喷涂技术得到广泛应用。

（2）表面形变强化

表面形变强化是提高金属材料疲劳强度的重要工艺措施之一。通过机械手段在金属表面产生压缩变形，使表面形成形变硬化层，此形变硬化层的深度可达0.5～1.5 mm。产生形变的机械手段包括滚压、喷丸和内挤压等。在形变硬化层中产生两种变化：一是组织结构上，亚晶粒极大地细化，位错密度增加，晶格畸变度增大；二是形成了高的宏观残余压应力。经喷丸和滚压后，金属表面产生的残余压应力的大小，不但与强化方法、工艺参数有关，还与材料的晶体类型、强度水平以及材料在单调拉伸时的硬化率有关。具有高硬化率的面心立方晶体的镍基或铁基奥氏体热强合金，表面产生的压应力高，可达材料自身屈服点的2～4倍。材料的硬化率越高，产生的残余压应力越大。在三种常用的表面形变强化中，以喷丸强化应用最为广泛。

1）滚压　目前滚压强化作用的滚轮、滚压力大小等尚无标准。对圆角、沟槽等可通过滚压获得表层形变强化，并能在表面产生约5 mm深的残余压应力，其分布如图5.5-1所示。

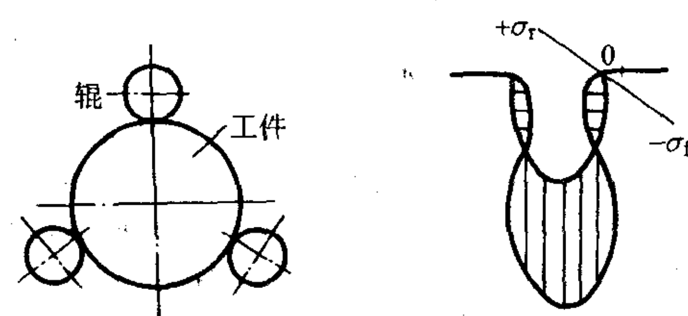


图5.5-1　表面滚压强化示意图

2）内挤压　内孔挤压是使孔的内表面获得形变强化的工艺措施，效果明显。

3）喷丸　喷丸是国内外广泛应用的一种再结晶温度以下的表面强化方法，即利用高速弹丸强烈冲击零部件表面，使之产生形变硬化层并引进残余压应力。喷丸强化已广泛用于弹簧、齿轮、链条、轴、叶片、火车轮等零部件，可显著提高抗弯疲劳、抗腐蚀疲劳、抗应力腐蚀疲劳、抗微动磨损、耐点蚀（孔蚀）能力。

（3）表面热处理方法

许多零件在工作中，表面层承受着比零件心部更高的应力和受到磨损。因而，要求零件的表面层具有高强度、高硬度、耐磨和抗疲劳能力，而心部则应保持较好的塑性和韧性，能经受重载荷作用和传递大的扭矩，为实现同一零件各部分性能不同的要求，表面改性处理是有效手段，其中表面热处理应用得较为广泛。

1）表面淬火　钢的表面淬火是一种不改变表层化学成分，但改变表层组织的处理方法。这种方法是快速加热使工

件表层奥氏体化，不等心部组织发生变化，立即快速冷却，表层起到淬火作用，其结构是表层获马氏体组织，而心部仍保持塑性、韧性均好的组织，工件各部性能都能满足使用要求。

表面淬火时的加热方法很多，如电感应、火焰、电接触、浴炉、电解液、脉冲能量等加热方法。目前采用较为普遍的是电感应加热和火焰加热。

① 感应加热表面淬火　感应加热表面淬火是把工件置于通有一定频率电流的感应器中，表面快速升温达淬火温度，随即将工件放入淬火介质中冷却的处理工艺。工件在感应器产生的交变磁场中，会形成涡流加热工件。通入感应器的电流频率越高，感应电流越向工件表面集中，被加热的金属层厚度越小，淬火后的淬硬层深度越小。

根据工件的情况适当选择感应加热器。对小型件（如小模数齿轮、中小型轴类）可选择高频（200 k ~ 300 kHz）感应加热，保证淬硬层深度在 0.5 ~ 2 mm 内即可满足要求。对于要求淬硬层深度在 2 ~ 8 mm 的大型轴类和大模数齿轮，适合选用中频（2 k ~ 2.5 kHz）的感应加热器表面淬火。对于要求淬硬层深度大于 10 mm 而又无高、中频感应加热的电源设备条件时，可采用工频（50 Hz）加热方式，用于大直径钢材的穿透加热和要求淬硬层深的大直径零件的表面淬火。对于采用高、中频感应加热难以实现沿零件轮廓表面淬火的零件（如中、小模数齿轮、花键轴、链轮等），可采用超音频（20 k ~ 40 kHz）感应加热，效果很好。

感应加热具有加热快、淬火组织细密、工件表面为残余压应力和生产率高，易实现机械化和自动化生产等优点，但加热需复杂昂贵的设备，形状复杂的零件，所需加热器制造困难，故不适合单件、小批量生产。

② 火焰加热表面淬火　火焰加热表面淬火是用高温火焰（如氧-乙炔火焰）快速加热工件表面达淬火温度，随即喷水使表面快速冷却，获得所需表面硬层的工艺方法，如图 5.5-2 所示。

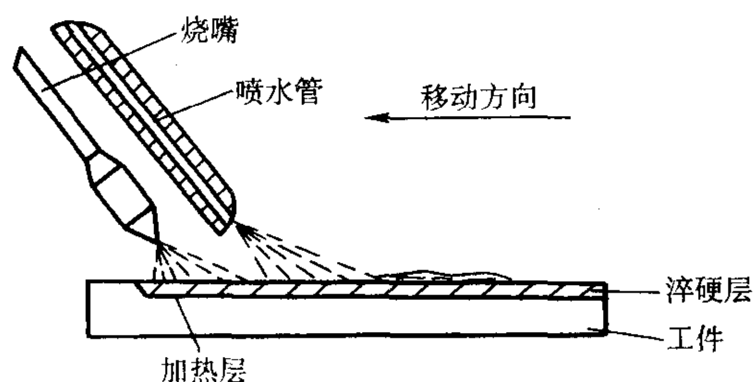


图 5.5-2　火焰加热表面淬火示意图

该法淬硬层一般为 2 ~ 6 mm，适用于用中碳钢、中碳合金钢制成的异型、大型或特大型工件的表面淬火。此法不需复杂设备，操作简便，但淬火效果不够稳定。

2) 化学热处理　化学热处理是将工件放在一定介质中加热和保温，使介质中的活性原子渗入工件表层，以改变工件表层的化学成分和组织，达到强化工件表面或保护工件表面的处理工艺。

① 渗碳　渗碳是指提高工件表面含碳量的工艺方法。目前工业中较为普遍的是固体渗碳和气体渗碳两种方法。

固体渗碳时采用木炭和碳酸盐（如 $BaCO_3$、Na_2CO_3 等），气体渗碳则采用煤油、丙酮、甲醇等有机物。渗碳剂在高温下都可分解出活性碳原子，被工件吸收，然后碳原子向表面深处扩散，从而使工件在一定深度的表层（0.2 ~ 2 mm）含碳量达 0.35% ~ 1.05%。渗碳中加热时间的长短、渗碳层的厚薄，视工件的大小和要求而定。

渗碳工件应选用低碳钢，经渗碳后表层与心部含碳量相差明显，否则不会表现出表层硬、耐磨和心部韧性和塑性好的特点。渗碳后的工件一定要进行淬火和低温回火处理，使表层具有回火马氏体结构和削除渗碳过程中形成的粗晶粒结构。

② 渗氮　渗氮是指提高钢件表面含氮量的处理工艺。工业中应用较广泛的是气体渗氮法。把氨气通入炉中，在 380℃以上的温度时，就可分解出活性氮原子，被工件表面吸收，并向内部扩散形成 0.1 ~ 0.6 mm 的渗氮层，即可明显改变工件表面的性能。硬度可达 69HRC，抗磨、耐腐蚀，并具有较好的抗疲劳性。

渗氮对钢种有要求，一般选用 38CrMoAl 进行渗氮效果较好。如果为获得高的耐腐蚀性可采用高温（达 720℃）、短时间处理工艺，获得 0.015 ~ 0.06 mm 厚的渗氮层，这种抗蚀氮化所适用的钢种可扩展到合金钢、碳钢和铸铁件。

③ 碳氮共渗处理　碳氮共渗处理是向钢的表面同时渗入碳和氮原子的处理工艺。其目的是提高钢的疲劳强度和工件表面硬度及耐磨性。

碳氮共渗时，是向炉内通入氨气和滴入煤油，在 820 ~ 860℃温度时，即可获得活性碳氮原子，被工件表面吸收，并逐渐扩散到内部，形成 0.2 ~ 1.0 mm 厚的共渗层。经淬火和低温回火后可获得所需性能。

碳氮共渗具有单独渗碳或单独渗氮的共同优点。表面硬度高且高硬区较深，表面脆性小，共渗过程中工件变形小，生产周期短。

1.3.2　表面防腐蚀方法

在材料表面施以覆盖层以达到防腐蚀的目的，常用的有电镀、化学镀、热浸镀、化学氧化、磷化、涂料涂装等处理技术。

（1）电镀

电镀是通过电化学沉积在基体金属表面形成金属镀层的一种表面加工方法，镀层的性能不同于基体金属，达到防护或装饰的目的。

镀锌主要用于钢铁等黑色金属的防腐蚀。电镀锌工艺可采用酸性和碱性电镀液两种，阳极使用纯锌。酸性电镀液价廉且电流效益高、电镀速度快，缺点是均镀能力差。碱性电镀液价格虽高，但均镀能力好。

镀镍可用作表面镀层，也可作为多层电镀的底层或中间层。

镀铬层具有美丽的光泽，耐腐蚀、硬度高且摩擦因数小，故用于装饰、耐磨损和耐蚀。

在阴极上同时沉积出两种或两种以上金属，形成结构和性能符合要求的镀层工艺过程，称为合金电镀。合金镀层具有许多单金属镀层所不具备的特殊性能。通过合金电镀还可以制取高熔点和低熔点金属组成的合金，以及具有优异性能的非晶态合金镀层。

（2）化学镀

化学镀是指在没有外电流通过的情况下，利用化学方法使溶液中的金属离子还原为金属并沉积在基体表面，形成镀层的一种表面加工方法。被镀件浸入镀液中，化学还原剂在溶液中提供电子使金属离子还原沉积在镀件表面。



$$M^{n+} + ne \rightarrow M \tag{5.5-2}$$

化学镀是一个催化的还原过程，还原作用仅仅发生在催化表面上，如果被镀金属本身是反应的催化剂，则化学镀的过程就具有自动催化作用。反应生成物本身对反应的催化作用，使反应不断继续下去。化学镀又称为自催化镀、无电电镀。

与电镀相比，化学镀具有其自身的特点。不管零件形状如何复杂，其镀层厚度都很均匀；镀层外观良好，晶粒细、无孔、耐蚀性更好；无需电解设备及附件；能在非金属（塑

料、玻璃、陶瓷等）以及半导体上沉积。化学镀的缺点是溶液稳定性差，使用温度高，寿命短。

由于化学镀层具有耐磨、耐蚀、高硬度、焊接性好、磁性等性能，在许多行业中得到广泛应用。

(3) 热浸镀

它是将工件浸在熔融的液态金属中，使工件表面发生一系列物理和化学反应，取出后表面形成金属镀层。工件金属的熔点必须高于镀层金属的熔点。常用的镀层金属有锡、锌、铝、铅等。热浸镀工艺包括表面预处理、热浸镀和后处理三部分。按表面预处理方法的不同，它可分为熔剂法和保护气体还原法。热浸镀的主要目的是提高工件的防护能力，延长使用寿命。

(4) 化学氧化和磷化处理

通过化学或电化学手段，使金属表面能形成稳定的化合物膜层，也称为化学转化膜。化学成膜的机理是金属与特定的腐蚀液接触而在一定条件下发生化学反应，由于浓差极化作用和阴极极化作用等，金属表面生成一层附着力良好的，能保护金属不易受水和其他腐蚀介质影响的化合物膜。由于化学转化膜是金属基体直接参与成膜反应而成的，因而膜与基体的结合力比电镀层和化学镀层大得多。根据形成膜时所采用的介质不同，可分为氧化物膜、磷酸盐膜、铬酸盐膜等。

钢铁在含有氧化剂的溶液中进行处理，使其表面生成一层均匀的蓝黑到黑色膜层的过程称为化学氧化，也称为“发蓝”或“发黑”，根据处理温度的高低，钢铁的化学氧化可分为高温化学氧化和常温化学氧化。

把金属放人含有锰、铁、锌的磷酸盐溶液中进行化学处理，使金属表面生成一层难溶于水的磷酸盐保护膜的方法，称为金属的磷酸盐处理，简称磷化。磷化膜层为微孔结构，与基体结合牢固，具有良好的吸附性、润滑性、耐蚀性、不黏附熔融金属（Sn、Al、Zn）性及较高的电绝缘性等。磷化膜主要用作涂料的底层、金属冷加工时的润滑层、金属表面保护层以及用作电机硅钢片的绝缘处理、压铸模具的防黏处理等。磷化膜厚度一般为5～20μm。

(5) 涂料涂装

用一定的方法将涂料涂覆于工件表面而形成涂膜的全过程称为涂装。涂料（或称漆）为混合物，一般由成膜物质、颜料、溶剂和助剂组成，可以涂装在各种金属、陶瓷、塑料、木材、水泥、玻璃等制品上。涂膜具有保护、装饰或特殊性能（如绝缘、标志等），应用十分广泛。

1.4　故障监测与检测技术

1.4.1　故障监测原理与方法

在线故障状态监测包括以下几个部分。

(1) 信号的在线检测

在线检测必须满足两方面的要求，一是在线，它是针对系统而言，对于连续运行的机械设备是指机器运行（生产）过程（系统）中的检测，是在生产线上进行，故属于在线检测。有些机械设备的运动，既有连续，又有中断，例如机床加工一个零件可看成是一个系统，则切削、换刀、上下料和测量都是系统的组成环节。虽然除了切削之外，在进行换刀、上下料等其他环节时，机床并不运动，刀具也不加工，但都属于系统的组成部分，故仍属在线；二是动态过程具有多方面的信息，没有必要都检测，所选择的信号应在机器上的部位都要能敏感地反映工况特征信息的变化。

传感器的精度和可靠性直接影响着整个监测系统的工作情况。根据输人量性质的不同，传感器分为加速度、速度、位移、温度、压力传感器等。按变换原理的不同可分为电阻式、电感式、电容式、压电式、磁电式传感器等。传感器的类型选择必须考虑在实际机器上安装的可行性。

(2) 信号的特征分析

鉴于直接检测信号大都是随机信号，它包括了大量的与故障无关的信息，但这种原始信号蕴含了系统状态的重要信息，不过人们很难直接地从它的外部变化观察出系统内部的规律，一般不宜用作判别量。需要用现代信号分析和数据处理方法把直接检测信号转换为能表达工况状态的特征量。对于某些具有规律的信号，也可以从波形结构上提取特征量。特征分析的目的是用各种信号处理方法作为工具，找到工况状态与特征量的关系，把反映故障的特征信息和与故障无关的特征信息分离开来，达到“去伪存真”的目的。因此，信号处理是特征分析的一种工具，但不是惟一的工具。用作特征分析的方法有频域分析、时域分析、统计分析、小波分析及波形结构分析等。

(3) 特征量的选择

用上述方法可以得到很多表达系统动态行为的特征量，但没有必要都用来判别工况状态。因为实际生产中，各个特征量对工况状态变化的敏感程度不同，应当选择敏感性强、规律性好的特征量，达到“去粗取精”的目的。对此，只有在系统建成之后，结合机组运行做实验，进行特征分析，才能知道哪一个特征量敏感或不敏感。实验室试验所得到的某种规律可作参考。选择对具体机器最敏感的特征量，才能加强监测诊断的针对性，提高诊断的准确性。特征量选择还要考虑判别的实时性，要求计算简单，如能在一定程度上表达工况状态的物理含义，就更有利于对工况状态变化原因的分析。用模式识别方法进行状态分类时，特征量的数量以2～3个为宜，一个太少，误判率大；而特征量太多，又使得判别函数复杂，计算量大，实时性差，且误判率并不因特征量的数量增多而单调地减少。在模式识别中，当特征量的数量增至3个以后，计算复杂，实时性差，而对降低误判率并无明显的改善。特征分析方法几乎包括现代所有的信息处理技术能提供的手段，如数字信号处理、时间序列分析、信息理论、图象识别及应用数学等。

(4) 工况状态识别

工况状态识别就是状态分类问题，分类与诊断往往是一个概念，此处从生产过程不同的目的考虑，把“分类”分成监测与诊断两个问题，工况监测的目的是区分工况状态是正常还是异常，或者哪一部分不正常，便于进行运行管理，强调在线和实时性。因为主要是正常与异常两种状态，用模式识别及模型参数判别都很有效。

1.4.2　故障检测技术

(1) 常用的检测技术

1) 直接观测法　应用这种方法时主要依靠检测人员的经验和感官，且限于观测到的或接触到的位置。这种方法可以获得第一手资料，更多的是用于静止的设备。在观测中有时使用了一些辅助的工具和仪器，如倾听机器内部声音的听棒、检查直接观测不到的零件表面有无表面缺陷的内窥镜、探查零件表面有无裂纹的磁性涂料及着色渗透剂等，来扩大和延伸检测人员的观测能力。详细内容可参考本篇第2章内容。

2) 参数测定法　根据设备运行的各种参数的变化来获取故障信息是广泛应用的一种方法。因为机器运行时各部件的运动必然会有各种信息，这些信息参数可以是温度、压力、振动或噪声等，它们都能反映机器的工作状态，如根据轴瓦下部油压变化可以了解转子对中情况，分析油液中金属碎屑情况可以了解轴瓦磨损程度等。在运转的设备中，振动是最重要的信息来源，在振动信号中包含了各种丰富的故障信息，任何机器在运转时的工作状态发生了变化，必然会从振动信号中反映出来。对旋转机械来说，目前在国内外应用

最普遍的方法是利用振动信号对机器状态进行判断，首先应从振动信号中提取有用的特征信号，即利用信号处理技术对振动信号进行处理。目前应用最广泛的处理方法是进行频谱分析，即从振动信号中的频率成分和分布情况来判断故障。

其他如噪声、温度、压力、变形等参数也是故障信息的重要来源。

3）磨损残余物的测定　机器零件在运行过程中的磨损残渣可以在润滑油中找到。测定的方法分为三种：第一种是直接检查残渣，以及测定油膜间隙内电容或电感的变化、润滑油混浊度的变化等方法以迅速获得零件失效的信息。第二种是残渣的收集，例如采用磁性探头、特殊的过滤器等收集齿轮、滚动轴承等工作表面疲劳引起的大块剥落颗粒。第三种方法是油样分析，可以确定机器中什么零件在磨损。

测定机器零部件如轴承、齿轮、活塞环等的磨损残渣在润滑油中的含量，也是一种有效的获取故障信息的方法。根据磨损残渣在润滑油中含量及焊料分布可以掌握零件磨损情况，并可预防机器故障的发生。

4）设备性能指标的测定　设备性能包括整机及零部件性能。通过测量机器性能及输入、输出量的变化信息来判断机器的工作状态也是一种重要方法。

(2) 典型故障检测技术

由于机器运行的状态千差万别，因而出现的故障也是多种多样，采用的检测方法也不相同。

1）振动和噪声的故障检测　机械在运行的过程中振动是诊断的重要信息，它反映了机器的状态。根据检测的振动和噪声信号，可采用以下方法对机器状态进行诊断。

① 振动法　对机器主要部位的振动值如位移、速度、加速度、转速及相位值等进行测定，与标准值进行比较，据此可以宏观地对机器的运行状况进行评定，这是最常用的方法。

② 特征分析法　对测得的上述振动量在时域、频域、时-频域进行特征分析，用以确定机器各种故障的内容和性质。

③ 模态分析与参数识别法　利用测得的振动参数对机器零部件的模态参数进行识别，以确定故障的原因和部位。

④ 冲击能量与冲击脉冲测定法　利用共振解调技术以测定滚动轴承的故障。

⑤ 声学法　对机器噪声的测量可以了解机器运行情况并寻找故障源。

2）裂纹及缺陷损伤的故障检测　材料中出现的裂纹可以采用以下方法进行检测。

① 超声探伤法　该方法成本低、可测厚度大、速度快、对人体无害，主要用来检测平面型缺陷。

② 射线探伤法　主要采用 X 射线。该方法主要用于检测体积型缺陷，适用范围较广，测量成本较高，对人体有一定损害。

③ 渗透探伤法　主要有荧光渗透与着色渗透两种。方法简单、成本低，应用范围广，可直观显示，但仅适用于有表面缺陷的损伤类型。

④ 磁粉探伤法　该法使用简便，较渗透探伤更灵敏，能探测近表面的缺陷，但仅适用于铁磁性材料。

⑤ 涡流探伤法　这种方法对封闭在材料表面下的缺陷有较高的灵敏度，它属于电学测量方法，容易实现自动化和计算机处理。

⑥ 激光全息检测法　它是 20 世纪 60 年代发展起来的一种技术，可检测各种蜂窝结构、叠层结构、高压容器等。

⑦ 微波检测技术　它也是近几十年来发展起来的一种新技术，对非金属的贯穿能力远大于超声波方法，其特点是快速、简便，是一种非接触式的无损检测方法。

⑧ 声发射技术　它主要对大型构件结构的完整性进行监测和评价，对缺陷的增长可实行动态、实时监测，且检测灵敏度高，目前在压力容器，核电站重点设备及放射性物质泄漏，输送管道焊接部位缺陷等方面的检测获得了广泛的应用。

3）设备零部件材料的磨损及腐蚀故障检测　这类故障除采用上述无损检测中的超声探伤法以外还可用以下方法。

① 内窥镜　它是利用特制的光纤内窥探测器直接观测到材料表面磨损及腐蚀情况。

② 油液分析技术　油液分析技术可分为两类：一类是油液本身的物理、化学性能分析；另一类是对油液污染程度的分析。具体的方法有光谱分析法与铁谱分析法。

4）温度、压力、流量变化引起的故障检测　机器设备系统的有些故障往往反映在一些工艺参数，如温度、压力、流量的变化中，在温度测量中除常规使用的装在机器上的热电阻、热电偶等接触式测温仪外，目前在一些特殊场合使用的非接触式测温方法有红外测温仪和红外热像仪，它们都是依靠物体的热辐射进行测量。

1.5　故障排除与修理技术

机械设备出现故障时，常常出现部分零件变形、断裂、腐蚀或磨损等情况，零件修复是机械设备修理的一个重要组成部分，应用各种修复新技术修理设备是提高设备维修质量、缩短修理周期、降低修理成本、延长设备使用寿命的重要措施，尤其对贵重、大型零件、加工周期长、精度要求高的零件，需要特殊材料特种加工的零件，意义更为突出。通常，修复失效零件与更换零件相比具有以下优点：节省时间；节约材料；可降低备件的消耗；可避免因其备件不足而停修；修复旧件一般不需要大、精、稀设备，牵涉人力少，易组织生产；利用新技术修复旧件还可以提高零件的某些性能，延长寿命。

零件修复方法种类很多，每种修复方法各有其优点，也有局限性，所以应根据自身条件和修理范围适当选择，使修复层与基体结合牢固，使修复工艺对基体金属的不良影响最小，使修复后的零件获得优良的性能，并使修复工作取得良好的经济效益。常用的修复技术按其所采用的工艺手段分为机械修复法、焊接修复法、电镀修复法、喷涂修复法及黏接修复法等。

(1) 机械修复法

利用机械连接，如螺纹联接、键、铆接、过盈联接等使磨损、断裂、缺损的零件得以修复的方法称为机械修复法。如镶补法、金属扣合法等，这些方法可利用现有的简单设备与技术，进行多种损坏形式的修复。其优点是不会产生热变形；缺点是受零件结构、强度、刚度的限制，难以加工硬度高的材料，难以保证精度要求高的材料。

1）金属扣合法　金属扣合法修复技术是借助高强度合金材料制成的扣合连接件（波形键），在槽内产生塑性变形来完成扣合作用，以使裂纹或断裂部位重新连接成一个整体。该方法适于不易焊补的钢件和不允许有较大变形的铸件，以及有色金属件，对大型铸件的裂纹或断面的修复效果更为突出。

金属扣合法的特点是：修复后的零件具有足够的强度和良好的密封性；修复的整个过程在常温下进行，不会产生热变形；波形槽分散和排列，波形键分层装人，逐片铆击，不产生应力集中，操作简便，使用的设备和工具简单，便于就地修理。该方法局限性是不适于修复厚 8 mm 以下的铸件及振动剧烈的工件，此外，修复效率低。

① 强固扣合法　该方法是先在垂直于裂纹方向或折断面的方向上，按要求加工出具有一定形状和尺寸的波形槽，然后将用高强度合金材料制成的其形状、尺寸与波形槽相吻合的波形键嵌入槽中，并在常温下铆击使之产生塑性变形而充满整个槽腔，这样，由于波形键的凸缘与槽的凹缘相互紧密的扣合，将裂开两部分牢固地连接成一体，如图5.5-3所示。这种方法适用于修复壁厚8～40 mm的、一般强度要求的机件。

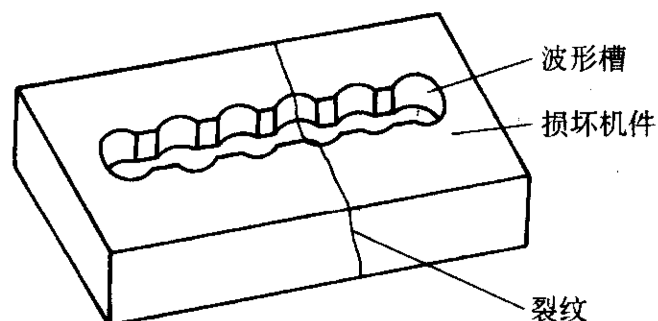


图5.5-3　强固扣合法

② 强密扣合法　对有密封要求的修复件，如高压汽缸和高压容器等防渗漏零件，应采用强密扣合法进行修复。这种方法是先用强固扣合法将产生裂纹或折断面的零件联接成一个牢固的整体，然后按一定的顺序在断裂线的全长上加工出缀缝栓孔。注意应使相邻的两缀缝件相割，即后一个缀缝栓孔应略切入上一个已装好折波形键或缀缝栓，以保证裂纹全部由缀缝栓填充，以形成一条密封的金属隔离带，起到防渗漏作用，如图5.5-4所示。对于承受较低压力的断裂件，采用螺栓形缀缝栓，对于承受较高压力、密封性要求较高的机件，采用圆柱形缀缝栓。

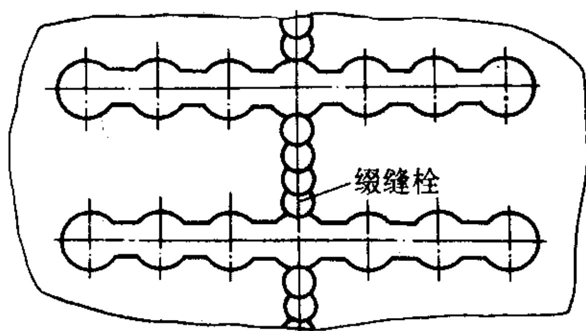


图5.5-4　强密扣合法

③ 加强扣合法　对承受高载荷的机件，只采用波形键扣合而其修复质量得不到保证时，需采用加强扣合法。其方法是：在垂直于裂纹或折断面的修复区上加工出一定形状的空穴，然后将形状尺寸与之相同的加强件嵌入其中。在机件与加强件的结合线上拧入缀缝栓，使加强件与机件得以牢固联接，以使载荷分布到更大的面积上。此法适用于承受高载荷且壁厚大于40 mm的机件。缀缝栓中心布置在结合线上，使缀缝栓一半嵌入加强件，另一半嵌入机件，相邻两缀缝栓彼此重叠0.5～1.5 mm，如图5.5-5所示。缀缝栓材料与波形键相同，其尺寸可参照波形键凸缘及波形键深度。加强块形状可根据载荷性质、大小、方向设计成楔形、十字形、X形、长方形等。

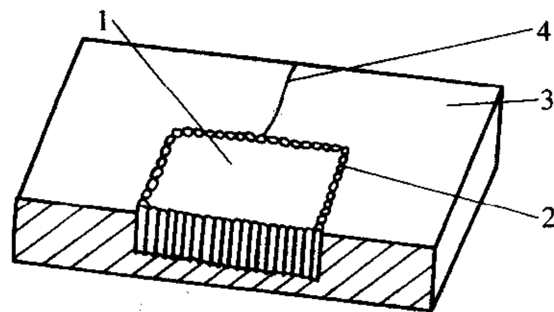


图5.5-5　加强扣合法
1—加强件；2—缀缝栓；3—机件；4—裂纹

④ 热扣合法　利用金属热胀冷缩的原理，将一定形状的扣合件经加热后扣入已在机件裂纹处加工好的形状尺寸与扣合件相同的凹槽中，扣合件冷却后收缩将裂纹箍紧，从而达到修复的目的。

2）镶补法　镶补法就是在零件磨损或断裂部位外补以加强板或镶装套等，使其恢复功能。一般中小型零件断裂后，可在其裂纹处镶加补强板，用螺钉或铆钉等将补强板与零件连接起来（图5.5-6）；对于脆性材料，应在裂纹端头钻止裂孔。这种方法操作简单、适用面广。

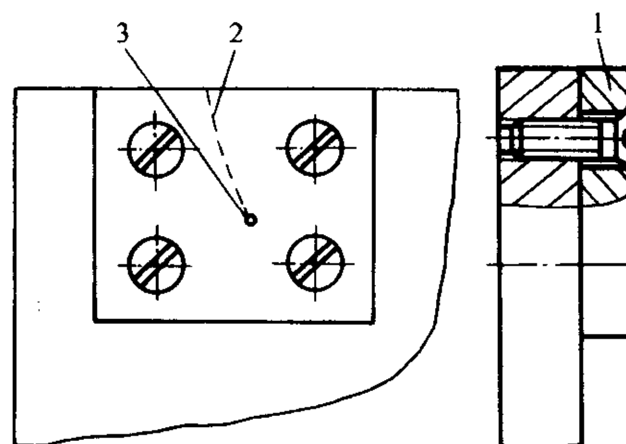


图5.5-6　补强板
1—补强板；2—裂纹；3—止裂孔

3）调头法　某些零件局部磨损后，可采用调头或转向的方法重新满足使用要求。如丝杆的螺纹局部磨损后，可调头重新使用。调头转向重新使用的零件必须是构造对称或经简单加工即可满足使用要求的零件。

（2）焊接修复法

对于已经磨损的零件或出现裂纹及已经断裂的零件，可以通过焊接的方法进行修复。焊接技术用于修补零件缺陷时称为补焊；用于恢复零件几何形状及尺寸，或使其表面获得具有特殊性能的熔覆金属时称为堆焊。这种修复方法与传统的焊接没有本质的区别，但它焊接的是失效后的机械零件，因此要根据不同的对象来选定不同的焊接工艺。这种方法适用性较广，大部分金属零件都可以用焊接修复法修复；焊接修复后结合强度高；实施时不受零件尺寸、修复场地的限制。但焊接修复温度较高，易引起金相组织变化并产生应力及变形，不易修复精度较高、细长及较薄的零件，同时易产生气孔、夹渣、裂纹等缺陷。

（3）热喷涂修复法

热喷涂修复法可修复工件磨损、缺陷，也可使工件表面获得耐磨、防腐、隔热、导电、绝缘、密封、润滑等特殊性能。适于现场修复、成本低、周期短。它的适用性强，几乎各种能加热到熔化或半熔化状态的材料，如金属、合金、陶瓷、塑料、氧化物、碳化物、硅化物、氮化物、有机树脂等均可制成粉末或丝状喷成涂层，被喷涂的工件可以是金属，也可以是非金属。喷涂时工件温度较低，因此，不易引起变形及相变。这种方法的缺点是涂层的结合强度不高，约40～90 MPa。喷涂时雾点分散、材料附着率低、损失较严重。喷涂时工件表面需粗糙化处理，会降低工件的强度。喷涂层为多孔组织，易存油、有利于润滑，但不利于防腐蚀。根据热源来分；热喷涂有四种基本方法：火焰喷涂、电弧喷涂、等离子喷涂和特殊方法。

（4）电镀修复法

对于磨损零件需要恢复尺寸时，可以采用电镀修复法，它同时也可以用于改善零件表面的性质，如提高耐磨、耐蚀及表面导电性、改善润滑条件等。用于零件修复时的镀铬层主要是硬铬层，其目的是利用铬的硬度高、耐磨性和抗腐蚀性好的特性。镀铬层与基体金属的结合强度特别高，但主要缺点是脆性大，一般允许厚度为0.2～0.3 mm。镀铁修复原料便宜，镀层厚度可达2 mm，镀层硬度不高，但韧性好、

镀铁层与基体金属结合良好。镀镍层的硬度为 150～500HV，可用于磨损、腐蚀零件的修复。电刷镀设备简单轻便，无需镀槽、挂具，可进行现场修复，工艺，镀层沉积速度快，镀层纯度高，结晶致密，氢脆性低，结合强度高，可有效控制镀层厚度，镀层种类多，适应材料广、污染小。

(5) 胶接修复法

胶接修复技术近年来发展迅速，它具有如下特点：①不受材质限制，可胶接金属、非金属，也可接异种材料；②不削弱基体的强度，不产生应力集中、耐疲劳性好，粘接的构件质量小、平整美观；③胶接工艺所需温度不高，不会引起基体热变形及组织变化，不易产生裂纹，可修复铸铁件、有色金属、合金件、极薄件、微小件和细长件；④胶接工艺简单、易操作，所用设备简单，便于现场修复，修复成本低；⑤可赋予胶接面密封、绝缘、隔热、防腐、防振、导电等性能。胶接的缺点是：①不耐高温，大多数胶接剂只能在150℃以下长期工作；②多数胶剂固化后脆性较大、抗冲击性、抗剥离、抗老化性能差。

2　抗失效的材料选择和材料设计技术与方法

抗失效的材料选择和设计主要包括化学成分及组织结构设计、加工工艺设计、经济成本分析等，从材料成分和性能角度考虑，主要有化学成分及组织结构设计以及加工工艺设计两种方法。材料的性能决定于化学成分和组织结构，在加工工程中材料的组织将发生变化，控制加工工艺往往能够获得好的结果，提高材料性能。

材料的成分与组织结构关系是建立在合金热力学、动力学、固体结构等基础理论之上的。运用这些基本理论，改变成分、组成相的数量、尺寸、形状、分布等便可以改变材料性能。因此，材料的成分与组织设计是抗失效材料设计的核心。

材料的化学成分是根据零部件使用性能确定的，但是仅凭材料成分还不能完全控制性能，因为组织结构起着更直接作用，所以运用组织性能关系，可以更有效地预测材料性能。

合金相图是确定材料成分与组织结构关系的依据。由于合金成分与组织结构互为因果关系，所以根据性能可以选定合金成分，从而就确定了组织状态；同样根据性能选定材料的组织状态，从而也就确定了合金成分。

抗失效材料的加工工艺设计较为复杂。对于尺寸较小、形状简单的零件制定加工工艺比较容易；而尺寸较大、形状复杂的零件则比较困难。不论采用什么加工方法，首先必须保证零件所要求的使用性能；其次是完成规定的生产效率，最后是低的经济成本。

材料加工工艺设计除考虑产品性能外，产品的形状、尺寸、质量，特别是产量是重要内容。机械产品质量包括材料质量与加工质量，两者既有差别，又有联系。产量不同采用的生产手段也不同。小批量生产可用简易设备，损伤技术较简单、投资较少，但是产量低，单件产品成本高，产品质量稳定性差。大量生产时采用机械化、自动化程度高的设备，虽然投资大，技术复杂，但是产品质量稳定，单件产品成本低。

2.1　抗变形失效材料的选择和设计方法

变形失效是机械装备使用时经常发生的，变形包括弹性变形和塑性变形，为了消除或减小变形失效的发生，可以从材料的选择和设计方面考虑，也可以从结构设计方面加以考虑。

2.1.1　抗变形失效材料成分的选择和设计

金属的弹性行为起源于其晶体点阵内原子间的相互作用，这种相互作用可以分为两大类：即吸引作用和排斥作用，对以金属键结合为主的晶体，粗浅地说，吸引力是金属正离子与公有电子间库仑引力作用的结果，这个力是长程的，也就是说在比原子间距大得多的距离处它就起作用。排斥力的来源有二，对于离子半径和原子半径比大的金属，原子间的排斥力主要是由于离子间相互作用引起的；对于离子半径和原子半径比小的金属，原子间的排斥力主要是由于原子接近时电子加速运动引起的。排斥力是短程的，也就是说只有两原子的距离在接近原子间距时，才能显示其重要作用。如果原子间距离进一步减小，斥力的增长速度将剧烈增大，使斥力大于吸力，因此，原子间两种力的大小是随原子间的距离而变化的。当金属中任何一个原子和其近邻的原子交互作用力达到平衡的时候，金属就处于最低的能量状态。当施加外力时，这种平衡被破坏，在外力和吸引力、排斥力之间建立起新的平衡，从而使原子稳定在新的平衡位置发生了原子间距变化，即出现了所谓的变形。如果此时这个外力还不足以在晶体内部产生其他质的变化，如位错不可逆运动或裂纹形成和扩展时，则一旦此外力去除后，由于这种新的平衡又被破坏，为了恢复平衡，原子将重新回复到原来所在的吸力和斥力保持平衡的位置。于是，宏观上物体又恢复原状了，这就是弹性变形的物理过程。

弹性变形可能是由机械应力造成的，也可能是由于材料受热后发生热膨胀造成的。

弹性模量是表征材料抵抗弹性变形的能力，其数值的大小反映材料弹性变形的难易程度，弹性模量值越大，在相同的应力下材料的弹性变形便越小。

不同材料的弹性模量的大小差别很大。这主要是由于各种材料具有不同的结合键和键能。由表 5.5-1 可以看出，具有强化学键结合的材料的弹性模量高；分子间仅由弱范德瓦耳斯力结合的材料则弹性模量很小。

表 5.5-1　不同材料弹性模量和键型

材料	E/MPa	键型
铁及低碳钢	20700	金属键
铜	12100	金属键
铝	6900	金属键
钨	41000	金属键
金钢石	114000	共价键
Al_2O_3	40000	共价键和离子键
石英玻璃	7000	共价键和离子键
电木	500	共价键
硬橡胶	400	共价键
非晶态聚苯乙烯	300	范氏力
低密度聚乙烯	20	范氏力
铸铁	(17.3～19.4) ×10^3	金属键
低合金钢	(20.4～21) ×10^3	金属键
奥氏体不锈钢	(19.4～20.4) ×10^3	金属键

热膨胀也会引起弹性变形，热胀冷缩是人们所共知的自然现象。线胀系数是表征这一特性的参数。不同材料具有不同的线胀系数。如果材料匹配不当，在温度改变时就可能引起弹性变形。例如，钢的线胀系数约为 12×10^{-6}℃$^{-1}$，是青铜的一半，如果用 2Cr13 不锈钢作轴套，用青铜作轴瓦，这样的结构在常温下可以很好地工作，但当温度很低时，就会因轴套的收缩远小于轴瓦的收缩而发生抱轴现象。因此，为

了避免因热膨胀而引起的弹性变形，需要选择线胀系数匹配的材料。

2.1.2　抗变形失效材料热处理工艺选择和设计

(1) 抗弹性变形失效材料热处理工艺选择和设计

金属弹性主要取决于它的键合强度，对组织结构的变化相对来说不太敏感，少量的合金化和热处理对弹性模量影响不大。

形成固溶体的合金，其弹性模量可以升高也可以降低。一般说来，固溶体中的溶质元素，在其周围晶体中引起畸变，使弹性模量降低，但是若溶质原子能使溶质与基体原子间键合强度较基体与基体原子间的键合强度更强时，则弹性模量将升高。实验观察表明，如果溶质原子降低基体金属的熔点，则也将使合金的弹性模量下降（如铜中若加入能显著降低其固溶体熔点的元素，其弹性模量明显下降）。就许多固溶体合金而言（如铜－镍等），随着溶质浓度的增加，合金弹性模量变化的趋势大致和固相线的走向一致。

第二相对弹性模量的影响大致可按两相混合物体积比例的平均值计算。对铝合金的研究表明，硬的第二相粒子可以提高合金的弹性模量，原因在于第二相具有较基体为高的弹性模量。事实上，Al-Si 或 Al-Cu 合金不论是淬火成固溶体还是时效处理产生沉淀相都能提高弹性模量。铍青铜时效后 E 值的提高可超过 20%，看来也是和铍的析出有关，因为铍有着高的弹性模量。

中高碳钢淬火后弹性模量下降通常不到 10%，随着回火温度的增加，弹性模量不断回升。如上所述，前者是因为淬火加剧了点阵畸变，削弱了表征键合强度的弹性模量；后者是因为畸变的点阵在回火中逐渐得到了恢复，同时还析出了高弹性模量的碳化物所致，不过由表 5.5-2 可见，这些变化在数值上都是很小的，这表明热处理不能显著改变弹性模量，因此，不能指望它来改变零件的刚度。

表 5.5-2　几种弹簧钢的切变模量 G　MPa

50CrMn		50CrVA		$60Si_2$	
热处理状态	G	热处理状态	G	热处理状态	G
退火	84340	退火	83850		
840℃淬火	80800	淬火	80610	淬火	79240
淬火 300℃回火	81600	淬火 300℃回火	81890	淬火 300℃回火	80610
淬火 400℃回火	82280	淬火 400℃回火	82580	淬火 400℃回火	81790
		淬火 500℃回火	83650	淬火 500℃回火	83360
淬火 600℃回火	84630	淬火 600℃回火	84930		

冷加工可降低金属及合金的弹性模量，但一般改变也不大（大量冷变形后 E 值可减小 5%），钢在强烈的有方向性变形后 E 值可大为增大，这是因为择优取向形成了方向性织构的关系。

总结上述诸影响因素讨论可知，弹性模量是金属的一个对成分组织都不敏感的性质，是金属最稳定的力学性质，很少受材料各种内在因素和零件外在服役条件的影响，因此，可以认为，至今作为人们改变金属力学性能的一些重要手段如合金化热处理和冷热加工在这个问题上是很难发挥作用的。

(2) 抗塑性变形失效材料热处理工艺选择和设计

纯金属单晶体的屈服强度从理论上讲与使位错开动的临界分切应力有关，其值由位错运动所受的各种阻力决定。这些阻力主要是点阵阻力、位错间交互作用产生的阻力、位错与其他晶体缺陷交互作用的阻力等。

点阵阻力是使一个位错在晶体中运动所需克服的阻力，也称为派－纳力，以 τ_{P-N} 表示，它与晶体结构和原子间的作用力等因素有关：

$$\tau_{P-N}=\frac{2G}{1-\nu}\exp\left(-\frac{2\pi w}{b}\right) \qquad (5.5\text{-}3)$$

式中，w 为位错宽度，$w=\frac{a}{1-\nu}$；a 为滑移面的面间距；b 为滑移方向上的原子间距。

从这个式子可以看出，位错宽度越大，派－纳力越小，因为这时位错周围的原子比较接近于平衡位置，点阵弹性畸变能低，位错易于移动。对于面心立方金属，w 较大，故派－纳力较小，屈服应力低，而体心立方金属的位错宽度较小，派－纳力较大，屈服应力高。

位错间交互作用力来自平行位错的长程弹性相互作用，相交位错产生会合位错，以及运动位错与穿过滑移面的位错交截产生割阶的作用，这些阻力均正比于 Gb，反比于位错间距离 l，可用下式表示：

$$\tau=a\frac{Gb}{l} \qquad (5.5\text{-}4)$$

因为位错密度 $\rho\propto\frac{1}{l^2}$，所以式 (5.5-4) 又可以写成

$$\tau=aGb\sqrt{\rho} \qquad (5.5\text{-}5)$$

式中，a 为与晶体本性、位错结构及分布有关的比例系数，面心立方 $a\approx0.2$，体心立方金属 $a\approx0.4$。

由此可见，位错密度增加，临界切应力也增大，所以屈服应力随之提高。因此，要提高屈服强度，应增加晶体中的位错密度，晶体经过剧烈的冷变形，其位错密度可增加 4～5 个数量级，从而显著提高屈服应力。

若将多晶体中的晶粒看作单晶体，则上述分析也适用于多晶体。但多晶体中存在晶界，多晶体的位错运动还必须克服界面阻力。因为晶界两侧晶粒的取向不同，因而其中一个晶粒滑移并不能直接进入邻近晶粒，于是位错在晶界附近塞积，造成应力集中，从而激发相邻晶粒中的位错源开动，才能引起宏观的屈服应变。这就是著名的 Hall-Petch 公式：

$$\sigma_s=\sigma_0+kd^{-\frac{1}{2}} \qquad (5.5\text{-}6)$$

由上式可见，细化晶粒是提高金属屈服强度的有效方法，而且细化晶粒还可以提高塑性和韧性。因为晶粒越细小，越难造成裂纹形核所需要的应力集中，而且裂纹在不同取向的各个晶粒内传播也更困难。因此细化晶粒是金属强韧化的好办法，是其他强化方法难以达到的。近年来更发展了超细晶粒处理方法来大幅度改善钢的性能。

纯金属的屈服强度很低，即使经过形变硬化和细晶强化，仍不能作为结构材料加以应用。常常使用合金材料，其组织多为固溶体基体加第二相质点。

固溶强化处理——纯金属中加入溶质元素，形成间隙型或置换型固溶体，显著地提高屈服强度，固溶体中溶质原子越多，固溶强化效果越高，而且间隙固溶体的强化效果比置换固溶体的大。

固溶强化是由许多方面的作用引起的，主要有溶质原子与基本的弹性交互作用、电化学作用、化学作用以及几何作用。

合金元素的作用主要是改善金属的力学性质，即提高强度或改善塑性。合金元素对于金属力学性质的影响是多种多样的。合金元素在基体中的分布状态也有好几种：均匀的单相固溶体、有成分偏聚的固溶体、有序化的固溶体及金属间化合物，还有通过脱溶沉淀或粉末冶金等方法所获得的复相合金。但总的来说，合金强化可以分为两种类型：固溶强化和沉淀强化。

固溶强化是利用点缺陷对金属基体进行的强化。具体的方式是通过溶入某种溶质元素形成固溶体而使金属强度、硬度升高，例如，将 Ni 溶入 Cu 的基体中，得到的固溶体的强

度就高于纯铜的强度。

固溶强化根据溶质原子占据的位置不同有间隙式和替代式的差异。间隙式固溶强化是将碳、氮等小溶质原子嵌入金属基体的晶格间隙中，使晶格产生不对称畸变造成的强化效应。间隙式原子在基体中还能与刃位错和螺位错产生弹性交互作用，并使两种位错钉扎，过一步强化了金属。替代式溶质原子在基体晶格中造成的畸变大都是球面对称的，因而强化效果要比间隙式原子小。

第二相强化——合金中第二相的存在可能有两种情况：聚合型与弥散型。聚合型第二相尺寸与基体晶粒尺寸处于同一数量级，常呈片状、块状，如钢中的珠光体，$\alpha+\beta$ 两相黄铜中的 β 相；弥散型第二相以细小弥散的质点均匀分布于基体相内。

对于聚合型合金，其强度决定于第二相对位错运动的阻力，一般，第二相阻碍滑移使基体产生不均匀变形，由于局部塑性约束而导致强化。

对于弥散型合金，第二相通常是增强相，比基体——固溶体的硬度高得多。第二相强化与第二相的数量、尺寸、形状、分布，第二相与基体的强度、塑性及形变强化特性，第二相与基体相晶体学配合等许多因素有关。第二相质点的强化作用主要是因为质点的成分与性质不同于基体，在质点周围形成应力场，而这些局部应力场对位错运动有阻碍作用。但是，这种阻碍作用能够发生的条件，是位错能沿着第二相质点引起的应力场弯曲，而取得最小的位能。当质点间距与位错可被弯成的半径相等，位错所遇到的障碍等于它所遇到的质点应力场的代数和，因而合金得到最大的强化。当质点间距小于位错可被弯成的半径时，位错的弯曲不能像质点应力场那样急剧，位错线两边的应力场部分地相互抵消，因而不能得到最大的强度。当质点间距大于位错的弯曲半径时，位错线绕过质点所需的应力降低，合金的强度下降。第二相质点的尺寸、间距以及形状和分布可通过热处理来调整。

复合材料提高基体强度的物理原因与第二相弥散强化类似：强度与增强体均匀分布的平均间距的平方根成正比。而基体或增强体中某一软相有利于应力弛豫则是韧性提高的原因。在外应力作用下，复合材料发生断裂的机理是增强体与基体之间界面的分离，增强体本身的断裂并最终从基体中拔出。显然，决定复合材料性质的不仅是基体和增强体的性质、分布、尺度，而且决定于两种材料结合的界面的完整性。为了最大限度地发挥复合材料的潜力，必须充分保证基体和增强体之间的物理和化学相容性，如两相之间良好的界面化学浸润性；热胀系数匹配，如陶瓷基复合材料的理想匹配是增强体的热胀系数小于基体的热胀系数，以使陶瓷基体处于压应力状态，增强体则处于张应力状态；弹性模量匹配，增强体的弹性模量和应变值大于基体的对应值时，是理想的匹配。

2.2　抗断裂失效材料的选择和设计方法

2.2.1　抗韧断失效材料的选择和设计方法

钢的强度是其抗变形和断裂的能力，韧性则是其在变形和断裂过程中吸收能量的量度。为了防止韧性断裂的发生，需要提高材料的强度和韧性。在不改变成分和生产工艺的情况下，提高强度往往会导致韧性的降低，提高韧性需要牺牲一定的强度。依据钢的强韧性与组织结构之间的关系，强韧化的技术途径是组织结构转化到最利于提高强度及韧性的理想状态，以便使钢的强度和韧性均可有获得大幅度提高的充分可能性。由合金化和冶金技术控制微观组织结构因素的技术途径有以下几方面。

1）细化晶粒和组织能同时提高钢的强度和韧性，故被广泛应用，具体方法为：在合金化方面是以少量强碳化物形成元素等进行微合金化；而在冶金技术方面主要有：①由控制终轧温度和轧后冷却速度等途径来细化奥氏体再结晶晶粒和冷却后的铁素体晶粒；②由循环热处理达到晶粒细化；③快速奥氏体化多级热处理等。

2）改善基体和强化相形态　改善 α 基体韧性的途径为：在合金化方面使用能保持或改善基体韧性的合金元素，如镍和稀土元素等；而在冶金技术方面主要有：①采用炉外精炼和二次重熔技术以减少钢中的气体和杂质元素等，从而净化基体；②通过快速奥氏体化多级热处理工艺改变碳、合金元素以及碳化物相合理而有利的分配，并改善强化相形成，使其呈细小、均匀而弥散的分布。

改善回火马氏体钢中马氏体形态，使其形成位错型条状马氏体。技术途径为：在合金化中，控制碳含量以及合金元素的类型和含量以提高 Ms 点；在冶金技术方面，采用超高温奥氏体（高于常规工艺 50～100℃）淬火等途径以得到位错型条状马氏体，再配以附加热处理以细化奥氏体晶粒。

3）引入韧性相　获得少量奥氏体的途径有：①在合金化设计中要加入适量的形成和稳定奥氏体的元素，如 Ni、Mn 等；②对中碳低合金超高强度钢采用超高温奥氏体化淬火获得含有少量奥氏体的马氏体、奥氏体复合组织；③对于具有奥氏体和马氏体可逆转变的高镍、高锰钢，如 9Ni 钢、Fe-5Mn 等，可在固溶处理得到马氏体后，在可逆转变临界区内回火以获得所需数量的奥氏体。

引入贝氏体的途径主要是等温热处理，一般是在下贝氏体转变区进行等温停留一定时间，从而获得所需数量的下贝氏体后立即淬火，使剩余的奥氏体转变为马氏体。

获得马氏体、铁素体双相组织的途径主要是临界区淬火，依据加热温度在 A_{C1} 和 A_{C3} 临界区位置的高低来控制铁素体的数量，由不同预处理工艺来控制铁素体的形态和尺寸。

4）减少气体杂质、控制夹杂物形态　为减少钢中气体、杂质和非金属夹杂物对强韧性的危害，主要途径是使用现代精炼技术，其中炉外精炼和二次重熔能有效地减少气体、杂质，后者还可改善铸锭结构和质量，而通过喷粉（SiCa 等）或在钢中加入微量稀土元素能有效地控制非金属夹杂物形态，使其成为颗粒状不变形的夹杂物。采用含有稀土氧化物的渣系在电渣重熔过程还原加入稀土的方法，能更充分发挥稀土的有利作用，显著提高其强韧性。

2.2.2　抗脆断失效材料的选择和设计方法

脆性断裂的发生常常引起灾难性的事故，通过合理选择材料和适当的热处理工艺，可以有效防止脆性断裂事故。

由断裂力学知识可知，当临界裂纹尺寸固定时，工作应力越高，断裂韧性值也应相应提高才能满足要求。要在提高强度的同时，又能改善韧性、降低脆性，可从三方面着手，其一改善合金的化学成分和冶炼生产方法，去除或固定对韧性不利的有害因素；其二获得可达到最佳韧性的显微组织和相分布；其三细化显微组织，细化晶粒。抗脆断失效材料的选择和设计时，也应从这三方面着手分析。

1）晶体结构　韧脆转变现象并非所有的铁基合金中都出现，只在部分体心立方晶型的铁基合金中才有，对其他面心立方晶型的材料，脆性转变的现象不是很明显。有些合金含量很高的钢，虽然是体心立方晶型，但含碳量极低，故在低温下仍是韧性断裂。因此，在其他条件允许的情况下尽量避免选用体心立方的低合金钢。

2）化学成分的影响　合金的化学成分是通过不同工艺，形成合金的各种组织结构来反映各种性能的，包括工艺性能和使用性能。钢中合金元素的主要影响如表 5.5-3 所示。

表 5.5-3 钢中合金元素的主要作用

元素	合金元素的作用
碳	钢中主要强化组元。间隙式固溶强化或形成碳化物造成弥散强化。但大颗粒碳化物也可成为裂纹源，若成网状分布时导致脆断。碳量增加时提高脆性转变温度，增0.01%C，约使转变温度提高4℃
镍	主要韧化组元。降低相变温度，增加淬透性，稳定奥氏体
铬	碳化物形成组元。在结构钢中增加淬透性，保证不锈钢抗腐蚀性能，有固溶强化作用
钼	强碳化物形成组元，增加结构钢淬透性，抑制回火脆性，有固溶强化作用
硅	脱氧剂。固溶强化组元，同时提高脆性转变温度，在结构钢淬火回火时提高 ε-碳化物转变温度
锰	脱氧剂。固定硫形成 MnS，可防止 FeS 形成热脆，增加结构钢淬透性，Mn/C 比增高可降低脆性转变温度
钴	在马氏体时效钢中钴有细化时效沉淀相的强化作用
钛	结构钢中加钛形成 TiN 或 TiC，抑制奥氏体晶粒长大。在马氏体时效钢中作为沉淀相元素形成 Ni_3（TiAl）起强化作用，在硼钢中起定氮保硼作用
铝	强脱氧剂。也可形成 AlN，钉扎奥氏体晶界抑制晶粒长大（直到980℃），由于能减少固溶体中 N 量、故可降低脆性转变温度，在马氏体时效钢中作沉淀相元素形成 Ni_3（TiAl）起强化作用

钢的化学成分中的有益元素含量在合金设计中应给予重视，而C、N、H、O、P等有害夹杂含量增加均会提高脆性转变温度，降低冲击韧性，所以应减少其含量。要对脆断事故进行分析，首先看是否含量超标，不超标时也要考虑合金配比是否合适，因为成分落在牌号规范内，但配比不合适（如 Mn、C 比），其工艺性能或使用性能上达不到要求并引起失效的事例是很多的，如在设计钢的成分时应尽可能地控制一些对钢的回火脆性影响较大元素的配比，使钢的回火脆性不致过大，以及向回火脆敏感性较大的钢中添加钼和钨，对回火脆性敏感性较大的铬镍钢、铬锰钢、硅锰钢、铬钒钢等加钼便是如此。

镍被认为是降低钢的脆性的转变温度作用最大的元素。低碳钢的脆性转变温度因镍量增加而降低，淬火后低温回火的镍钢含4.5%镍，高温回火的钢含2.5%镍时，脆性转变温度可降到最低。在合金结构钢中，铜的作用与镍相似。

锰对钢的脆性转变温度的影响因含量不同而异。低碳钢含锰在1.5%以下时可使脆性转变温度降低，所以，降低碳增加锰可以减少钢中的碳化物、细化晶粒，有利于低碳钢获得较高的低温冲击韧性，但含锰高时又使钢的脆性转变温度提高。

此外，钢中偏析、夹杂物、白点、微裂纹等缺陷越多，韧性越低。

3）细化晶粒　细化晶粒是控制材料韧性，避免脆断的重要手段。粗晶粒的钢脆性转变温度较细晶粒的为高，如粗晶粒的中碳钢的脆性转变温度，可较细晶粒钢的高40℃。细晶粒强化符合 Hall-Petch 关系，当晶粒越细，晶界面积就越大，晶界对位错运动的阻碍也越大，从而使强度升高。此外，晶粒越细，在一定体积内的晶粒数目越多，在同样的变形量下，变形分散在更多的晶粒内进行，晶粒内部和晶界附近的应变量相差较小，变形较均匀，相对来说引起应力集中减小，使材料在断裂之前能承受较大的变形量，因而有较好的塑性；又因为晶粒越细，晶界的曲折越多，越不利于裂纹的传播，从而在断裂过程中可吸收更多的能量，表现出较高的韧性；当晶粒细小时，晶界面积增加，又使晶界杂质浓度减少，避免产生沿晶脆性断裂。在铸造生产中可通过加大过冷度，在铸铁中加入硅铁或硅钙合金，在铝合金中加入钛、锆、钒等或在合金钢中加入碳化物形成元素（Ti、V 等），阻止加热时晶粒长大，从而细化晶粒提高韧性。

4）形变热处理　形变热处理是形变强化与热处理强化相结合的一种复合强化工艺。高温形变热处理可细化奥氏体的亚结构，因而细化淬火马氏体，使强度和韧性提高。低温形变热处理除了细化奥氏体亚结构外，还可增加位错密度，促进碳化物弥散沉淀，降低奥氏体含碳量和增加细小板条马氏体的数量，因而提高强度和韧性。形变热处理还可消除钢的回火脆性，即使钢加热至 A_{C3} 温度以上进行变形并立即淬火、回火，这样可使钢的回火脆性消除，并得到纤维状断口。

5）亚温淬火　亚温淬火可以提高低温韧性和抑制高温回火脆性，并显著降低脆性转变温度，这与改善晶粒度密切相关。亚温淬火时固溶温度处在两相区，可以形成很细的奥氏体和未溶铁素体两相组织，铁素体-奥氏体相界面比一般淬火的奥氏体晶界面积大许多倍，因而单位相界面上杂质浓度减少；此外，亚温淬火的未溶铁素体比奥氏体能溶解较多的杂质含量，进一步降低奥氏体晶界的杂质偏聚浓度，因而提高钢的韧性，抑制高温回火脆性。

6）低碳马氏体强化　在相同强度水平下，低碳马氏体（板条马氏体）比高碳马氏体的韧性好得多，即在具有较高强度、硬度的同时，还具有相当高的塑性和韧性。这是由于板条马氏体中含碳量低，可以产生“自回火”，且碳化物分布均匀；其次是板条马氏体的胞状位错亚结构中位错分布不均匀，存在低密度位错区，为位错提供了活动余地，由于位错运动降低局部应力集中，可延缓裂纹萌生而对韧性有利；此外，含碳量低，晶格畸变小，淬火应力小，不存在显微裂纹，裂纹通过马氏体条也不易扩展，因此，低碳马氏体具有很高的强度和良好的韧性，同时还具有脆性转变温度低、缺口敏感性小等优点。通过热处理可以改变马氏体的形态，增加板条马氏体的相对数量，从而可显著提高钢的强韧性，如选用低碳或中碳钢，通过高温加热淬火和低温回火来获得板条马氏体组织，这是一条充分发挥钢材潜力的有效途径。

2.2.3 抗疲劳失效材料的选择和设计方法

1）提高疲劳极限的材料选择和设计　材料的疲劳强度和抗拉强度之间，在一定条件下存在着较密切的关系，因此，在一定条件下凡能提高抗拉强度的因素（如合金化、热处理等），一般来说，均可使疲劳强度提高。钢中加入碳、锰、镍、铬、钼、钒、铜和硼等能提高抗拉强度，也可使钢的疲劳强度提高。但是，研究表明，无论是碳钢还是合金钢，上述元素中对疲劳强度影响最大的是碳，钢的疲劳极限随钢中碳含量的增加而升高。在结构钢中，合金元素所引起的主要作用是改变钢材的淬透性和马氏体点，以控制热处理后的组织状态，对中等强度的结构钢，在钢材有足够淬透性的条件下，合金元素一般对疲劳极限无显著影响。

热处理对材料疲劳强度的影响比材料成分的影响要大得多，同一成分的材料，由于热处理不同，疲劳强度可在相当大的范围内变化。热处理对疲劳强度的影响，实质上就是显微组织的影响，钢材经过不同的热处理，虽然可以得到相同的静强度，但由于组织不同，却有着不同的疲劳强度。例如，片状珠光体的共析钢和粒状珠光体的共析钢，在相同的强度水平时，由于碳化物形态的差异，可导致疲劳极限的明显不同。

同一种钢通过等温淬火得到的下贝氏体组织的疲劳强度比同等强度的淬火回火组织为高，虽然铁素体和珠光体型的组织，持久比（σ_{-1}/σ_b）较马氏体高，但由于马氏体本身的σ_b较其他组织高许多，因此马氏体组织仍可获得高的疲劳强度。

复合组织的存在会使钢材的疲劳强度降低，一般认为其原因是所谓“冶金缺口”的作用，它可能是粗大珠光体、自由铁素体、合金元素偏析或淬火时残留下来的奥氏体所引起。

细化晶粒可以提高疲劳极限，因为晶界是裂纹扩展的一种障碍。组织均匀性好的回火马氏体和回火屈氏体，其疲劳极限较高，但淬火组织中的未溶铁素体和过多的残余奥氏体等非马氏体组织将使疲劳极限下降。

冶炼、浇注及成材过程中可能引起化学成分不均、夹杂物、气体、疏松及各种缺陷组织等材质问题，其中夹杂物对疲劳性能的影响最大。钢中夹杂物使疲劳性能明显下降，材料的强度水平越高，夹杂物的影响越大，它不仅可以作为疲劳裂纹的核心，还会使裂纹扩展的速率增快。减少夹杂物的含量能够明显改善疲劳性能，因此，对于重要零件的用材，在一般熔炼基础上，还应采用真空熔炼、真空除气甚至电渣重熔等工艺，以降低材料的夹杂物含量。

2）延缓疲劳裂纹萌生的材料选择和设计　疲劳破坏是在交变载荷的反复作用下产生的，疲劳破坏过程经历了疲劳裂纹的萌生和扩展阶段。

疲劳裂纹的形成时包括材料循环硬化/循环软化和裂纹萌生这样二个阶段。在循环硬化/循环软化的过程中，要材料不发生软化和有很好的稳定性，宜使静载硬化指数$n \geqslant 0.1$。但形变硬化、沉淀强化、马氏体相变强化的高强材料通常都会出现疲劳软化，其中基体变形呈波状滑移（即容易滑移）时尤其如此。所以，在提高静强度的同时，要保证在循环加载时不失去强度，宜遵循以下原则：①对形变硬化材料，应采用层错能较低、呈平直滑移型的基体；②对沉淀硬化材料，沉淀相应为稳定相，不会重新溶入基体，或有意加入弥散质点以阻止长程的位错运动；③如果基体的层错能高，是波状滑移型材料，则应采用弥散强化或纤维增强，当然，这类措施对平直型滑移的基体材料也很有效。

疲劳断裂常常是从零部件表面的某些缺陷造成的应力集中处开始，因而材料的疲劳极限不仅对材料的组织结构敏感，而且对工作条件、加工处理条件等外因也很敏感。因此，能引起任何表面组织结构的变化的表面处理方法都会影响到疲劳裂纹形成的抗力。这种表面处理方法大致有三类：①机械处理，如喷丸、冷滚压、研磨和抛光；②热处理，如火焰和感应加热淬火；③渗、镀处理，如氮化和电镀等。

3）降低裂纹扩展速率的材料选择和设计　近门槛区的裂纹扩展速率主要决定于ΔK_{th}值，在相同的ΔK下，提高ΔK_{th}值，使裂纹扩展速率大大降低。提高ΔK_{th}值的方法如下。

① 晶粒粗化，ΔK_{th}值提高，但晶粒粗化会大大降低屈服强度。

② 亚共析钢中的珠光体增加，裂纹扩展门槛值下降。

③ 适当增加钢中的残留奥氏体量对提高ΔK_{th}值有利。

当裂纹扩展，特别是进入Ⅱ区时，裂纹扩展速率主要取决于裂纹扩展系数B，而B值取决于材料的性能和裂纹在Ⅱ区的扩展机制。当高碳钢具有粒状珠光体组织时，裂纹以条带机制扩展，因而裂纹扩展速率低。当高碳钢具有片状珠光体时，裂纹以微区机制扩展，而高碳钢的塑性低，故B值大，裂纹扩展速率大大提高。

由此可见，钢的显微组织变化引起了裂纹的Ⅱ区扩展机制的改变，从而改变了B值，则会大大地影响裂纹扩展速率。若显微组织的变化不引起Ⅱ区裂纹扩展机制的改变，则对裂纹扩展速率无明显的影响。要降低裂纹的Ⅱ区的扩展速率，应控制显微组织使裂纹在Ⅱ区以条带机制扩展。通常第二相呈粒状，不破坏韧性相的连续性，裂纹可能以条带机制扩展。若裂纹以非条带机制扩展，提高的强度和塑性可降低裂纹在Ⅱ区的扩展速率。

2.2.4　抗高温断裂失效材料的选择和设计方法

由金属材料的蠕变变形和断裂机理可知，提高蠕变抗力的关键在于：一方面使受阻位错不容易由交滑移和攀移得到松弛；另一方面则要大力强化晶界。

（1）合金化和晶体结构

在温度较低（在等强温度以下）、变形速度大的条件下，金属变形以滑移机制为主，这时对形变抗力和强化因素的考虑与室温相同。但在使用温度提高、应变速率小的情况下，金属内部的扩散常常控制着蠕变。蠕变过程的激活能等于原子自扩散的激活能，因此金属的熔点越高，晶体结构越紧密，其高温性能越好。如T_S表示工作温度，T_M表示材料熔点，则体心立方铁素体耐热钢$T_S/T_M \approx 0.5$；面心立方的奥氏体耐热钢$T_S/T_M \approx 0.6$；而镍基超合金$T_S/T_M \approx 0.7$；高温陶瓷Al_2O_3、SiC、Si_3N等则具有更高的耐高温潜力。

在基体金属中加入铬、钼、钨、铌等元素，可以显著提高其抗蠕变性能，合金元素的强化方式有两种：即固溶强化和弥散粒子强化。

1）固溶强化　大多数合金元素在固溶状态都能增强抗蠕变能力，即减小蠕变速率和延长蠕变断裂时间。在高温蠕变中，溶质原子的作用主要在于对回复过程的影响，从而降低了蠕变速率，溶质原子对位错攀移的影响有两种可能的机制：合金原子的加入使层错能降低、位错变宽，因而使攀移困难。例如，镍铬合金的层错能比纯镍低，抗变能力就比较高。在镍铬合金中添加钴，使层错能又进一步降低，抗蠕变能力也进一步提高；溶质原子与溶剂原子的结合力使扩散激活能增高，因而降低了攀移速率。溶质元素的熔点越高，效应就越显著。当溶质原子含量较低时，不可能显著改变整个基体的扩展系数。但溶质原子优先附着于位错的割阶上，因而能够显著降低攀移速率。少量合金原子对回复的显著影响，可能就是这种机制。

2）弥散粒子强化　实际使用的耐热合金，大都含有高度弥散的第二相粒子。这种强化方法比固溶强化更为有效。试验证明这种强化作用可以维持到$0.7T_M$，而固溶强化在$0.6T_M$以上就显著减弱。第二相粒子的弥散度越高，而且稳定性越高，则强化效应越好。高度弥散的第二相粒子，可以有效地阻碍晶内位错的滑移和攀移以及晶界的滑动，因而强化效果好，弥散粒子的稳定性决定于弥散相与基体间的界面能，以及弥散相在基体中的溶解度。弥散相与基体间的界面能越低，则稳定性越高。因为界面能是粒子长大的推动力，界面能低，粒子就不容易长大。弥散相在基体中的溶解度越低，则粒子长大或再溶的过程越慢，稳定性也越高。有人观察到铬钼耐热钢在蠕变后沉淀粒子周围的位错乱网。由于乱网中位错密度很高造成强烈的应变硬化，同时粒子又阻碍这

些位错的滑移与攀移，因而不易产生回复软化。这是弥散粒子提高耐热强度的一个重要原因。有些沉淀过程在变形过程中发生，沉淀相多在位错上成核，这样的沉淀相对位错的阻碍特别有效。

(2) 晶粒度和晶界结构

由于一般耐热合金的正常使用温度大致都在等强温度以上，所以晶界滑动对蠕变的贡献将有较大增长，采用粗晶对提高蠕变极限和持久强度都很有利。不过，晶粒也不可过大，以免使持久塑性和冲击韧性下降。所以对奥氏体耐热钢和镍基合金，一般以2～4级晶粒度为好。由于细晶对蠕变强度不利，所以晶粒度不均匀也是应当避免的。此外，由于垂直于拉应力的晶界通常是空洞和裂纹的形核位置，所以采用定向凝固工艺使柱状晶沿受力方向生长，减少横向晶界，就可以大大提高构件的持久强度。目前，这种工艺已在涡轮叶片上得到很好的应力。

杂质元素对高温合金的晶界很敏感，改进冶金质量，能大大提高蠕变极限和持久强度。由于有害杂质元素S、P及Pb、Sn、Bi、Sb等，即使其含量甚微（十分之几），当其在晶界集聚后，会导致晶界严重弱化，使高温性能急剧降低。因此冶炼中要尽量减少有害杂质。除此之外，还应当减少非金属夹杂物和冶金缺陷（例如气孔），因为它们也严重降低材料的高温性能。冶炼时在合金中添加适量的硼和稀土元素能增加晶界扩散激活能，既能阻碍晶界滑动，又增大形成晶界裂纹的表面能，因而可以提高蠕变极限和持久强度。

(3) 热处理工艺的影响

珠光体耐热钢一般采用正火加高温回火工艺。正火温度应较高，以促使碳化物较充分而均匀地溶于奥氏体中。回火温度应高于使用温度100～150℃，以提高其在使用温度下的组织稳定性。

奥氏体耐热钢或合金一般进行固溶处理和时效，使之得到适当的晶粒度，并改善强化相的分布状态。有的合金在固溶处理后再进行一次中间处理（二次固溶处理或中间时效），使碳化物沿晶界呈断续链状析出，则可使持久强度和持久塑性进一步提高。

采用形变热处理改变晶界形状（形成锯齿状），并在晶内造成多边化的亚晶，使新生成位错攀移困难，则可使合金进一步强化。如GH2038、GH78型铁基合金采用高温形变热处理后，在550℃和630℃的100 h持久强度分别提高25%和20%左右，而且还保持有较高的持久塑性。

2.3 抗腐蚀失效材料的选择和设计方法

2.3.1 抗应力腐蚀失效材料的选择和设计方法

应力腐蚀断裂事先没有明显征兆，易造成突发性灾难后果，工程上最易产生应力腐蚀的金属材料是奥氏体不锈钢、低合金钢、铜合金、高强度铝合金等。最常见的应力腐蚀现象有碱脆、氯脆和硝脆等。

抗应力腐蚀失效材料的选择与设计主要考虑材料的化学成分、组织和性能的影响。

(1) 化学成分与应力腐蚀的关系

一般而言，纯金属不发生应力腐蚀，但含有杂质时也可能产生应力腐蚀。例如，质量分数为99.999%的Cu在$Cu(NH_3)_2^{5+}$溶液中也会产生应力腐蚀。

低碳结构钢价格低、用途广、用量大，因而发现的应力腐蚀断裂问题也多，在浓碱、硝酸盐、硫化氢、氰化物等的水溶液中以及氨水和无水的液氨等介质中不断发生开裂或断裂事故。低碳结构钢含碳量约低于0.25%，大多在正火或退火后使用，其金相组织为铁素体基体中有各种形貌的渗碳体。碱脆是火力发电厂最有害的断裂事故。锅炉用水一般用碱软化，虽然含碱量甚微，但在铆接处NaOH含量可从原有的10^{-6}数量级富集到20%以上。NaOH的浓度越高，则碱脆的敏感性越大。

在生产氮肥及硝酸盐的工厂，低碳钢在浓硝酸盐中使用时会出现应力腐蚀开裂或断裂的现象，这种现象叫做“硝酸盐脆化”，简称“硝脆”

虽然硝脆与碱脆的化学反应不同，但是它们的机理却是一样的，都是属于沿晶溶解型，对于这两种应力腐蚀，建议使用碳量约为0.20%的镇静钢（20 g）。

(2) 组织与应力腐蚀的关系

材料的组织状态与应力腐蚀关系研究的还不够多，一般认为，接近平衡状态的组织具有较高的抗应力腐蚀能力。碳钢及低合金钢化学成分一定时，各种组织对应力腐蚀敏感性由大到小排列顺序为，淬火马氏体、贝氏体、回火马氏体、回火贝氏体、托氏体、索氏体、正火珠光体、退火珠光体。从碳化物形态分析，碳化物越圆整，抗应力腐蚀能力越强。

碱脆研究结果表明，钢的晶粒大小对断裂应力的影响符合Hall-Petch关系。

(3) 材料的强度与应力腐蚀的关系

许多研究指出，钢的屈服强度和抗拉强度对应力腐蚀敏感性有重要影响。一些研究者提出了σ_s与材料的K_{ISCC}关系式。4340钢、D_6AC钢、H－11钢等在蒸馏水、盐水等介质中σ_s与材料的K_{ISCC}关系式为

$$K_{ISCC}=\frac{RT}{2V_H}\ln\left(\frac{C_{cr}}{C_0}\right)-\frac{\sigma_s}{2a} \quad (5.5\text{-}7)$$

式中，R为气体常数；T为热力学温度；V_H为氢摩尔体积；C_0为开始平均氢浓度；C_{cr}为临界氢浓度；σ_s为钢的屈服强度；a为常数。

从上式可知，随σ_s提高钢的K_{ISCC}降低。

通常都认为，决定钢的应力腐蚀特性的最重要因素是钢的强度水平，强度越高抗应力腐蚀能力急剧下降，特别是当$\sigma_s>1\,400$ MPa时，应力腐蚀敏感性显著增加。为了防止发生应力腐蚀，应当将在湿H_2S等介质中服役的低合金钢零件的硬度控制在22HRC以下，可显著降低应力腐蚀敏感性。

易引起金属材料产生应力腐蚀的介质见表5.5-4。

表5.5-4 易引起金属材料产生应力腐蚀的介质

金属材料	腐蚀介质
碳素钢、低合金钢	NaOH水溶液，硝酸盐水溶液，HCl水溶液，$CO+CO_2+H_2O$，含$CO_2+HCN+H_2S+NH_3$液体，含NH_3溶液，H_2S水溶液（高强钢），$H_2SO_4+HNO_3$混合酸
奥氏体不锈钢	含氯化物水溶液海水，高温水，苛性碱水溶液，H_2SO_4+NaCl，HCl，$H_2SO_4+CuSO_4$
马氏体不锈钢	海水，NaCl水溶液，$NaCl+H_2O_2$水溶液，NaOH水溶液，含NH_3水溶液，$HNO_3+H_2SO_4$，$H_2SO_4+HNO_3$水溶液，H_2S水溶液，高温高压水，高温碱水
Au-Cu-Ag，Ag-Pt	$FeCl_3$水溶液
Mg-Sn，Mg-Al	$NaCl+K_2CrO_4$水溶液，Na_2SO_4或$NaCl+K_2CrO_4$水溶液
Mg-Al-Zn-Mn	水
Al-Zn	空气，$NaCl+H_2O_2$水溶液

续表 5.5-4

金属材料	腐蚀介质
Al-Mg Al-Mg，Al-Cu-Mg，Al-Mg-Zn，Al-Zn-Cu Al-Zn-Mg-Mn，Al-Zn-Mg-Cu-Mn	NaCl 水溶液 海水，NaCl 水溶液，NaCl + H_2O_2 水溶液 海水
Cu-Al Cu-Zn，Cu-Zn-Sn，Cu-Zn-Pb Cu-Zn-Ni，Cu-Sn Cu-Sn-P，Cu-As	NH_3，水蒸气 NH_3 蒸气，氨水溶液 NH_3，$NH_3 + CO_2$ 空气
Cu-Zn-Si，Cu-Zn-Sn-Mn	水，水蒸气
蒙乃尔合金	沸腾 75%NaOH 水溶液，水蒸气（699 K 以上），HF，汞，有机氯化物，氟化氢
Ni（质量分数 99%）	NaOH 及 KOH（熔态及溶液），水蒸气（>699 K 以上），HF，S（533 K 以上）
Ti 及 Ti 合金	发烟 HNO_3，硫酸铀，HCl，熔盐（NaCl），海水，NaCl 水溶液，三氯乙烯
Cr-Ni-Fe 合金	HF，NaOH 水溶液（533～699 K），浓缩炉水（533～699 K），水蒸气 + SO_2，浓 Na_2S 水溶液
Pb	Pb $(CH_3COO)_2 + HNO_3$，大气，土壤

2.3.2　抗氢脆材料的选择和设计方法

由氢脆的机理可知，氢和应力的联合作用是导致氢脆发生的根本原因，因此，减少氢进入钢中的数量，如在浇注时增加烘干铸模的工序，以防止水气的进入，或者采用真空浇注或真空处理的方式。另外，在焊接和电镀时采用低氢焊接、低氢电镀、体系进行电化学保护，提高电位，使之在析氢电位以上，对防止氢脆都有明显的效果。

在零件设计时尽量降低设计应力，减少应力集中程度，消除残余应力，使局部应力小于氢脆临界应力（$\sigma < \sigma_{th}$ 或 $K_I < K_{th}$）。

在材料选择时，应采用氢脆敏感性低的材料。钢中加入一些合金元素可防止氢渗入或降低氢扩散能力。实验证明，体心立方金属氢脆敏感性大于面心立方金属。

氢脆受控于材料捕获的氢量 C_T 与引起此类缺陷开裂的临界氢浓度 C_{Cr} 之间的相对大小。当 $C_T < C_{Cr}$ 时，不会引起氢致开裂；当 $C_T = C_{Cr}$ 时，开始氢致开裂；当 $C_T > C_{Cr}$ 时，裂纹扩展。因此，减少氢脆敏感性的途径在于提高 C_{Cr} 和降低 C_T。

降低 C_T 的途径，减少内氢和限制外氢的进入两方面。减少内氢可通过改进冶炼、热处理、焊接、电镀、酸洗等工艺条件，以减少氢进入金属。对含氢材料也可进行脱氢处理。例如，高强钢可用疏孔镀镉或镀 Cd-Ti 合金，镀后在 180～200℃以下烘烤除氢 19～24 h 的方法消除氢脆。

另外，可通过添加陷阱使它们均摊一些氢量，以达到降低 C_T 的目的。但这些陷阱必须本身具有较高的 C_{Cr}，否则先在这些地方引发裂纹。其添加陷阱数量要足够多，且具有不可逆陷阱的作用，而且均匀分布在金属基体中。能满足这些条件陷阱很多，例如原子级尺寸的陷阱（以溶质原子形式存在）有 Sc、La、Ca、Ta、K、Nd、Hf 等，碳化物和氮化物形成元素（以化合物形式存在），如 Ti、V、Zr、Nb、Al、B、Th 等。但应注意，加入这些元素还应综合考虑对其他性能的影响。

限制外氢的进入主要从建立障碍和降低外氢的活性两方面入手。

利用物理、化学、电化学、冶金等方法在上施以镀层，镀层应具有低的氢扩散性和溶解度，从而构成氢进入金属的直接障碍。例如 Cu、Mo、Al、Ag、Au、W 等金属覆盖层，或经表面热处理而生成致密的氧化膜。有时可涂覆有机涂料或衬上橡皮或塑料衬里，防止与氢或致氢介质接触，起到隔离作用。

通过加入某些合金元素，延缓腐蚀反应或生成的腐蚀产物可抑制氢进入基体，可起到间接障碍的作用。例如含 Cu 钢在 H_2S 水介质中生成 Cu_2S 致密产物，可降低氢诱发开裂倾向。

降低外氢的活性，如在气相的 H_2S、H_2 中加入适量的氧，在腐蚀介质中加入适当的缓蚀剂抑制阴极析氢或者加入促进氢原子复合成氢分子的物质，都可减少外氢的危害。

存在外氢时，特别是在足够的应力作用下，终将使某类最有害缺陷的 C_T 达到其 C_{Cr}。因此，提高 C_{Cr}，即控制对 C_{Cr} 起作用的一些参数，对减少氢脆敏感性是很重要的。例如降低杂质含量，As、P、Sn、Bi、Se 等杂质在晶界的偏析；细化晶粒；控制有害夹杂物（如碳化物、氧化物）和碳化物的类型、数量、形状、尺寸和分布；适当的变形处理等可以提高临界氢浓度 C_{Cr}，从而降低氢脆敏感性。

在常用热处理操作中，经处理的钢抗氢蚀性最好，正火 + 回火状态次之，退火钢的抗氢蚀性能最差。这是因为处理可获回火索氏体，这种组织中的碳化物呈球状，表面能较低，比较稳定，故可提高抗氢性。

2.3.3　抗腐蚀疲劳材料的选择和设计方法

由于腐蚀作用，使疲劳裂纹萌生所需时间及循环周次都有所减少，使裂纹扩展速度增大。腐蚀疲劳的特点如下。

1）在空气中的疲劳存在着疲劳极限的碳钢及低、中强度钢在腐蚀疲劳的情况下已不存在疲劳极限。腐蚀疲劳是交变应力与腐蚀环境共同作用的结果。

2）与应力腐蚀不同，纯金属也会发生腐蚀疲劳，而且发生腐蚀疲劳不需要材料 - 环境的特殊组合，只要存在腐蚀介质，在交变应力作用下就会发生。金属在这种介质中可以处于钝态，也可以处于活化态。

3）金属的腐蚀疲劳强度与其耐蚀性有关。耐蚀材料的腐蚀疲劳强度随抗拉强度提高而提高；耐蚀性差的材料尽管它的疲劳极限与抗拉强度有关，但在海水、淡水中的腐蚀疲劳强度与抗拉强度无关。

4）腐蚀疲劳裂纹多起源于表面腐蚀坑或表面缺陷，裂纹源往往数量较多。腐蚀疲劳裂纹主要是穿晶的，但也可能出现沿晶的或混合的，并随腐蚀发展裂纹变宽。

5）腐蚀疲劳断裂属于脆性断裂，没有明显的宏观塑性变形。断口既有腐蚀的特征，又有疲劳特征，但由于腐蚀的作用而比较模糊。

合理选材是提高零件或结构抗腐蚀疲劳的一个重要因素，因此，在材料选择时应考虑材料耐蚀性、组织结构和表面状态等。

1）耐蚀性较高的金属，如钛、铜及其合金、不锈钢等，对腐蚀疲劳的敏感性较小；耐蚀性较差的金属，如高强铝合金、镁合金等，对腐蚀疲劳的敏感性较大。因而只是那些改善材料耐蚀性的合金化才是有益。对不锈钢来说，较好的耐点蚀性能使其在海水中的耐腐蚀疲劳性能也较好，故增加 Cr、Ni、Mo 等元素含量是有益的。

2）碳钢、低合金钢提高强度的热处理可以提高疲劳极限，但对腐蚀疲劳行为的影响较小，甚至有降低腐蚀疲劳强度的倾向。对不锈钢来说，某些提高强度的处理可以提高腐蚀疲劳强度，敏化处理则是有害的。细化晶粒可以提高钢在

空气中的疲劳强度，对腐蚀疲劳也可起到类似的效果。

3）表面残余应力为压应力时，耐腐蚀疲劳性能较表面残余拉应力为好；施加保护涂层可以改善材料的腐蚀疲劳性能。

2.4 抗磨损失效材料的选择和设计方法

磨损是一种相当复杂的过程。它是机械零部件的三种主要破坏形式之一。磨损的分类方法很多，按表面接触性质可分为金属－磨料磨损、金属－金属磨损和金属－液体磨损；按环境和介质分类分为干磨损、湿磨损和液体磨损；按照磨损的破坏机理进行分类，一般可分为：黏着磨损、磨粒磨损、腐蚀磨损和表面疲劳磨损。也有人认为按破坏机理分为黏着磨损、磨料磨损、腐蚀磨损、接触疲劳磨损、冲蚀磨损、微动磨损和冲击磨制等七类。

2.4.1 影响磨损的主要因素

（1）影响黏着磨损的主要因素

黏着磨损是一种主要的磨损形式，影响黏着磨损的主要因素有：

1）材料的塑性大小。一般脆性材料的抗黏着磨损能力比塑性材料高。塑性材料的黏着破坏发生在离表面一定深度处，磨屑较大，而脆性材料的黏着破坏主要是剥落，磨屑深度浅也易脱落。根据强度理论：脆性材料破坏由正应力引起，而塑性材料的破坏决定于切应力，最大正应力作用在表面，最大切应力却出现在离表面一定深度。材料的塑性越好，加工硬化强烈，最后剪断的位置离黏着结合点越远，表现出的黏着磨损越严重。因此，生产上要注意一对摩擦副的配对。不要用淬火钢（HRC60）与软钢（HRC20）配对；不要用软金属－软金属配对。选用两个高硬度的淬火钢配对。或淬硬钢－灰铸铁配对，会取得良好的效果。

2）金属性质越是相近的，构成摩擦副时黏着磨损也越严重。反之，金属间互溶程度越小，晶体结构不同，原子尺寸差别较大，形成化合物倾化较大的金属，构成摩擦副时黏着磨损就较轻微。滑动轴承就是这样的例子，选用淬火钢轴与锡基或铝基轴与轴瓦配对。在受力较小时，选用金属与塑料配对能减小黏着磨损。

3）通过表面化学热处理，如渗硫、硫氮共渗、软氮化等热处理工艺，使表面生成一层化合物薄膜，或为硫化物、含氮的化合物，使摩擦因数减小，起到减摩作用也减小黏着磨损。

4）改善润滑条件，如在润滑油中添加极压剂。极压剂是含氯硫磷的有机化合物，例如氯化石蜡、二烷基二硫代磷酸锌等，在高温高压条件下（如当摩擦副过载造成过热时），极压剂分解出活性元素与金属表面生成一层低抗剪强度的金属化合物膜，可减小黏着磨损。

5）黏着磨损严重时表现为胶合。出现胶合时，原光滑表面程度剧烈增加，磨痕很深达 0.2 mm 左右，摩擦表面温度很高，摩擦因数也急剧升高。胶合磨损出现在高速重载和润滑不良的情况下，在齿轮、蜗轮蜗杆、滚动和滑动轴承中都可见到这种失效形式。现对齿轮的胶合研究最多，例如Almen曾对美国通用汽车公司生产的汽车后桥圆锥齿轮的胶合失效情况进行统计，提出防止胶合磨损的经验公式

$$p_0 v_S \leqslant C \tag{5.5-8}$$

式中，p_0 为最大接触应力；v_S 为相对滑动速度；C 为实验常数。

这就是说，当接触压应力和滑动速度乘积小于某一数值时，可不发生胶合。因此，为避免胶合，生产上多采用限制表面压力和滑动速度的办法。

（2）影响磨粒磨损的因素

1）磨料的硬度、大小及形状，磨粒的韧性、压碎强度等。

2）外界载荷大小、滑动距离及滑动速度。

3）材料自身的硬度及内部组织。

对钢来说，基体组织对耐磨性的影响，依以下顺序递增：铁素体、珠光体、贝氏体和马氏体。由于马氏体可获得最高硬度，故耐磨性最高。但在硬度相同的条件下，下贝氏体较回火马氏体更耐磨，这在工具钢及球墨铸铁中均得到证实。虽然等温淬火获得下贝氏体时常伴随着一定数量的残留奥氏体，适量的残留奥氏体对耐磨性确有良好的作用，但在残留奥氏体甚少的情况下，下贝氏体也能显示有较高的耐磨性，因此这只能归结于下贝氏体固有的组织特征。像犁铧、农业中耕机的刮刀刮板，都采用等温淬火。

在马氏体基体中，适当增加碳化物的数量对耐磨性是有好处的。但在低合金钢中，碳化物均为渗碳体型 M_3C，其硬度较低约 950HV；如钢中含有 12%以上的铬，可使碳化物改变为 M_7C_3，硬度约 1 500HV，耐磨性大大提高；如在钢中加入适当的钒，可形成碳化物 MC，硬度则更高约 200HV。可使耐磨性进一步提高。

有色金属中第二相对耐磨性的影响，大体上分为两种情况：如析出的第二相与基体保持共格，第二相可被位错切过，虽然这种沉淀强化能显著提高材料的屈服强度和硬度，但其耐磨性并不比原来单相过饱和固溶体好多少；能提高耐磨性的第二相，是尺寸较大的，与基体已没有共格关系，被位错绕过的颗粒。

具有单相奥氏体组织的高锰钢，被认为是抵抗凿削磨损的理想材料。它有很高的加工硬化能力，特别适用于冲击载荷工况。现今颚式破碎机的齿板、挖掘机斗齿、钢轨道叉等都普遍用高锰钢。高碳和高锰量使奥氏体十分稳定，不形成珠光体和贝氏体，即使冷至低温或经过冷变形，也不转变成马氏体。水淬后得到的硬度只有 230HB 左右，但一经承受冲击磨损，其加工硬化速率急剧提高，在表面 1～2 mm 深度内，硬度迅速达到 500HB。对凿削磨损的试验表明，抗凿削磨损的能力和破碎机齿板的表面硬度有关，当表面硬度能达到 500HB 或 52HRC 以上时，磨损速率急剧降低。

（3）影响接触疲劳的因素

接触疲劳首先取决于加载条件，特别是载荷大小。此外还与许多其他因素有关。

1）非金属夹杂物　轴承钢中的非金属夹杂物有塑性的、脆性的和不变形（球状）夹杂物三种类型。其中塑性夹杂物影响较小，球状夹杂物（钙硅酸盐和铁锰酸盐）次之，危害最大的是脆性夹杂物，（氧化物 Al_2O_3、氮化物、硅酸盐和氰化物等）。因为它们无塑性，和基体的弹性模量不同，容易在和基体交界处引起高度应力集中，导致疲劳裂纹早期形成。研究表明，这类夹杂物的数量越多，接触疲劳寿命下降得越大。

分析夹杂物周围的应力场表明，夹杂与基体间膨胀系数的差别是影响疲劳强度的重要因素。膨胀系数小于基体的，淬火后界面产生拉应力，降低疲劳强度，氧化物即属于此。膨胀系数大于基体的，如硫化物，淬火后界面不会产生拉应力，因此对疲劳强度无害，甚至有利。硫化物的无害有利作用还有可能是由于将氧化物夹杂包住形成共生夹杂物，从而减轻氧化物的坏作用所致。改善钢的冶炼方法，进行净化处理，是减少夹杂物的根本措施，其中电渣重熔和真空电弧炉效果最好。

2）马氏体含碳量　根据在轴承钢上已进行的研究工作表明，在剩余碳化物相同的条件下，马氏体含碳量在 0.4%～0.5%左右时接触疲劳寿命最高，出现峰值。这可能是由于马氏体含碳量太低时，马氏体本身强度（主要指抗剪强度）低，而马氏体含碳量过高时又会引起马氏体脆化所致。

3）剩余碳化物大小和数量　将马氏体含碳量固定在0.5%附近，改变轴承钢中剩余碳化物颗粒大小，研究其对接触疲劳寿命的试验工作指出，细颗粒的碳化物（平均大小在0.5～1.0 μm）的寿命比粗颗粒碳化物（1.4 μm以上，一般在2.5～3.5 μm）的高。当然，碳化物颗粒和接触疲劳寿命不可能只是平均颗粒大小的问题，显然还和碳化物的数量、形状和分布有关。当然分布越均匀越好，形状的圆整度越高越好。至于碳化物数量问题，实验指出，如果不是为了提高耐磨性，最好是不要有剩余碳化物，因为实验观察到裂纹都是在碳化物和马氏体界面上传播。至少也要使剩余碳化物数量调整在6%以下，否则对接触疲劳无好处。例如对高碳钒钢采取措施，使其未熔碳化物作到颗粒小、数量少、均匀、圆整度好，即达到“小、少、匀、圆”的要求后，多次冲击接触疲劳会得到显著改善。

4）硬度　一般来说，在中低硬度范围内，零件的表面硬度越高，接触疲劳抗力越大。但是近十余年来的研究表明，在高硬度范围内，这种对应关系并不存在。近年来国内在轴承钢的大量试验统计工作中也证实了这一点。在20CrMo钢渗碳淬火后，调整回火温度，从而得到不同表面硬度，进行多次冲击接触疲劳试验时也证实了这一点。因为硬度主要反映材料塑变抗力高低和一定程度上反映材料切断抗力的大小。如果说接触疲劳裂纹的生成主要取决于材料塑变抗力和切断抗力（即抗剪强度），那么接触疲劳裂纹的发展则除抗剪强度外，还与材料的正断抗力有关。而材料成分组织变化引起正断抗力的变化在硬度值上是反映不出来的。

2.4.2　耐磨损材料成份选择和设计方法

影响材料耐磨性的因素很多，相互间的关系也很复杂。要提高材料的耐磨性主要可以从三个方面入手。

1）改善摩擦副之间的接触状态。使用润滑剂使在摩擦副表面形成润滑介质膜、表面吸附膜或表面反应膜，可以达到减轻磨损的目的；严格控制摩擦副表面几何形貌，也可以减少磨损。

2）选择适当的材料配合。

3）对摩擦副表面进行表面强化，提高耐磨性的效果十分显著。

评定磨损行为时常用磨损量、相对磨损量、磨损系数和耐磨性等参数。

磨损量分为长度磨损量、体积磨损量和质量磨损量，其中质量磨损量比较容易测量，然后再换算成体积磨损量。

相对磨损量可用磨损速度和磨损率表示，磨损速度指磨损量与磨损时间之比，磨损率指单位磨损距离的磨损量。

材料耐磨性常用磨损率的倒数来表示：

$$\varepsilon = \frac{1}{W} \tag{5.5-9}$$

式中，ε 为耐磨性；W 为单位时间或单位运行距离下产生的磨损量或磨损体积。

磨损系数表示为

$$K = \frac{磨损体积 \times 硬度}{载荷 \times 距离}$$

或

$$K = \frac{磨损体积}{载荷 \times 距离} \tag{5.5-10}$$

由于耐磨性不是材料的固有特性而是磨损工况等有关的系统属性，故很难利用某种试验中测定出的耐磨性来指导具体工程项目的设计和选材。因此，评定耐磨性时均采用一种“标准”材料作为参考试样，在相同磨损条件下进行测试，取其耐磨性比值定义为相对耐磨性，即

$$\varepsilon_{相} = \varepsilon_{试}/\varepsilon_{标} = W_{标}/W_{试} \tag{5.5-11}$$

式中，W 为磨损率；“试”及“标”分别指试样及标件。

虽然有上百种摩擦磨损试验设备可供使用，但正确地选择试验方法却是获得使用可信数据的关键。首先应根据材料的使用工况来选择磨损试验装置以便尽可能模拟出具体工件磨损失效的模式。要想在较短的试验时间内测得较准确的磨损率，除对运行参数如载荷、速度、接触应力等在不改变磨损机制的前提下做一定的强化外，重要的是选定磨损量的测定方法。可供选用的方法有失重法、外形尺寸变化和形貌测定法、刻痕测定法及放射性同位素测定法。

（1）选用耐磨材料

专为抗磨损而发展的钢有轴承钢（1.5%Cr及1.0%C），高锰钢（>1%C，11%～14%Mn）。前者是因其接触疲劳强度高而用于滚珠轴承，因此应严格控制其冶炼工艺以降低甚至消除其非金属夹杂物并通过热处理以保证高硬度和质地均匀；后者则具有强冲击载荷下出现的加工硬化的能力，以适应矿石破碎、岩石挖掘等机械装备中某些经受凿削式磨粒磨损的工况。这两个钢种具有的共同特点是含碳量较高而且还分别含有易形成硬质碳化物的元素（如铬等），以及能扩大钢的奥氏体区并能控制其稳定性和马氏体转变点的锰。这样便可以提高合金硬度或形变强化能力，从而得到优良的抗疲劳和冲击磨损能力。工具钢是另一类耐磨钢，它们都含有硬质碳化物并嵌在强韧基体组织中。这类钢的含碳和合金元素较高，经过热处理足以提高它们的硬度，特别是红硬性和强韧性。

工作于高温及腐蚀环境中的零件多选用Ni-Cr系列奥氏体不锈钢，如1Cr18Ni9Ti等，但它们易出现黏着磨损而造成严重擦伤，为此而发展出沉淀硬化型及马氏体系列的不锈钢。这些钢可以用热处理方法淬硬，其中应用最广的是马氏体不锈钢，它适宜外科工具和330℃以下工作的滚动轴承等。

另一类铁基耐磨合金是各类铸铁，其中以专为各种磨粒磨损工况而发展出的合金白口铸铁最引人注目，如各类介质泵用的镍硬铸铁和高铬白口铸铁。它们的显微组织中有大量的初生碳化物，当Cr含量>8%后，在$(FeCr)_3C$存在的同时还有不少$(FeCr)_7C_3$，这时其耐磨性有较大的提高。由于其硬度高、质地脆，难以车削加工成尺寸精密的零部件，且又不能承受较重的冲击载荷，故只能以铸件服役于碾磨类破碎机机械中。

非铁合金中常用的耐磨材料有铜基、铝基、锌基及钴基合金，例如铍青铜及铝青铜和著名的司太立合金，还有一些钴基合金以特殊的金属间化合硬质点强化，即含有较多Laves相，被称为Tribaloy合金。它也主要用于堆焊而服役于高温和腐蚀介质的耐磨零件。

以碳化物为主的高硬度金属陶瓷如钴黏结碳化钨和近30年来发展出高速钢刀具再气相沉积超硬膜层，都表现出很高的耐磨性能。

碳材料的质地轻、化学稳定性好，高温强度高，热导及电导性能优良，其中的石墨是理想的固体润滑材料。石墨浸渍合金及树脂的密封环和电机用的碳刷都是在摩擦磨损工况下应用成功的例子。

滑动接触中高分子聚合物显示出金属与陶瓷所不具备的摩擦学特性，例如，不易发生擦伤，化学稳定性好，具有自润滑性等。例如环氧系列、酚醛系列树脂再加上玻璃纤维或其他聚合物甚至金属纤维增强其性能，可以作为一类耐磨材料使用。值得关注的高聚物耐磨和减摩材料是聚四氟乙烯和高密度聚乙烯。PTFE既可用作固体润滑剂，又能以它为基加金属粒子或纤维和无机填料作为耐磨减摩复合材料使用。高密度聚乙烯对泥浆有较低的滑动摩擦因数，还可用注塑工艺制成形状复杂的零件。由于其化学稳定性好，已用作人工髋关节的股骨窝植入人体中。

（2）抗黏着磨损的材料匹配应遵循的规律

1）两种材料的固态互溶性低，不易黏着。一般晶格类型相近、晶格常数相近的材料互溶性较大，最典型的应是相同材料很易黏着。根据材料与钢匹配时抗黏着磨损能力分为优、良、可、差四类，即：

① 优——锗、银、镉、铟、锡、锑、铊、铅、铋；

② 良——碳、铜、硒、镉、碲；

③ 可——镁、铝、铜、钡、锌、钨；

④ 差——铍、硅、钙、钛、铬、铁、钴、镍、锆、铌、铑、钯、铈、钽、铱、铂、金、钍、铀。

固溶性>5%的材料匹配，抗黏着磨损能力很差；化学元素周期表中B族元素及非金属元素与钢铁匹配，有较好的抗黏着磨损能力。

2）两种材料间形成金属间化合物，可减少黏着效应，因为金属间化合物具有脆弱的共价键。

3）塑性材料比脆性材料易发生黏着现象。塑性材料之间，接触区形成的焊点强度常大于母体金属，因而滑动过程中焊点的撕裂常发生于次表层，产生的磨损颗粒大，表面粗糙。脆性材料相反。

4）材料熔点、再结晶温度、临界回火温度越高，或表面能越低，越不易黏着。

5）多相组织比单相组织不易黏着。例如珠光体就比奥氏体或铁素体不易黏着。金属中化合物相比单相固溶体黏着效应差。六方晶体结构优于立方晶体结构。

6）金属与非金属的匹配优于金属与金属的匹配。

7）提高材料表面硬度，表面不易塑性变形，因而不易黏着。对于钢HV700以上不会产生黏着。通常推荐硬件550~750HV，软件200~300HV。如果采用表面硬化，则心部硬度应大于350HV。对二元铜基合金，提高弹性模量有助于提高抗黏着能力。

8）合金钢中各种合金碳化物具有不同的抗黏着性能，大约为 $Cr:W:Mo:V \approx 2:5:10:40$。

9）各种金属材料当压力 $p>1/3$（HB）时开始塑性变形，因此摩擦副表面压强 $p \approx 1/3$（HB）时出现黏着现象，大大加速磨损。

10）奥氏体不锈钢互配极易黏合；钢轴与含Sn、Pb的轴承合金相配，具有良好抗黏着能力；轴瓦表面镀Ag、In有利于抗黏着；脆性灰口铸铁，抗黏着性能优于钢；聚四氟乙烯与钢匹配很耐黏着，且摩擦因数低、表面温度低；耐热的热固性塑料（聚酰亚胺、环氧树脂、酚醛树脂等）较热塑性塑料为好，后者怕热、易软化。

（3）抗磨粒磨损的材料的匹配应遵循的规律

1）纯金属和未经热处理的钢（退火钢），其耐磨性与硬度成正比。

2）经过热处理提高硬度能提高耐磨性。

3）冷作硬化提高硬度，也能提高耐磨性，尤其是塑性材料和较软材料，因为它们很难通过热处理大幅度提高其硬度。

4）对于淬硬钢，如果硬度相同，则含碳量高的钢耐磨性好。

5）马氏体耐磨性优于珠光体，珠光体又优于铁素体；回火马氏体耐磨性比不回火的高；下贝氏体可与马氏体耐磨性相当；低碳马氏体有良好耐磨性；马氏体中的碳化物细小均匀分布比粗大为好；网状碳化物不好。

6）单相组织不如多相组织耐磨。

7）同样硬度值，合金碳化物比渗碳体耐磨；碳化物的合金元素原子越多就越耐磨。

8）钢中加入碳化物形成元素，有利于提高耐磨性，例如Ti、Zr、Hf、V、Nb、Ta、W、Mo等元素优于Cr、Mn等元素。若合金元素不能增加碳化物的数量，即使改变其分布，也不能改变其组织成分；或者合金元素仅溶于铁素体中则不起作用。

9）对三体磨损（除两个摩擦表面外，还有介质颗粒参与摩擦过程），当表面硬度（Hm）约为1.4倍介质颗粒硬度（Ha）时效果最佳，再高则无效；当 $Hm<0.6Ha$ 时，提高Hm无效；Ha低于500~1 000HV或者高于2000HV时均为好；Ha在1500HV左右时磨损率最高。三体磨损的介质粒度小于100 μm 时，越小则表面磨损率越低；粒度大于100 μm 时，粒度与磨损率无关。

10）关于动、静件的硬度匹配关系，动件硬度是最主要的，但太脆却不耐磨。

11）一些常用的抗磨粒磨损的材料有：

① 低碳低合金钢（如16Mn、15MnMoVN等）低碳马氏体，在低于100~150℃下使用，耐磨性高于低碳钢，而工艺、价格相近。

② 中碳钢（40、45钢）和中碳低合金钢（40Cr、42CrMo等）具有较好的强度和耐磨性，常用作齿轮、轴等。

③ 高碳低合金钢GCr15和CrWMn等更为耐磨。

④ 如果要求耐磨和耐锈蚀性，可用马氏体不锈钢（3Cr13、4Cr13）淬火后低温回火的组织，在液体介质中工作的泵或叶轮等可用ZGCr8CuMo等。

⑤ 高碳高合金钢常用于对耐磨性要求极高的零件，如挖掘机的斗齿、颚式破碎机衬板、坦克拖拉机履带等常用ZGMn13，这种钢有很强的加工硬化能力。

⑥ 铸造石墨钢耐磨性优于ZGMn13，且价格低，兼有铸钢和铸铁的性能，可用作球磨机衬板等，但导热性差。

⑦ 耐磨铸铁耐磨性好，经济性佳，但比合金钢脆、强度也差。适合于矿山、水泥、煤、农机等耐磨件。

⑧ 硬质合金、人造金刚石等烧结材料价格较高，但极坚硬，常用于高速切削工具或特殊轴承。

（4）抗接触疲劳的材料的匹配应遵循的规律

1）表面硬度与抗接触疲劳能力大体呈正比关系，故提高硬度有利于抗接触疲劳。若硬度过高，太脆，反而使接触疲劳寿命下降。滚珠轴承钢硬度为62HRC时，接触疲劳抗力最高。

2）对于软表面（低于350HB），摩擦副之间硬度差为50~70HB时，两表面易于磨合、服贴，有利于抗接触疲劳。例如齿轮传动，小齿轮调质，大齿轮正火，其接触疲劳力明显优于调质对调质的匹配。

3）为严格控制初始裂纹和非金属夹杂，需严格控制冶炼和轧制过程。

4）钢中碳化物宜细小、分布均匀；当钢含碳量高于0.53%时，马氏体易粗大，脆性增加。

5）高锰钢做齿轮抗接触疲劳性能好，具有良好淬透性、表面硬度高、晶粒细小；铬钢、钒钢也有防止晶粒长大、淬透性好等优点；需要抗腐蚀时，宜用马氏体不锈钢，有较好的耐磨性及耐接触疲劳抗力；灰口铸铁虽然硬度低于中碳钢，但由于石墨片不定向，且摩擦因数低，故有较好的接触疲劳抗力；陶瓷材料具有高硬度和良好的抗接触疲劳能力，而且高温性能好，但多数不耐冲击。

（5）抗微动磨损的材料选择

适于抗黏着磨损的材料匹配也适于抗微动磨损，此外抗蚀性好，可形成软的氧化屑或磨屑，抗磨粒磨损好，都能有利于抗微动磨损。实践证明：灰口铸铁与不锈钢匹配（加 MoS_2 润滑），灰口铸铁对铜、灰口铸铁对灰口铸铁（加 MoS_2 润滑），以及工具钢对工具钢，钢对聚四氟乙烯等都有很好的抗微动磨损能力；而铝对灰口铸铁、铝对不锈钢、镀铬表面对镀铬表面、工具钢对不锈钢等匹配性质较差。

（6）抗腐蚀磨损应选用耐腐蚀的材料

在其表面形成的氧化膜能与基体结合牢固，氧化膜韧性好且致密的材料，具有优越的抗腐蚀磨损能力。通常用含 Ni 和含 Cr 的材料，而含 W 与含 Mo 的材料能在 500℃以上的高温条件下生成保护膜，并降低摩擦因数，可以作为高温耐腐蚀材料，WC 与 TiC 等硬质合金均有很好的耐腐蚀磨损性能。

总之，选择耐磨材料时，首先要确定磨损的类型，考虑材料本身的性质及配件材料的性质，材料匹配的相容性有时还与环境条件以及工作参数有关。

2.4.3 表面处理方法选择和设计

(1) 表面强化处理方法分类

表面强化处理是在选用通用材料的基础上，用工艺手段达到使材料表面强化、耐磨损的目的，使高性能与经济性较好地结合起来。常用的强化处理有三类。

1) 机械加工（包括切削加工、冷热成形加工、滚压、喷丸等）与热处理。这一类强化的特点是不改变表面的化学成分，靠改变材料表面的显微组织和结构，或改变加工表面的几何形貌来达到强化、耐磨损的目的。

2) 扩散处理。这一类强化的特点是要改变表面的化学成分，靠渗入、注入某些元素的办法，或同时附加热处理的手段，使表面得以强化、耐磨损。例如各种化学处理、化学热处理。

3) 表面涂覆。这类强化的特点是直接在材料表面镀、涂或用各种物理、化学方法覆盖上一层强化耐磨层。覆盖层分硬涂层和软涂层两种，硬涂层经常是镀铬、堆焊以及喷涂碳化物与陶瓷等所得，软涂层则常常是针对黏着磨损而制成的，目的是降低摩擦因数及提高耐温性等，软涂层包括涂覆铜、铟、金、银等软金属，也包括涂覆聚四氟乙烯和 MoS_2 等固体润滑剂。

选择表面强化时，应全面考虑以下各方面因素：对不同磨损类型的强化效果，对各种材料或零件（大小、形状及其部位）的适用性，零件的工况及环境条件，表面强化对材料（尤其金属材料）性能的改变，表面强化对后续加工的影响，表面强化对零件尺寸的改变、成本等等。

(2) 表面强化处理的性能

1) 机械加工。它改变金属表面的组织和结构，致使表面硬度提高，并且通常使表面层产生相当大的残余压应力，因而使摩擦表面的耐磨粒磨损性能和耐接触疲劳性能得到提高。机械加工后的强化性能见表 5.5-5。

表 5.5-5 机械加工后的强化性能

加工方法	表面硬度提高/%	表面残余压应力/MPa	强化层深度/mm
切削	20 ~ 30	3 ~ 7	0.05 ~ 0.5
喷丸	20 ~ 40	4 ~ 8	0.4 ~ 1.0
滚压（柱）	20 ~ 50	6 ~ 8	1 ~ 20
滚压（球）	20 ~ 50	6 ~ 8	0.3 ~ 5
振动冲击加工	20 ~ 40	3 ~ 6	0.1 ~ 0.7
滚筒抛光	10 ~ 15	1 ~ 2	0.05 ~ 0.3
超声波强化加工	50 ~ 90	8 ~ 10	0.1 ~ 0.9
金刚石碾压	30 ~ 60	3 ~ 7	0.01 ~ 0.2

对于冷加工强化来说，原料塑性越好、软，强化的效果越明显。

2) 关于加工表面形貌对耐磨性能的影响，表面几何形貌一方面影响润滑的有效性，另一方面也影响润滑膜破坏后表面粗糙峰之间碰撞的概率和应力水平、变形大小，因此，它对磨损起着极为重要的作用，有时其对提高零件的耐磨性功效远超过改善材料成分和热处理。

3) 热处理。通过热处理可以改变金属表面的显微组织和结构，达到提高耐磨性的目的。各种常见的显微组织的特性见表 5.5-6。

表 5.5-6 显微组织的特性及其对耐磨性的影响

显微组织和形态	硬度 HV	耐 磨 性
铁素体：在碳钢中 合金铁素体	50 ~ 135 100 ~ 270	不耐磨 不太耐磨
奥氏体：在合金钢中	150 ~ 220	本身硬度不高，但能在冷塑性变形下得到强化，所以当受到冲击、强压或大塑性变形时，能提高耐磨性
碳化物：渗碳体 合金钢中复杂碳化物	1 000 ~ 1 150 1 100 ~ 1 300	根据不同的形成条件，而有不同的形态与弥散度，耐磨性也不同 根据不同的形成条件，而有不同的形态与弥散度，耐磨性也不同
珠光体：球状 片状 索氏体型 索氏体 托氏体	160 ~ 190 190 ~ 230 230 ~ 260 250 ~ 300 300 ~ 400	碳化物数量越多，硬度也就越高，越耐磨 片状比球状珠光体耐磨
马氏体	800 ~ 900	高强度、高硬度、耐磨，作很多耐磨铁碳合金的基体
石墨 灰铸铁中共晶体： 石墨型（F + G） 磷化物型（P + Fe3P + Cm） 碳化物型（Cm + P）	不论什么形态都很软 130 775 900 ~ 1 000	能起润滑和减摩作用

4) 热-机械处理。在高温状态下先施以较大的塑性变形，然后再淬火冷却，或者在低温下（过冷奥氏体）对轴承钢进行变形处理，这两种都是将热处理和塑性变形相结合以提高强度、硬度和耐磨性、耐接触疲劳的措施。

5) 扩散处理。常用的扩散处理，其适用性、特性及耐磨性等见表 5.5-7。

表 5.5-7　扩散处理的适用性、特性和耐磨性

名称	适用性	强化层硬度及厚度	耐磨性
渗碳	低碳钢或低碳合金钢（例如<0.2%C）	55~65HRC 厚：0.5~1.5 mm	改善对接触疲劳的抗力，改善残余应力的状态。耐腐蚀性不如原来
液体氮化	低碳不锈钢和合金钢，铸铁，工具和模具钢	21~40HRC（当593℃回火后） 91HRB~33HRC（当649℃回火后） 化合物层厚：0.005~0.008 mm，扩散层厚0.25 mm	低碳钢100%、中碳钢80%接触疲劳抗力有提高。耐腐蚀性一般均略有提高。Falex试验载荷容量为3 400~4 750 N
气体氮	含Al、Cs、Mo、V、W等合金元素总量在1%以上的合金钢（通常用1% Al、1.5% Cs、0.3% Mo的合金钢），不锈钢和高速工具钢	58~70HRC 化合物层厚：≤0.08 mm（合金结构钢）≤0.01 mm（工模具钢） 不锈钢无化合物层 扩散层厚：0.2~0.64 mm（合金结构钢），0.03~0.2 mm（不锈钢及工模具钢）	提高了接触疲劳抗力。对耐腐蚀性在有白层的情况下通常是提高的，但对不锈钢耐腐蚀性则是降低的
辉光（等离子体或离子）氮化	灰铸铁、合金钢、工模具钢、不锈钢、钛合金	化合物层：1 000~1 500HV 厚：0.03 mm 扩散层：500~1 500HV 厚：1 mm	显著提高接触疲劳抗力，对耐腐蚀性则合金元素含量越少越耐腐蚀
碳氮共渗	碳素钢（正常低碳0.2%C）合金钢中适于渗碳的类型	58~63HRC 厚：0.08~0.51 mm	改善了接触疲劳抗力，对耐腐蚀性无影响
氰化及液体渗碳	低碳钢及低碳合金钢	55~65HRC 厚：0.03~0.25 mm（氰化） 0.25~1.52 mm（液体渗碳）	改善了接触疲劳抗力对耐腐蚀性无影响
渗硼	低碳钢、镍基和钴基合金、难熔合金、工模具钢	73~85HRC（钢件） 厚：0.013~0.127 mm	对接触疲劳影响不大。耐腐蚀性很好。Falex试验载荷容量为2 200 N（未渗前为2 850 N）。四球机试验载荷容量5 000 N（未坏），而未渗前800 N胶合
扩散（Cr、Si、Al、Zn、Cs等）	碳钢与合金钢，不锈钢、电工与工具钢、耐高温钢	70~76HRC（渗铬） 厚：0.15~1.0 mm（渗Al） 0.03~0.05 mm（渗Cs） 0.05~0.08 mm（渗Zn） 0.1~0.8 mm（渗Si）	改善耐腐蚀性

6）表面涂覆。常用的表面涂覆技术，包括电镀、化学镀、阳极处理、喷涂、离子镀、涂覆固体润滑、气相沉积、电焊等，其适用性、特性、耐磨性等，见表5.5-8。

电焊硬表面合金的方法，有很广泛的适应性和耐磨性，表5.5-9中按照不同的磨损类型推荐相应的硬表面合金。

表 5.5-8　表面涂覆的适用性、特性和耐磨性

名称	适用性	覆盖层硬度及厚度	耐磨性
阳极电镀	生铝和熟铝、Mg、锌合金、Ti、Be（铸或热压件）	硬度取决于厚度、密度与孔隙度，约60~80HRC 厚：Al—1.27~7.62 μm Al—12.7~114 μm Mg—2.5~25 μm Zn—32~41 μm Ti—102 μm（耐磨损） Be—2.5~7.6 μm	降低了铝的疲劳抗力，因此最好先进行扩散处理然后再镀。在钛表面形成钛化铁对疲劳抗力有利。耐腐蚀性：镀Be能耐900℃（大气条件下）
镀铬	铁合金及Ti、Mo、Al	300~1 000HB（1 000HB时最耐磨） 厚：0.01~0.31 mm	通常降低疲劳抗力，对耐腐蚀性有提高
镀银	钢	约100HV 厚：0.5~1.5 mm	通常对耐腐蚀性有提高，但如果镀层有微孔隙，且银成为钢的阴极，则会加速腐蚀
铬酸盐化学镀	Zn、Cd、Al、Mg、Be	1 000~1 800HV（载荷9.25 N） 厚：0.5~1.5 mm	镀铍：≤649℃抗氧化 镀铝：抗蚀性极好 镀镉及锌：降低腐蚀性疲劳倾向 抗磨粒磨损性不好

续表 5.5-8

名称	适用性	覆盖层硬度及厚度	耐磨性
化学镀（无电电镀）	铁及铁合金（含高硅除外）、铜合金、铝及铝合金、钛合金等	500HV左右 厚：0.08~0.13 mm	耐腐蚀性优于电镀，不宜承受冲击或反复弯曲变形
磷酸盐化学镀	铁合金及钢（硬度 < 48HRC），铝及铝合金、锌、镉	厚：0.004~0.008 mm	耐蚀性中等到优良，耐疲劳方面无作用
固体膜润滑剂	钢、铝及铝合金、铜及铀合金、不锈钢、钛、铬镍轴承表面	硬度：很软 厚：< 0.01 mm	对耐疲劳能力无作用，可在固体膜润滑剂中添加抗腐蚀剂，由于固体膜的减摩作用，在 Falex 试验中，当载荷为 4 450 N 时，能维持 4 h 以上
聚四氟乙烯	钢、铜及铝	厚：0.000 1 mm（在不锈钢表面可沉积钛层 0.005 mm 厚）	在模具表面沉积 TiC 后，使用寿命从 25 万次提高到 100 万次（并且原来每使用 2 万次还要清理一次表面，而沉积后不必）
电焊	原则上，任何相同成分的材料均可焊接，常用的硬度表面合金分五种：①铁基合金；②钴基合金；③镍基合金；④铜基合金；⑤碳化物复合材料	硬度可以很高，厚度可以很厚	很好

表 5.5-9 硬表面合金的选择

磨损类型	推荐的硬表面合金
黏着磨损	
中度磨损	含 6% ~ 16% Cr 的铁基合金；含 15% ~ 30% Cr 的钴基或镍基合金
严重磨损	钴基合金或镍基合金
磨粒磨损	
低应力磨损	在铁、钴或镍合金中含碳化物颗粒，含 3.5% ~ 4%C 的高铬铁
高应力磨损	含 2.0% ~ 3.5% C 的高铬马氏体铁，含 0.4% ~ 1.5%C 的马氏体合金钢
犁沟（并冲击）磨损	含 2.0% ~ 3.5% C 的高铬铁的奥氏体锰钢
冲蚀磨损	
低冲击角	在铁、钴或镍合金中含碳化物颗粒，含 3.5% ~ 4%C 的高铬铁，过共晶钴或镍合金
高冲击角	含 2.0% ~ 3.5% C 的高铬铁，亚共晶钴基合金
气蚀磨损	亚共晶钴基合金
微动磨损	钴或镍基合金
腐蚀-冲蚀磨损	钴或镍基合金，含 12% ~ 18% Cr 和 0.2% ~ 1.2%C 的高铬钢
氧化（热）腐蚀	含有残余相的镍或钴基合金，含碳化物的镍或钴基合金
高温磨损	钴基合金，含有残余相的钴或镍基合金

喷涂也是一种较常用的表面强化技术，它通过爆炸或等离子弧把粉状物加热并高速打到材料表面，固化形成一层耐磨表面层。此覆盖层厚度一般不超过 0.5 mm。

3 抗失效的机械设计技术和方法

3.1 抗脆断设计技术和方法

3.1.1 抗脆断设计思想

早在 20 世纪 40 ~ 50 年代，防止脆性断裂的问题就已作为重大技术问题提了出来。那时，由于工业的进步，结构向大型化发展，同时所受的载荷越来越苛刻；一些结构采用焊接连接，不可避免的在接头中带有缺陷；以及许多飞行结构片面地追求高强度材料而忽视韧性，因此产生了许多惊人的脆性断裂事故。例如在电站设计中，重数百吨的高速旋转的构件脆断，造成灾难性事故，影响正常发电，并需要长时间的修复，造成严重后果。又如在压力容器和压力管道方面，脆断可使容器或管道爆炸，特别是天然气储罐的爆炸、燃烧，可使周围变为一片废墟。

因此，一些先进国家先后开展了防脆断研究，因而推动了线弹性断裂力学理论的发展，并逐步建立了防止脆断的工程设计方法，使它成为保证结构可靠性的有效方法。通过实践检验，许多国家相继制定了防止脆断的设计规范。

防止脆性断裂的工程方法基于线弹性断裂力学理论，认为 K_I 是表征裂纹尖端应力应变状态、控制裂纹失稳扩展的参量。以含裂纹结构的应力强度因子的弹性解，结合从材料标准试件测得的断裂韧性与温度的关系曲线，可判断结构是否将发生脆断，确定临界裂纹尺寸的大小，评定许用缺陷尺寸，按工程要求、使用环境及条件设计结构，使其满足：

$$nK_I < K_{IR} \qquad (5.5\text{-}12)$$

并按裂纹扩展规律定出服役产品的检验周期和设计产品，使其在检查时仍能满足上式的要求，以保证产品的安全。

由于在线弹性范围内，J_{IC}、COD 与 K_{IC} 的一致性，也可按 J_{IC} 的 δ_c 来设计防脆断结构。

设计方法的要点及程序示于图 5.5-7 中。

3.1.2 抗脆断设计分类

1）按断裂判据分　K 判据法、COD 法、J 积分法、应变能密度因子法等。

2）按许用缺陷尺寸设计分　多用于评定焊接结构的缺陷尺寸。有当量裂纹尺寸法、临界裂纹尺寸法等。

3）按脆性转变温度设计分　断裂分析图法、确定能量水平法等。

4）按一定设计要求分　例如压力容器中的“先漏后破”准则。

5）半经验方法　例如寻找断裂韧性与常规性能的关系，并按经验给出常规性能的限制。

3.1.3 防脆断结构设计方法

通过材料的选择和设计，是防止断裂的一个手段，而在结构设计上防止断裂的出发点是降低工作载荷下零件的最大应力。这是因为在工作载荷一定的情况下，最大应力决定于

零件的尺寸。

(1) 设计降低最大应力的零件结构

1) 非圆形截面梁和支座 矩形、方形及工字形截面是主要受弯矩的梁常用的截面。在截面积相同时，立矩形的抗弯截面系数是方形的两倍以上，而工字形截面的抗弯截面系数是方形的5.5倍。其最大应力，当然是同样的倍数。要达到相等的抗弯强度，采用工字形截面所需材料的质量仅是方形截面的1/5～1/8。

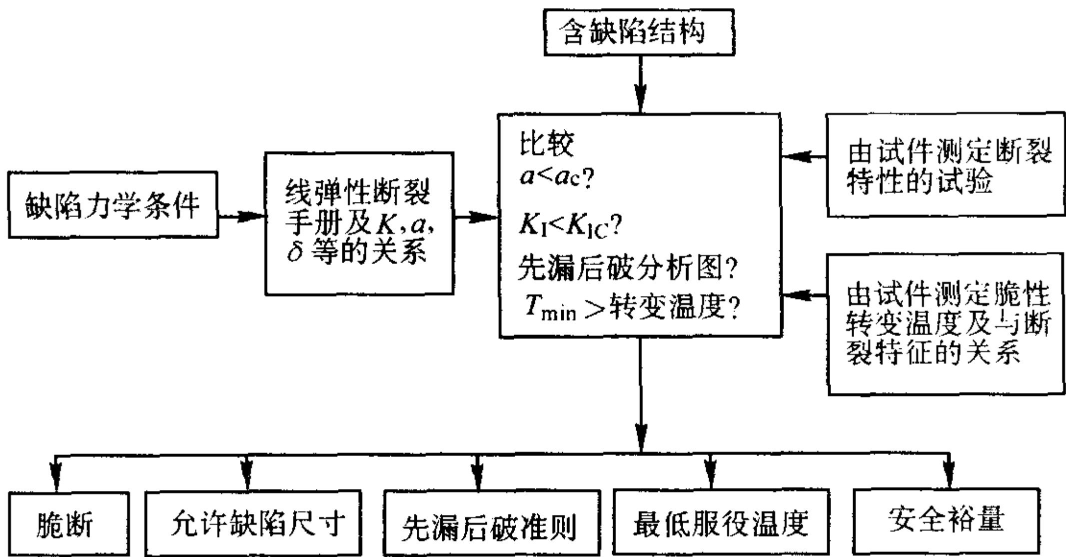


图 5.5-7 检验缺陷结构脆断的工程设计方法

钢材的抗压强度等于抗拉强度的1.2～1.6倍。铸铁的抗压性能更优于它的抗拉性能。在有些情况下，设计支架结构或与此类似的结构时，应根据受力状况，将材料的特性与应力分布结合起来考虑，使结构设计得更合理。如图5.5-8所示图中两支架的受力和应力分布状况可以看出，图b中的拉应力小于压力应力，符合材料的特性，图b支座的结构更为合理。

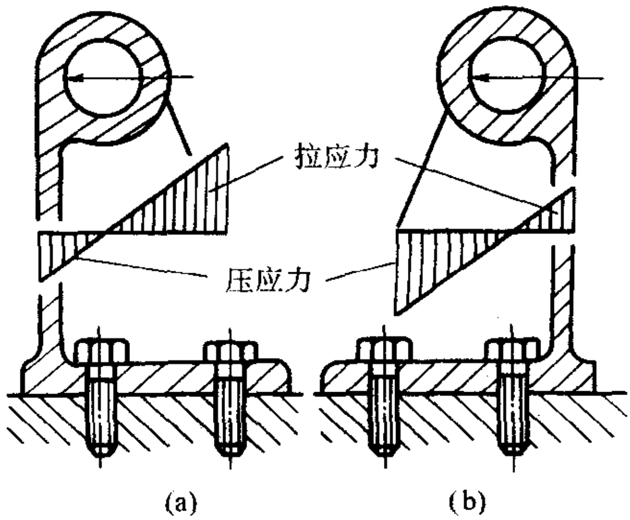


图 5.5-8 两铸铁支架的比较

2) 轴类零件 轴类零件多受弯曲和扭转力矩。虽然一般转轴为交变或脉动应力，但仍以最大应力为强度计算指标。为了节约材料、减小质量、又降低最大应力，须合理设计它们的支承和截面。

对于悬臂支承的齿轮轴，应尽量减小伸出部分的长度和合理选择支承间的距离。支承间的距离及其与伸出长度的比值直接影响支承受力和轴之弯矩的大小。悬臂轴改为中间受力结构时，其受力状况能得到更好的改善。对于大直径圆截面轴类零件，做成空心环形截面能使在受弯矩时的切应力得到合理分布，使材料得以充分利用。用相同质量的材料，空心轴可获得较实心轴大得多的强度和刚度。

轴类零件的轴肩或两不同直径轴段的过渡部分常用圆角或越程槽。当载荷大而轴的直径尺寸又受限制时，须仔细设计过渡部位的结构，减轻应力集中对轴强度的影响。

(2) 改善零件的受力状况

改善零件受力状况是从结构设计上提高零件强度的一种重要措施。从这方面采取措施常可获得显著的效果。零件受力状况的改善使内部应力降低，从而提高零件的强度。

1) 载荷分担 即把作用在一个零件上的载荷，经采取结构措施后，分给两个或更多的零件承担，从而减轻单个零件的载荷。一个典型的例子是卸荷皮带轮，它常用于机床传动箱外的带传动。

载荷分担的方法也可用于滚动轴承支承的设计。对于承受中等的径向载荷及中等的或较小的轴向载荷，通常采用向心推力轴承就可以了。对于承受径向载荷及较大的或大的轴向载荷时，就需采用向心轴承和推力轴承分别承担径向力和轴向力。

弹簧族和弹簧组合的设计也是采用了载荷分担的方法。如果用一个圆柱形螺旋弹簧承担大的载荷，弹簧丝要很粗，而钢丝直径越粗它的许用应力越低，弹簧的绕制也更困难。若设计成直径不同的两三个套在一起组成的弹簧族，由几个弹簧共同承担一个大的载荷，则可收到减小外廓尺寸、减小弹簧质量的良好效果。

2) 载荷均布 将一个集中载荷分为几个较小的集中载荷或均布载荷，将使轴或梁的受力状况得到改善。齿轮的轮齿常发生沿齿宽载荷分布不均甚至严重偏于一端的现象。行星齿轮传动装置的几个行星轮也存在使载荷均布的问题。

重型齿轮减速器中，齿轮位置的合理布置能改善齿轮本身、轴和轴承的受力状况。对于一般齿轮减速器，为了使结构简单，多采用非对称布置，将齿轮置于远离动力输入端，轴的扭转变形可以部分地改善因轴弯曲变形导致的轮齿沿宽度方向的载荷分布不均。

渐开线齿轮传动若沿齿宽具有相同的截面，啮合齿沿接触线长度上的载荷将集中在偏于两端的部分。若将轮齿修整成鼓形，使齿的中间部分和端部的厚度差25～35 μm，则可以靠齿面受力的弹性变形使沿齿宽方向的载荷分布得比较均匀。

为了不使轮齿的载荷集中在两端，还可以采用减小齿轮端部的局部刚度，使端部在受力后产生较大的弹性变形，从而载荷分布趋于均匀。

轴颈和它上面的齿轮之间传递扭矩的方式，也会影响轮齿载荷的分布。

行星齿轮传动中，使行星轮之间的载荷均匀分布是必须达到的要求。使用行星轮浮动套有很好的均载效果。在浮动套与行星轮间有较大的配合间隙。转动中将油吸入间隙使具有挤压膜效应，油膜起弹性元件的均载作用，使行星轮受载均匀。

重型设备、重要部位、受较重载荷的螺母，要注意从螺

母的结构设计上解决啮合螺纹间的均载问题。普通结构的螺母由于靠近被联接件的螺母，刚度最大，沿着离开被联接件方向，刚度依次减小，所以各螺纹间的载荷分布很不均匀。

总之，上述列举的各种使载荷均布的结构措施，其中大部分的共同点是从改变刚度或借助弹性变形来调节载荷分布的。

3）载荷抵消　载荷抵消是采取措施使外载荷部分地或全部地相互抵消，从而使相关零件所受的载荷减小。例如，在同一根轴上有两个斜齿圆柱齿轮，设计者可以选定两个齿轮轮齿的旋向，使二者产生的轴向力相互抵消。这样，可以减小轴承的轴向载荷。

在高速回转机械中，旋转惯性力必须靠结构的措施及动平衡的方法使之降低到允许的大小。回转件的质量须相对于回转中心呈对称分布，将回转件在动平衡机上做动平衡试验，则是为了测出并消除超出允许值的不平衡质量。

做往复运动的机械，如连杆机构，也可在设计中采取结构措施和动平衡的方法，使其在运转时产生尽可能小的惯性力。

(3) 弹性强化和塑性强化

用弹性强化的方法提高强度，是使零件在受外载荷之前，产生与外载荷的作用相反的弹性预变形，受外载荷后，预应力部分地抵消外载荷产生的应力，从而降低最大应力。

塑性强化（也称过载强度）是使零件在工作状态下应力最大的那部分材料预先经间隔性变形产生与工作应力符号相反的残余应力，以此来部分地抵消工作应力。塑性强化的塑性变形产生在最大应力附近的材料中，残余应力抵消工作应力也是在最大应力附近的区域内。所以塑性强化的效果，是降低最大应力、使应力分布趋于均布。

应当注意，只有当材料具有足够的塑性时才能采用过载强化，在脆性材料中，越限应力会在受拉层引起使零件损坏的微观裂纹和断裂。在塑性材料中，过度的变形也会导致发生同样的现象。因此，为限制塑性变形，超限应力不得超过 $(1.1\sim1.2)\ \sigma_{0.2}$。

还应该考虑到：每一种方式的越限应力只能强化作用载荷对面那一个方向的材料；若作用载荷反向，原先的越限应力反而是削弱了材料。因此，这种方法只适用于定向载荷、脉动载荷和具有一个方向优势的交变载荷。

对零件在毛坯阶段、冷态或半塑性状态受压缩的方法，使工作载荷下为拉应力的那部分材料受深度的压缩。这样，在零件中存在与工作应力符号相反的应力，最终的效果十分显著。

3.1.4 几种主要的防脆断设计方法

(1) 基于 K_{IC} 的设计方法

这种方法提供了构件防止脆断的许用载荷。此方法基于线弹性断裂力学的原理，并使考察点的应力强度因子 K_I 小于构件材料的断裂韧性的下限值 K_{IR}。

这一方法的典型代表是ASME的锅炉与压力容器规范第Ⅲ篇附录G和第Ⅺ篇附录A的方法。它的基本要点是：

1）测定材料的断裂韧度，确定参考临界应力强度因子值 K_{IR}。

2）测定材料的参考无塑转变温度，它是用落锤试验决定的材料无塑转变温度，再附加夏比冲击试验的要求后确定的温度。

3）作参考临界应力强度因子值 K_{IR} 与参考温度参数的曲线。

4）确定最大缺陷尺寸　如果已有产品，可通过探伤并进行规则化处理后确定最大缺陷尺寸（规则化处理见本篇第4章）。但是多数情况下，在设计时并无产品，因此只能根据以往的经验假定一个缺陷进行设计。

5）计算应力强度因子　应包括结构承受的各种载荷所引起的应力强度因子，一般采用 $K_I = Y\sigma\sqrt{\pi a}$ 计算，其中 Y 为形状因子。

6）选取安全系数　为了保证安全，在计算主要应力强度因子时，还要打上安全系数。

7）许用载荷条件：$\sum nK_I < K_{IR}$。

(2) 按许用缺陷尺寸的设计方法

如在制造过程或使用中，用各种检查方法发现存在缺陷，则需评定缺陷是否容许。这种方法常用于评定焊接结构中存在的焊接缺陷。典型的有国际焊接协会的“按脆断破坏观点建议的缺陷验收标准”、日本焊接协会的“按脆性起裂的焊接缺陷评定标准”等。这种设计方法有两种设计步骤。

1）缺陷尺寸的规则化。

2）决定缺陷的特征尺寸　将第一步决定的简化缺陷尺寸，化为统一的等效穿透缺陷，并以穿透缺陷的半长作为缺陷的特征尺寸，以便比较。

3）确定缺陷处在无缺陷的作用应力或应变值　此方法在线弹性阶段可用应力或应变评价。超过弹性时此方法仍可应用，但均按应变评价。

4）材料断裂韧度的确定　在所需温度条件下，用三点弯曲试样测量断裂韧度。

5）允许裂纹尺寸的确定　根据线弹性断裂力学确定临界裂纹尺寸。

6）进行评定　将检测的缺陷尺寸与允许裂纹尺寸比较，当检测缺陷尺寸小于允许裂纹尺寸时合格，结构不会发生脆性断裂；

检测缺陷尺寸大于允许裂纹尺寸时不合格，缺陷不能被接受。

以裂纹张开位移COD为评定判据的称为COD法。这一方法的设计要点如下。

1）缺陷尺寸的规则化；

2）决定缺陷的特征尺寸；

3）确定缺陷处在无缺陷的作用应力或应变值等步骤与上面的方法相同。

4）计算张开位移　从第2）步求得的缺陷尺寸及第3）步求得的应变，可以求出张开位移。

5）确定材料的断裂韧度。

6）评定　当张开位移小于临界张开位移时，合格；当张开位移大于临界张开位移时，缺陷不能接受。

(3) 按一定设计要求（或准则）的防脆断设计方法

按“先漏后破”准则设计压力容器。压力容器的“先漏后破”准则是ESSO公司首先提出的。这个方法假设一个半圆形表面裂纹，由于疲劳而扩展，当扩展到穿透壁厚时，用深为壁厚、长为二倍壁厚的裂纹表示，并要求所使用材料的断裂韧度相对这一尺寸的裂纹，在应力达到屈服时也不会引起脆性破坏，而是首先泄漏，同时这一泄漏是可以检验的，因此，可向使用者早期告警，及时采取措施，避免灾难性的脆断发生。

(4) 按脆性转变温度设计的方法

材料的断裂韧度随温度发生变化，因此，可以按脆性转变温度进行防脆断的设计。可以选择的脆性转变温度有：

1）规定某一能量水平对应的温度；

2）无塑转变温度；

3）50%纤维断口对应的温度。

在应用脆性转变温度进行防脆断设计时，应特别注意板厚、裂纹尺寸及加载速率的影响。

(5) K_R 或 G_R 阻力曲线方法

用于评定平面应力型的断裂情形，它是阻力曲线中的一种。R阻力曲线方法主要用于防止韧性断裂的评定，但是，和COD法一样，由于 $G_I = \frac{K_I^2}{E'}$，因此也可用于防脆断设计。其方法是：由实验确定材料的 K_R 或 G_R 阻力曲线，计算裂纹结构受载荷下的 G 或 K_I。当载荷小时，R阻力曲线与 G 或 K_I 曲线相交，此时 $\frac{\partial G}{\partial a} < \frac{\partial G_R}{\partial a}$ 或 $\frac{\partial K_I}{\partial a} < \frac{\partial K_R}{\partial a}$。随着载荷的增大，$G$ 曲线与 G_R 曲线相切，裂纹达到临界值，此时如果载荷再增大，裂纹就会失稳扩展。裂纹失稳扩展的判据是：

$$\left.\begin{aligned} G &= G_R \\ \frac{\partial G}{\partial a} &= \frac{\partial G_R}{\partial a} \end{aligned}\right\} \quad (5.5\text{-}13)$$

或

$$\left.\begin{aligned} K_I &= K_R \\ \frac{\partial K_I}{\partial a} &= \frac{\partial K_R}{\partial a} \end{aligned}\right\} \quad (5.5\text{-}14)$$

G 曲线与 G_R 曲线的切点，或 K_I 曲线与 K_R 曲线的切点是平面应力断裂韧度 K_C 或 G_C，由此可求出与 K_C（G_C）对应的临界裂纹尺寸。

(6) 经验方法

由于 K_{IC}在防脆断设计中的重要性，因此需要测定材料的 K_{IC}。但 K_{IC}试验的试件制作比较困难，而且试件尺寸要求满足平面应变条件，因而变得很大，少量材料有时会得不到满足条件的数据。从常规力学特征求 K_{IC}估计值有一些经验方法。

另外，我国的压力容器含缺陷安全评定规范中提供了一个合乎使用要求，防止脆性和疲劳破坏的工程评定方法，详细内容请参考本篇第4章。

防脆断设计的方法主要基于线弹性断裂力学的原理，这要求裂纹尖端存在较大范围的线弹性应力场或 K 控制场，也就是说要求裂纹尖端塑性区尺寸相对裂纹深度很小。因此，工程上对应力 $\sigma < 0.75\sigma_s$（平面应变状态）及 $\sigma < 0.45\sigma_s$（平面应力状态），可采用线性断裂力学的方法，超过此范围则需应用弹塑性断裂力学的理论。

3.2　抗韧断设计技术和方法

3.2.1　抗韧断设计思想

在生产实际中，结构广泛采用中、低强度钢制成。如果结构带有缺陷，在断裂前，早已大范围屈服。同时，有的结构存在高应变集中的部位，受载时这些部位产生塑性变形。另外许多加工工艺，如焊接引起结构中产生残余应力等，这些都已超出了线弹性断裂力学的适用范围，必须开展弹塑性断裂力学的研究。此外，大量断裂事故及实验表明：用线弹性断裂力学方法来分析问题，其结果与实际差距很大。

弹塑性断裂的分析很复杂，表现为：

1）材料变形及结构几何形状变化的非线性；

2）弹塑性断裂过程复杂，经过裂纹尖大量钝化、开裂、缓慢的裂纹扩展和失稳扩展；

3）裂纹在平面和斜面上扩展，为混合型断裂。

因此弹塑性断裂的分析和计算是很复杂和困难的，亟待建立对工程设计及应用都适合的简单方法。

抗韧断设计工程方法基于 J 积分理论、J 积分控制裂纹扩展理论，用已有的含裂纹结构的 J 积分弹性解和 J 积分全塑性解内插，联合建立并推导出弹塑性解。经与从材料标准试件测得的 J 积分阻力曲线和由本方法建立的裂纹推动力图共同使用，可预计缺陷结构的开裂、裂纹扩展和失稳裂纹扩展。

3.2.2　抗韧断设计方法

韧性断裂的实质是真实应力超过材料的屈服强度所致，因此，零件防止韧性断裂需要对金属零件韧性破断失效原因分析，应对失效件进行设计、材质工艺和实际使用载荷进行复核；分析是否存在设计结构不当、选材不当，原材料尺寸、表面质量、冷热加工缺陷以及严重超载导致失效。相关内容请参考本章的2.2.1节和3.1.3节。

断裂力学的发展，是有可能对含缺陷的机件或结构的承载能力进行定量分析评定，这对于保证机件和构件安全运行，防止事故发生具有重要意义。这方面的研究成果已纳入一些国家的工程技术规范，成为必须执行的指导性文件。由于某些方面还不很完善，有些文件尚为草案。进行这样的评定的总步骤如下。

1）探伤确定缺陷部位、分布、几何形状和尺寸，并进行简化处理。

2）用断裂力学方法确定缺陷处的应力强度因子或其他断裂力学参量COD和J等。

3）试验测定材料（母材、焊缝、热影响区、大型铸锻件不同部位）的断裂韧度。

4）依所选用的评定规范进行安全评定。

现行的主要规范如下。

1）美国ASME锅炉及压力容器规范第Ⅲ篇附录G的方法。

2）美国ASME锅炉及压力容器规范第Ⅺ篇附录A的方法。

3）国际焊接学会IIW的评定标准草案。

4）英国标准BS7910。

5）日本焊接工程学会JWES评定草案。

6）美国电力研究所EPRI弹塑性断裂估算方法。

7）我国的《在役含缺陷压力容器安全评定规程》。

评定方法已在本篇第4章中进行了详细的介绍，这里不再重复。

3.3　抗疲劳设计技术和方法

3.3.1　抗机械疲劳设计技术和方法

(1) 抗机械疲劳结构设计

1）减小构件的应力集中　实际零件不可避免地带有多种形状的缺口，如台阶、键槽、螺纹、孔等。缺口造成应力集中，使疲劳极限降低。缺口越尖锐，疲劳极限下降越多。所以防止疲劳失效的最简单方法是尽量减少不必要的应力集中。粗糙的加工表面也会引起应力集中，同一材料经不同切削加工工艺（车、铣、刨、磨、抛光等）后，表面粗糙度相关很大，因此疲劳寿命也差别甚大。例如抗拉强度为980 MPa的CrMo钢，抛光后的疲劳极限为540 MPa，磨削则为500 MPa，精车为400 MPa，粗车后仅为351 MPa。对于带有裂纹的构件，裂纹的发展起了支配作用，构件的应力集中或表面粗糙度就不再影响疲劳性能。

2）表面处理及表面应力状态　疲劳源多产生于零件表面，因此采用表面强化处理是提高疲劳强度的有效途径。通常对零件进行的表面处理，有表面滚压、喷丸、表面淬火、渗碳、渗氮、碳氮共渗以及激光热处理、离子注入等。这些处理一方面提高了材料表层强度，推迟裂纹萌生，另一方面，可能是更重要的，在表层形成有利的残余压应力。

(2) 抗机械疲劳设计方法

抗疲劳设计准则已由无限寿命设计发展到损伤容限设计，但各种准则各有其不同的应用范围，并不能完全取代。

1）无限寿命设计　无限寿命设计是最早的抗疲劳设计准则，它的出发点是，零件在设计应力下能够长期安全使

用。它的强度条件是，对于应力幅和平均应力不随时间变化的稳定交变应力状态（等幅应力），零件的工作应力小于其疲劳极限；对于应力幅和平均应力随时间变化的不稳定交变应力状态（变幅应力），当交变应力中超过疲劳极限的过载应力数值不大、作用次数很少时，可将这些应力忽略，而按其余次数较多的交变应力中的最大者小于零件疲劳极限的强度条件进行设计。由于零件在疲劳极限的应力作用下具有无限寿命，因此，当零件的设计应力小于疲劳极限时，零件能够长期使用。

对于很多安装在地面上不常搬动的民用机械，对它们的主要要求是使用可靠和寿命长，而对机器的自重常常没有严格限制。对于这些机器往往使用无限寿命设计法。这里，机器的质量虽然比使用有限寿命设计法为重，但由于使用寿命比有限寿命设计法大大增加，从经济上看还是比用有限寿命设计法合算。特别是那些只生产一台或几台的重型机械，机器成本很高，不允许造备品，而一台出了事故，就要长时间停产，造成的损失很大。对于这些机器，为了确保能够长期安全使用，更加应当使用无限寿命设计法。

使用无限寿命设计法时，常常是先用静强度设计确定出零件尺寸，再用这种方法进行疲劳强度校核。

2）名义应力法　有时候需要机器在一定的使用期限内安全使用，这时的设计方法是有限寿命设计法。这种方法允许零件的工作应力超过疲劳极限，机器的质量比无限寿命设计法为小，是当前许多机械产品的主导设计思想。如飞机、汽车等对自重有较高要求的产品，都使用这种设计方法进行疲劳设计。

名义应力有限寿命设计法是一种使用较早的有限寿命设计法，已经使用多年，使用经验比较丰富。这种设计法常称为安全寿命设计，它是无限寿命设计法的直接发展，二者的基本设计参数都是名义应力。其设计思想也大体相似，都是从材料的 $S-N$ 曲线出发，再考虑各种影响因素，得出零件的 $S-N$ 曲线，并根据零件的 $S-N$ 曲线进行抗疲劳设计。所不同的只是，无限寿命设计法使用的是 $S-N$ 曲线的右段——水平部分，亦即疲劳极限；而有限寿命设计法使用的是 $S-N$ 曲线的左段——斜线部分，亦即有限寿命部分。另外，由于斜线部分的疲劳寿命各不相同，因此在对材料 $S-N$ 曲线进行修正时，要考虑循环数对各影响系数的影响。此外，无限寿命设计的设计应力低于疲劳极限，因此，比设计应力低的低应力，对零件的疲劳强度和寿命无影响，设计计算时不管实际的工作应力如何变化，只需按照最高应力进行强度校核即可。如果在最高应力下不会发生疲劳破坏，再加比最高应力为低的其他应力也不会发生问题。而有限寿命设计的设计应力一般都高于疲劳极限，这时就不能只考虑最高应力，而需要按照一定的累积损伤理论估算总的疲劳损伤。

3）局部应力应变法　常规疲劳设计法是以名义应力为基本设计参数，按名义应力进行抗疲劳设计。而实际上，决定零件疲劳强度和寿命的是应变集中（或应力集中）处的最大局部应力和应变。因此，以应变集中处的局部应力、应变为基本设计参数的抗疲劳设计方法称为局部应力应变法。它的设计思路是：零件的破坏都是从应变集中部位的最大应变集中处起始，并且在裂纹萌生以前都要产生一定的局部塑性变形，而局部塑性变形是疲劳裂纹萌生和扩展的先决条件，因此，决定零件疲劳强度和寿命的是应变集中处的最大局部应变。只要最大局部应力应变相同，疲劳寿命就相同。因而，有应变集中的零件的疲劳寿命，可以使用光滑试样的循环应力-应变曲线和应变-寿命曲线进行计算，也可以使用局部应力应变相同的光滑试样进行疲劳试验来模拟。

局部应力应变法估算出的是裂纹形成寿命。这种方法常常与断裂力学方法联合使用，用这种方法估算出裂纹形成寿命以后，再用断裂力学方法估算出裂纹扩展寿命，两阶段寿命之和即为零件的总寿命。

局部应力应变法已经在许多部门广泛应用。这种方法很快得到广泛应用的原因在于：①应变是可以测量的，而且已被证明是一个与低周疲劳相关的极好参数，根据应变分析的方法，就可以将高低周疲劳寿命的估算方法统一起来。②使用这种方法时，只需知道应变集中部位的局部应力应变和基本的材料疲劳性能数据，就可以估算零件的裂纹形成寿命，避免了大量的结构疲劳试验。③这种方法可以考虑载荷顺序对应力应变的影响，特别适用于随机载荷下的寿命估算。④这种方法易于与计数法结合起来，可以利用计算机进行复杂的计算。

4）损伤容限设计　上述抗疲劳设计方法都是以材料内没有缺陷和裂纹为前提的。但是，实际零件在加工制造过程中，由于种种原因，往往已经存在着这样或那样的缺陷。另一方面，随着高强材料或高强结构的广泛应用，零件所受的动应力越来越高，裂纹也往往很快萌生。对于这种有缺陷和有裂纹的零件，仅仅用分散系数来考虑，仍有可能发生事故。为了考虑这些初始缺陷或裂纹的影响，便在断裂力学理论和破损-安全设计原理的基础上，提出了一种全新的抗疲劳设计方法——损伤容限设计。

简单说来，损伤容限设计就是以断裂力学为理论基础，以无损检测技术和断裂韧度的测定技术为手段，以有初始缺陷或裂纹零件的剩余寿命估算为中心，以断裂控制为保证，确保零件在使用期内能够安全使用的一种抗疲劳设计方法。

损伤容限设计允许零件有初始缺陷，或在使用寿命中出现裂纹，发生破损，但在下次检修前要保持一定的剩余强度，能正常使用，直至下次检修时能够发现，予以修复或更换。

损伤容限设计的关键问题，是正确估算裂纹扩展寿命。断裂力学对解决裂纹扩展问题、从而对合理的裂纹扩展寿命估算提供了一条有效的途径，推动了损伤容限设计的发展。

进行损伤容限设计时，必须在结构上采取安全措施，并要有一定的检修制度，确保使用安全，亦即必须对结构进行断裂控制。

5）概率疲劳设计　概率疲劳设计是根据零构件的工作应力与疲劳强度相联系的统计方法而进行的抗疲劳设计方法，是概率统计方法与抗疲劳设计相结合的产物，也称为疲劳可靠性设计。

对于实际使用的零件，应力和疲劳强度数据都有一定的离散性，因而，仅仅按照平均值进行设计是不安全的。这时，必须根据以往的使用经验，引用较大的安全系数。只有当工作安全系数大于许用安全系数时，零件才能安全使用。由于安全系数的选取纯粹是经验性的，没有反映问题的实质，因而这些方法不能作出合理的设计。而概率疲劳设计由于考虑了工作应力和疲劳强度的分布，能够给出可靠度指标，因而可以克服上述缺陷，做出既安全可靠又质量小的合理设计。

3.3.2 抗腐蚀疲劳设计技术和方法

腐蚀疲劳的发生除了与材料有关外，还与应力和腐蚀环境有关，因此，从结构设计上应当考虑降低应力，当表面残余应力为压应力时，耐腐蚀疲劳性能较表面残余拉应力为好；另外还要考虑减少腐蚀的措施。

（1）抗腐蚀疲劳结构设计方法

1）结构形状尽可能简单、合理　形状简单的结构易于采取防腐蚀措施。形状复杂的结构件，往往存在死角、缝隙（图5.5-9）、接头，在这些部位易造成腐蚀介质积存、排放不尽、浓缩等现象，从而引起腐蚀。

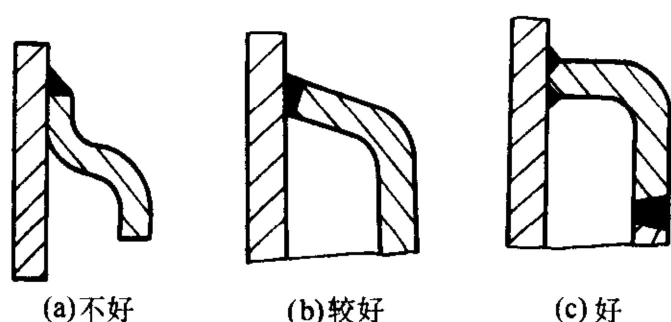


图 5.5-9　避免有缝隙

在无法简化结构的情况下，可考虑将腐蚀严重的部位与其他部位分离的方法，并且使其便于拆卸，以利于维修或更换。

2）防止残存水分或腐蚀介质的腐蚀　对于水溶液腐蚀来讲，没有水分或腐蚀性电解质溶液是不会发生腐蚀的，因此，在设计中，应尽可能地采取各种构型、排液孔等方式，来避免在构件中残存液体，如图 5.5-10、图 5.5-11 所示。

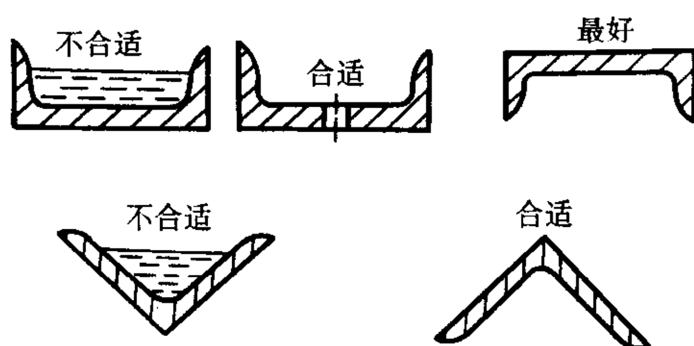


图 5.5-10　几种结构合适和不合适的布置剖面图

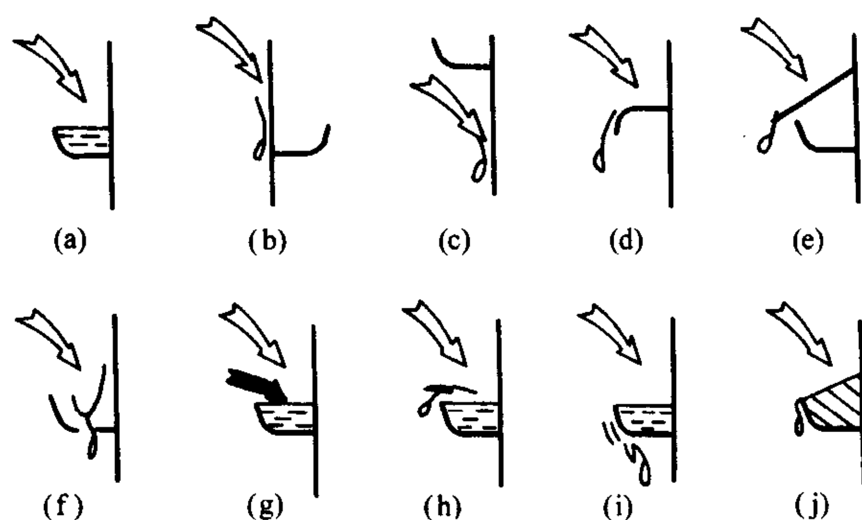


图 5.5-11　防止水分和尘粉积聚的结构设计示意图
(a) 不好，可能积聚水分和尘粒；
(b)、(c)、(d)、(e) 把可能积聚水分的结构置于介质作用之外；
(f) 采用排水孔；(g) 加入缓蚀剂；(h) 保证清扫；
(i) 装自动清水结构；(j) 可积水处用填料填充

3）铆钉联接的抗腐蚀疲劳设计　大量采用铆钉进行联接的结构，联接件应尽量采用对接，避免采用搭接形式。采用对接时只采用薄的刚性垫片就可以减少偏心距，而对于搭接形式的铆钉联接，必须采用刚性加强件（图 5.5-12），才可以提高疲劳强度，但仍不如对接形式好。铆钉的布置通常采用平行排列的方式，不宜采用交错排列的铆接。因为根据“力流线”的概念，并排排列的铆接“力流线”歪曲较小，抗疲劳性能较好。并且，各排铆钉的大小和数量应尽可能相同。从研究铆钉排数对应力集中系数的影响可知，双排与三排铆钉的应力集中系数比单排的小，且疲劳强度高。为了提高疲劳强度应采用双排或三排联接。

4）焊接接头的抗腐蚀疲劳设计　在结构的主传力零件上要避免采用焊接。若要采用焊接时，应把焊缝安排在低应力区。因为焊接材料一般比基体材料疲劳强度低，焊接在基体上造成的热影响区会使疲劳强度下降。由于热胀的不均匀性所造成的残余拉应力，极易使材料出现裂纹。

在焊接中起焊和停焊位置、焊缝端部以及引弧点部位都是常见的腐蚀疲劳破坏部位，因为腐蚀疲劳裂纹首先产生在应力集中区，如图 5.5-13 所示。

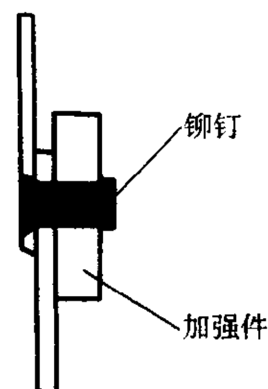


图 5.5-12　搭接铆接处的加强

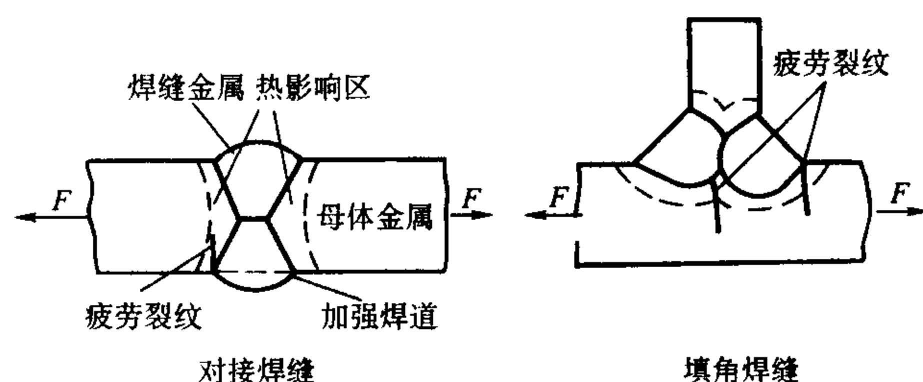


图 5.5-13　疲劳裂纹萌生部位

5）轴键联接抗腐蚀疲劳设计　常和轴结合在一起的元件，如用平键联接的键槽、缺口销或夹紧套筒上的孔，会改变应力分布，而使局部出现应力峰值，在外界腐蚀介质作用下会产生腐蚀疲劳破坏。为了减少应力集中，结构外形尽可能地使应力分布均匀，可采用花键轴或多轴代替平键联接。

6）防止腐蚀疲劳破坏的扰流挡板和缓冲板　应该使溶液里各成分良好混合、均匀分布，在搅拌器里的扰流挡板就起到这个作用。但是搅拌器里旋转的液体对导向板和扰流挡板及容器器壁施加应力，随脉动过程增加的拉应力促进了腐蚀疲劳破坏。

对于列管的管端，如果需要涨管，在涨管过程中，变形的管端出现内应力，操作时还要受到工作应力的作用。如果列管受到侧面液流振动的影响，还会出现附加应力。这些叠加的振动载荷可使锥型涨管区出现腐蚀疲劳裂纹。

为了削除可能产生的腐蚀疲劳裂纹，在涨管区应尽可能减小由操作载荷引起的应力。在侧向液流入口区加一与容器联接在一起的缓冲板，以减轻液流对管子的冲击作用。实际上，在列管式热交换器中，为了消除流动介质以及热膨胀引起的交变载荷而导致的腐蚀疲劳破坏，同样要设计安装导向板。

（2）减少腐蚀的措施

减少腐蚀常用涂层、缓蚀剂及电化学保护。

防腐蚀措施的选择应强调整体性和系统性，即应按设备、装置整体来选择，而不是按单独的组合零部件或单体来考虑。除此之外，防腐蚀措施的选择还应与设备和其主要部件的几何形状、部位以及使用的难易和防腐蚀效果综合起来考虑。尽管有时所采用的防腐蚀措施价格昂贵，并与较廉价的、可大量更新的材料费用或合理更改几何结构所需费用几近相等，但是，前者有利于补偿设备强度的降低，减少经常性维修，使操作更安全可靠。大量实践证明，采用廉价的防腐蚀措施是不经济的。

1）涂层保护措施　金属覆盖层可分为阳极防腐镀层和阴极防腐镀层。镀锌层就是一种阳极防腐镀层，在电化学腐蚀过程中，镀锌层是腐蚀电池的阳极，铁是阴极，受到保护。这类镀层存在孔隙也不影响它的防腐蚀作用。阴极防腐

镀层则不然，例如，镀锡层，在大气中腐蚀时，它是腐蚀电池的阴极，若有孔隙，则加速孔隙处铁基体的腐蚀，因此，镀层制做要求比较严格，不能存在孔隙。

金属覆盖层的制做方法很多，主要有：电镀法、热浸镀法、化学镀法、扩散镀法、喷涂法、化学气相沉积（CVD）法和磁控溅射等。

可以根据表5.5-10选择金属镀层。

表5.5-10 金属镀层的性能及作用

工艺过程	镀层材料	电位	扩散层厚度/μm	用途或限制
渗铝	铝	A	25～152	工厂加工
刷镀	19种金属	V	0.25～152	波导管、工地施工
阴极喷镀	金属陶瓷	V	4	特殊用途
化学还原	钴 铜 镍 钯 金属合金	N N N N V	2.5～25	特殊用途 印刷电路板 特殊用途
爆炸喷镀	金属陶瓷	V	25～305	质量最好，特殊用途硬化表面
扩散渗镀	金属硅酸盐	V	0.25～760	特殊用途
电泳渗镀	铝 钼 镍	A N N	25.4～255	小零件
电镀	铝 镉 铬 铜 黄铜 金 银 铁 铅 镍 铂 钯 铑 锡 锌	A A N N N N N V N N N N N N A	6.5 4～12.5 0.15～510 0.25～760 1.8～2.5 0.75～20 2.5～25 >3 175 6.5～1250 2.5～55 2.5～25 2.5～5 0.25～0.5 5～50 2.5～25	钢丝、薄板、小零件
爆炸黏结	金属	V	1 525～19 050	板、管、钢带
火焰喷涂	铝 锌 锡 金属陶瓷	A A N V	100～205 50～125 75～380 125～1 525	多孔、需密封 多孔、需密封 低熔点合金
熔融黏结	金属	V	1 525～19 050	板、管
电镀	锌	A	12.5～125	最长80 ft（24 m）

续表5.5-10

工艺过程	镀层材料	电位	扩散层厚度/μm	用途或限制
气体渗镀	金属	V	0.25～1 780	特殊用途
浸镀	铜 镍 银 锡 铅	N N N N N	1.25	特殊用途
离子渗镀	金属	V	薄膜	特殊用途
铅渗镀	铅	N	4.7	特殊用途
金属包覆	铝 黄铜 铜 铅 镁 镍合金 钯 铂 银 不锈钢 锡 钛 钽	A N N N A N N N N N N N N	250～760 1 525～19 050 1 525～19 050 1 525～19 050 127～19 050 790～10 160 510～3 175	薄板、板、带、管过渡连接特殊用途过渡连接，特殊用途
氢氧喷涂	镍-铬-铝	N	12.5～25	特殊用途
金属粉末扩散渗镀	铝 镉 铅 锡 锌	A A N N A	50	特殊用途
等离子喷涂	金属陶瓷	V	0.25～2 540	质量较好、高温熔合金属
固体渗锌	锌	A	25～75	小零件
擦镀	金属	V	0.25～150	特殊用途
镀铅锡合金	铅/锡	N	0.25～25	薄钢板
浸锡	锡	N	0.15～30	特殊用途
真空沉淀	金属陶瓷	V	0.25～75	特殊用途
气相沉淀	铝 铬 铁 镍 石墨	A N — N N	12.5～25 2.5～25 25～2 540	特殊用途
机械镀	镉 锡 锌	A N A		特殊用途
表面硬化	金属	V	760～10 160	特殊用途

注：A—阳极性；N—阴极性；V—不定。

在实际中还经常使用一些有机涂料，它是一种有机高分子混合物，用以保护和装饰材料的表面，使其免受外界环境侵蚀。常用的防腐涂料大多数属于树脂类或橡胶类涂料。一些典型涂料的耐蚀性比较见表5.5-11。

表5.5-11　典型涂料的耐蚀性

品种	耐腐蚀性能								
	酸	碱	盐	溶剂	水	耐氧化性	耐候性	耐磨性	耐热性
丙烯酸类	8	8	9	5	8	9	10	10	8
醇酸类	6	6	8	4	8	3	6	6	8
沥青类	10	7	10	2	8	2	4	3	4
氯化烃类	8	8	8	2	3	2	4	3	4
氯化橡胶类	10	10	10	4	10	0	8	6	8
环氧类	9	10	10	9	10	6	8	6	8
环氧-聚酯类	10	1	7	3	7	2	6	6	7
乳胶	2	.	6	1	2	1	10	6	5
含油料类	1	1	6	2	7	1	10	4	7
酚醛类	10	2	10	9	10	7	9	5	10
苯氧基类	3	9	10	5	10	6	4	4	8
硅酮类	4	3	6	2	8	4	9	4	10
乙烯类	10	10	10	5	10	10	10	7	3
氨基甲酸酯类	9	10	10	9	10	9	8	10	8

注：1表示保护性最差；10表示保护性最好。

2）缓蚀剂的选择　在腐蚀介质中添加一些少量的缓蚀剂可以显著地阻止或减缓金属的腐蚀速度。这些缓蚀剂不改变金属构件的性质和生产工艺。在使用时用量很少，一般添加的质量分数在0.1%～1%之间可起到防蚀作用。使用缓蚀剂时方法简单，无需特殊的附加设备。

缓蚀剂的选择应根据金属设备使用的环境来确定。

① 石油工业中使用的缓蚀剂

a）采油和气的时候常采用酸化处理工艺，可以达到增产措施，但是这一措施虽然增加了采油收得率，但对设备的腐蚀也是相当严重的。为了抑制腐蚀作用，常采用有机化合物做为缓蚀剂，如甲醛、咪唑及其衍生物、季铵盐类等。

b）为防止油罐底部沉积水腐蚀常用水溶性缓蚀剂，常用的无机缓蚀剂有亚硝酸盐，当水中含有硫化合物时可以用有机缓蚀剂苯甲酸铵。为防止与油层接触的金属发生腐蚀常用酰化肌氨酸及其衍生物。为防止油罐上部与空气接触的金属腐蚀采用气相防锈剂，常用亚硝酸二环已铵。

c）输油管广泛使用有机化合物做为缓蚀剂，常用的有喹啉、环已胺、吗啉及二乙胺。

② 工业循环冷却水中使用的缓蚀剂　工业用水量最大的是冷却水，约占工业用水量的60%～65%，而在化工、炼油、钢铁等工业则占80%以上。因此，工业生产中大量使用循环冷却水系统，它又分为敞开式和密闭式两种。

a）敞开式循环冷却水系统由于水与空气充分接触，水中含氧量很高，具有较强的腐蚀性。而且，由于冷却水经多次循环，水中的碳酸氢钙等无机盐逐渐浓缩，再加上水中微生物的生长，水质不断变坏。在这种冷却水系统中经常采用重铬酸盐，它是最有效的阳极型缓蚀剂。当水中含有Cu^{2+}等金属离子时，添加聚磷酸盐效果更好。通常聚磷酸盐和铬酸盐混合使用是最佳的复合缓蚀剂。

b）密闭式物质循环冷却水系统的腐蚀环境更为苛刻。采用的缓蚀剂有聚磷酸盐、锌盐、硅酸盐等。亚硝酸铵的浓度达到120×10^{-6}时，具有较好的缓蚀效果，缓蚀率达98%。锌盐也是使用较多的复合缓蚀剂。但锌盐属于有毒物质，用量应限制在排污要求范围内。

③ 大气腐蚀缓蚀剂　大气腐蚀是金属腐蚀最广泛的一种腐蚀。在防止大气腐蚀时使用缓蚀剂要考虑不同环境因素和使用范围。

a）油溶性缓蚀剂。这类缓蚀剂能溶于油中，形成通常所说的防锈油，在制品表面形成油膜，缓蚀剂分子容易吸附于金属表面上，阻滞因环境介质渗入在金属表面上发生的腐蚀过程。各类可溶性缓蚀剂对金属的适应性能如表5.5-12所示。

表5.5-12　各类油溶性缓蚀剂对金属的适应性能

序号	缓蚀剂的种类	对金属的适应性	性　能
1	羧酸类	适用于黑色金属	高分子长链羧酸类，具有防潮性能，复合使用效果更好
2	磺酸类	黑色金属较好，对有色金属不稳定，低分子磺酸盐能使铁表面生成锈斑，分子量在400以上，防锈性能较好	有良好的防潮和抗盐雾性能
3	酯类	与胺并用对黑色金属有效，个别对铸铁有效	作为助溶剂与其他缓蚀剂并用有防潮作用
4	胺类及含氮化合物	适用于黑色和有色金属，对铸铁也有效	耐盐雾、二氧化硫、湿热等性能
5	磷酸盐或硫代磷酸盐	大多数适合黑色金属，一般与其他添加剂并用	抑制油品氧化过程所生成的有机酸，大多数作为辅助添加剂或润滑的缓蚀剂

b）水溶性缓蚀剂。这类缓蚀剂是指以水为溶剂的缓蚀剂。可方便地作为机械加工过程的工序间防锈。大多数的无机盐、如亚硝酸钠、硼酸钠、硅酸钠等是优良的缓蚀剂。它们的优点是节约能源，防锈被膜除去简单、安全，价格便宜。

c）气相缓蚀剂。这种缓蚀剂具有足够高的蒸气压，即在常温能很快充满周围的大气中，吸附金属表面上而阻滞环境大气对金属的腐蚀过程。气相缓蚀剂种类很多，常用的有6类：有机酸类、胺类、硝基及其化合物、杂环化合物及胺有机酸的复合物和无机酸的胺盐。对钢有效的有：尿素加亚硝酸钠、苯甲酸胺加亚硝钠等。对铜、铝、镍、锌有效的有：肉桂酸胍、铬酸胍、碳酸胍等。

3）电化学保护　电化学保护是指通过施加外电动势将被保护金属的电位移向免蚀区或钝化区，以减小或防止金属腐蚀的方法。电化学保护技术目前已广泛应用于舰船、海洋工程、石油及化工等部门。电化学保护按作用原理可分为阴极保护和阳极保护。其中阴极保护可以通过外加电流和牺牲阳极法两种途径来实现。

① 阴极保护方法的主要特点及应用介质　任何金属结构要进行阴极保护，应具备以下条件。

a）环境介质必须是导电的。因为这些介质将构成阴极保护系统的一部分，这样保护电流才能通过导电介质流动，形成一个完整的电回路。在土壤、海水、酸碱盐溶液等介质中都可实施阴极保护，但在气体介质中则不行。气液界面、干湿交替部位的保护效果也不好。在强酸浓溶液中，因保护电流消耗太大，一般也不宜使用阴极保护方法。

b）金属材料在所处介质中应易于阴极极化，否则因消耗的电流大而不适宜采用阴极保护方法。

c）被保护金属结构的几何形状不要过于复杂，否则保护电流分布不均，会出现局部保护不足或保护过度的现象。

阴极保护方法应用介质。

a）淡水及海水中。防止船舶、码头、平台、闸门和冷却设备等的腐蚀。

b）碱及盐类溶液中。防止贮槽、蒸发罐、熬碱锅等的腐蚀。

c）土壤及海泥中。防止管线及电缆等的腐蚀。

② 牺牲阳极保护方法的主要特点

a）不需要外加直流电源，适用于无源地区和小规模、分散的保护对象。

b）驱动电压低、输出功率低、保护电流小且不可调节。阳极有效保护距离小，使用范围受介质电阻率的限制。但保护电流的利用率较高，一般不会造成过保护，对邻近金属设施干扰小。

c）阳极数量较多，电流分布比较均匀。但阳极质量大，会增加结构质量，且阴极保护的时间受牺牲阳极寿命的限制。

d）系统牢固可行，施工技术简单，单次投资费用低，不需专人管理。

③ 外加电流阴极保护方法的主要特点

a）需要外部直流电源。

b）驱动电压高，输出功率和保护电流大，能灵活调节、控制阴极保护电流，阳极有效保护半径大。因此在恶劣的腐蚀条件或高电阻率的环境中也适用。但有可能造成过保护，也可能对附近金属设施造成干扰。

c）阳极数量少、系统质量小、结构质量增加不多。难溶和不溶性辅助阳极消耗低、寿命长，可作长期的阴极保护。但由于系统使用的阳极数量少，保护电流可能分布不均匀。

d）在恶劣环境中，系统易于操作。设备安装、施工、维护较复杂，一次投资费用高。

阴极保护方法的选择应根据供电条件、介质电阻率、所需保护电流的大小、运行过程中工艺条件变化情况、寿命要求、结构形状等决定。通常情况下，对无电源、介质电阻率低、条件变化不大、所需保护电流较小的小型系统，宜选用牺牲阳极保护。相反，对有电源、介质电阻率大、所需保护电流大、条件变化大、使用寿命长的大系统，应选用外加电流阴极保护。

④ 牺牲阳极保护的设计　牺牲阳极设计步骤如下。

a）保护参数的选定　根据被保护结构的材质和工作条件、表面涂层的有无及质量优劣确定保护电流密度和保护电位等保护参数。很多情况下可按有关规定或标准选定。例如，我国国家标准中规定钢质船舶在海水中的保护电位范围为 −0.75～−0.95 V。保护电流密度的大小与被保护金属的种类、腐蚀环境条件、保护系统中电路的总电阻等因素有关，这些因素有时能使保护电流密度在很宽的范围内变化。例如在下环境中保护未加涂层的钢结构，其保护电流分别为：

土壤	10～100 mA/m^2
淡水	20～50 mA/m^2
静止海水	50～150 mA/m^2
流动海水	150～300 mA/m^2

b）保护面积和保护电流的计算　计算被保护结构需要保护的总表面积 S（m^2），并根据表面积 S 和保护电流密度 $j_p(mA/m^2)$ 确定总保护电流 I_p。

$$I_p = j_p S \tag{5.5-15}$$

c）牺牲阳极材料和型号的选择　对牺牲阳极材料有明确规定的情况，设计中应遵守相应规定。若没有明确规定，或提供几种材料供选择，设计者应根据介质的电阻率、被保护结构的材质、涂层状况和保护要求，结合牺牲阳极的特性和货源情况确定牺牲阳极材料。一般镁阳极适用于电阻率较大的介质环境，锌阳极适用于电阻率较低且比较潮湿的土壤环境。在海洋环境中，通常船用牺牲阳极、港口设施用牺牲阳极可以采用锌合金或铝合金阳极。但铝合金和钢结构撞击会产生火花，所以铝合金阳极的使用在某些场合受到一些限制。

d）计算所需的阳极数量。首先根据阳极消耗率、阳极设计寿命和保护电流按下式求出所需阳极的总质量。

$$W = \alpha B I_p T \tag{5.5-16}$$

式中，W 为所需阳极的总质量，kg；B 为阳极消耗率，kg/(A·a)；I_p 为保护电流，A；T 为阳极设计寿命，a；α 安全系数，与阳极形状及使用条件有关，一般可取 1.10～1.25。

根据计算得出的阳极总质量以及单支阳极的质量，即可算出所需阳极的数量。

对于已经标准化、系列化的阳极，各种规格的阳极都有发生电流量的数据，根据阴极保护所需的保护电流 I_p 即可求出所需的阳极数目：

$$N = \beta \frac{I_p}{I_f} \tag{5.5-17}$$

式中，N 为所需牺牲阳极的数目，支；I_p 为阴极保护所需的保护电流，A；I_f 为单支牺牲阳极的发生电流，A/支；β 为遮蔽系数。由被保护设备的结构、阳极间距离和阳极材料的性质等决定，一般取 $\beta = 1.5～3.0$。

对于已标准化的系列产品，I_f 的数值可查阅产品目录。非标准化产品，可利用有关的经验公式计算。

计算得出的阳极数目只是一个大概的数目。实际安装后，还应测量被保护设备的电位及电位分布，并据此调整阳极数目或阳极布置。

e）牺牲阳极的发生电流量的计算。牺牲阳极的发生电流量可利用某些经验公式计算。

$$I_f = \frac{\Delta E}{R} \times 1\,000 \tag{5.5-18}$$

式中，I_f 为牺牲阳极的发生电流量，mA；ΔE 为牺牲阳极的驱动电压，V；选用锌合金阳极时，$\Delta E = 0.2$ V；选用铝合金阳极时，$\Delta E = 0.25$ V；R 为系统电阻，Ω，系统电阻主要为阳极的技术电阻，它决定于介质的电阻率和阳极的几何尺寸。

f）牺牲阳极的寿命计算。牺牲阳极的寿命可按下式计算。

$$T = \frac{mQ \times 1\,000}{8\,760 I_m} \frac{1}{K} \tag{5.5-19}$$

式中，T 为牺牲阳极使用寿命，a；m 为每块牺牲阳极的质量，kg；Q 为牺牲阳极实际发生电量，A·h/kg；I_m 为牺牲阳极平均发生电流量，mA；对于保护年限较长的设施，$I_m = (0.50～0.55)\ I_f$；对于一般的舰船，$I_m = (0.6～0.7)\ I_f$；$1/K$ 为牺牲阳极的利用系数，通常取 85%。

g）复验阳极质量。计算出阳极总质量，如果不能满足要求，应重新选择阳极的规格型号。

h）阳极的布置和安装。根据分散能力和保护半径决定阳极的布置。一般说来，为了使保护电流均匀地分布在被保护的结构上，阳极的分布应该是均匀的。

i）监测系统设计。选定参比电极和电位测量仪表；设计检测点和测量点；制定检测方案。

⑤ 外加电流阴极保护设计

a）保护参数的选择。在外加电流阴极保护设计中，保护电位和保护电流密度是两个最重要的参数。

保护电位域与所选用的参比电极的种类有关。因此应首先确定该保护系统中参比电极的种类，并据此确定相应的保护电位域。

保护电流密度的大小与被保护结构的表面状态及周围的环境条件有关。此外，保护电流密度还与采用的阴极保护方法有关。通常，采用外加电流阴极保护时的保护电流密度要比采用牺牲阳极保护时大得多。

外加电流阴极保护系统的寿命通常要比牺牲阳极的寿命长得多。安装外加电流阴极保护系统的船舶或钢结构，通常其涂层将多年不维修，故随着涂层破坏程度的增加，所需的保护电流密度也将增加。

b）保护面积和保护总电流的计算同牺牲阳极保护设计步骤b)。

c）辅助阳极材料的选择。根据介质腐蚀性、污染规定、辅助阳极结构及分布，结合辅助阳极特性选择合适的辅助阳极材料。

d）确定阳极尺寸及数量。可以先确定阳极的尺寸，再计算阳极的数量；也可以先确定阳极的数量，然后计算阳极的尺寸。

e）决定阳极的布置和安装方式。阳极的布置应考虑满足分散能力的要求，以使电流均匀分布。对于埋地结构的保护，设计中还须选择合适的阳极床位置，它们应该处于：电阻率低的地方；适当地靠近电源；干扰小的地方；电缆少受干扰的地方。

f）参比电极的选择和布置。参比电极可根据被保护结构所处的介质环境条件和参比电极的性能作出适当的选择。位置如果离阳极太近，有可能造成离阳极较远的部件保护不足；离阳极太远，则可能造成阳极附近区域的过保护。

g）电源容量的计算。电源容量主要是指电源的额定输出电流和电压。电源的额定输出电流一般是将计算所得的最大保护电流量加上适当的裕量，并取一个整数。通常阴极保护电源的额定输出电流在100 A以下，额定电压在20～40 V以下。

h）确定控制方式。系统阴极保护要求不高时，可以采用固定槽压法；需要精确控制时，要用恒电位法或恒电流法。

i）根据电流和电压容量、控制方式以及阳极的汇流接线选择直流电源类型、规格型号及台数。

j）设计电联接和绝缘防渗结构。外加电流法辅助阳极与电缆连接处必须进行绝缘防渗处理，否则漏电，造成金属芯的阳极溶解，很快蚀断。

3.4 抗环境失效设计技术和方法

3.4.1 抗蠕变失效设计技术和方法

传统的高温零件强度设计是建立在经典强度理论基础上的，其设计依据是危险截面处的工作应力要小于材料许用应力。根据服役条件不同，其高温性能可选用蠕变极限，以保证零件高温蠕变不超过允许值，或选用持久强度，以保证高温构件在规定条件下不发生断裂，但不限制其变形量。安全系数的选择视情况而定，对10万小时变形1%的蠕变强度设计，其安全系数选1.0～1.25，对铸造合金用1.2～1.5；对10万小时变形合金的持久强度设计，其安全系数用1.5～1.65，对铸造合金选2。

因此，提高金属零件蠕变破断失效抗力和预防措施主要有如下几种：

1）降低零件实际感受温度，如采用隔热涂层等。

2）削除燃烧不均匀现象，避免局部工作温度过高，如正确调整喷嘴雾化效果等。

3）改进设计避免应力集中和局部早期产生微裂纹。

4）制造中严格控制热加工工艺，如GH4037合金二次固化后缓冷等。

5）修理中去除叶片排气边表面缺陷后，应仔细抛光。

6）制造或修理中，采用表面强化或预防措施，如喷丸强化、渗铝等。

3.4.2 抗应力腐蚀设计技术和方法

合理选择材料可以防止金属的应力腐蚀，除此之外，从结构设计上考虑，还有许多防止或减缓应力腐蚀的措施。

（1）减小应力集中的结构设计

应力集中是产生应力腐蚀断裂的最重要的因素，有应力集中的地方往往就是发生断裂的部位。在实际结构中完全避免应力集中是不可能的，但可以在结构设计和加工等过程中尽是减少应力集中的现象。

1）合理的开口形状　结构件中的开口应开在低应力区部位。开口形状和方向对控制峰值应力起到重要的作用。

2）圆角和凸台半径　圆角和凸台在结构上是广泛存在的。要求圆角半径有一定数值，半径大小会导致应力集中而引起断裂，一般规定 $R \geqslant 10a$，如图5.5-14所示。

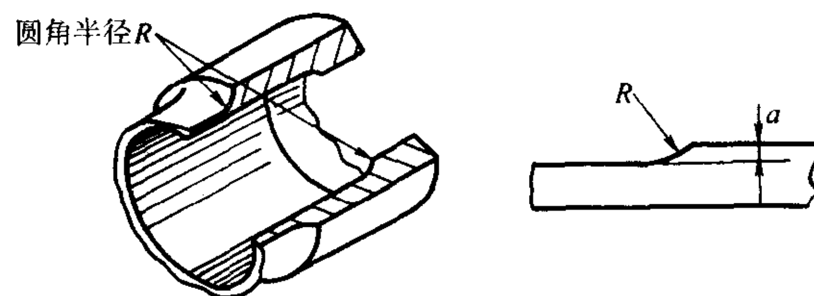


图5.5-14　圆角和凸台的要求

3）铸件截面变化　为了达到等强度的设计目的，要求截面按一定规律变化，截面均匀也可使应力分布均匀，减少应力集中，内外圆角半径 $R > 0.25 \times$ 壁厚，在特别部位应加衬套，如图5.5-15所示。

（2）加工过程应力控制

1）钻孔　结构件的断裂问题约有80%发生在孔边的位置，钻孔技术中最重要的一环是退刀过程，在退刀时产生沿孔轴线的孔壁划伤，这些划伤在板材受拉－拉疲劳载荷时就很可能是裂纹源；改进钻孔方法，在退刀时仍保持钻状旋转，所造成的孔壁划伤方向和受力方向接近平等，因此不易形成裂纹源，从而提高了使用寿命。

2）焊接　如果把不同壁厚的管子焊在一起时，在邻近焊缝的区域内，由于焊接应力与操作应力叠加而出现很高的拉应力，在一定介质作用下就会发生应力腐蚀断裂，所以当不同壁厚的管子需要直接焊在一起时，应将焊接处厚壁管径逐渐减小到与薄壁管径相同，以使焊缝和过渡区分开，如图5.5-16所示。当法兰和容器壁焊接时，如果两者之间留有缝

隙，液体就会渗入其中。当机械载荷很高时应选用使断面变化及焊缝彼此远离的结构，如图 5.5-17 所示。当支架和底板采用对接联接时不能只焊一侧，否则当腐蚀介质进入焊缝时，可能产生应力腐蚀断裂。应该将支架底板完全焊在一起，以采用 K 型焊缝最有利，如图 5.5-18 所示。

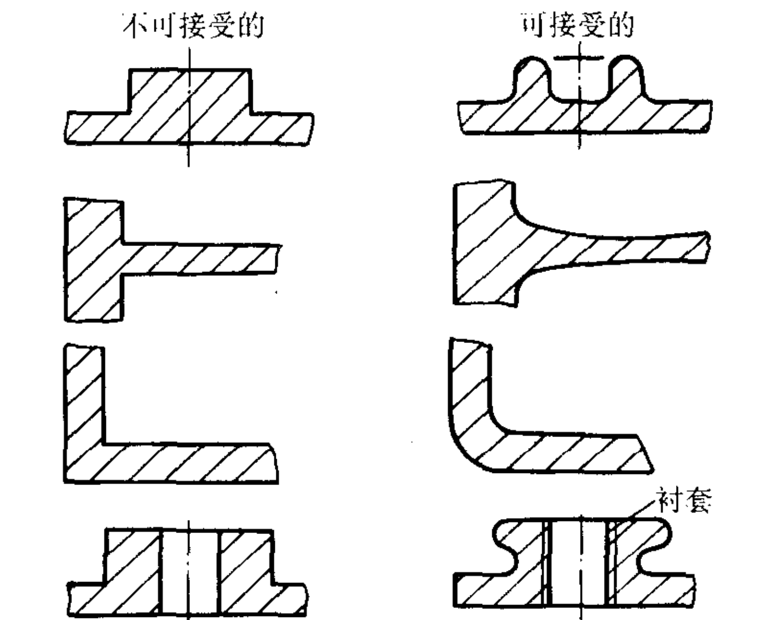


图 5.5-15　截面变化设计

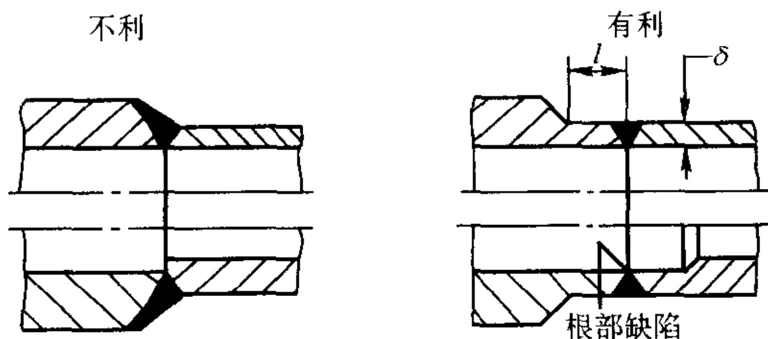


图 5.5-16　不同壁厚的管件的焊接

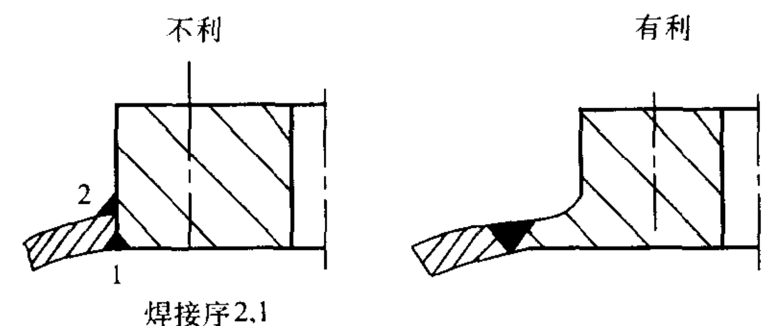


图 5.5-17　法兰与容器的焊接

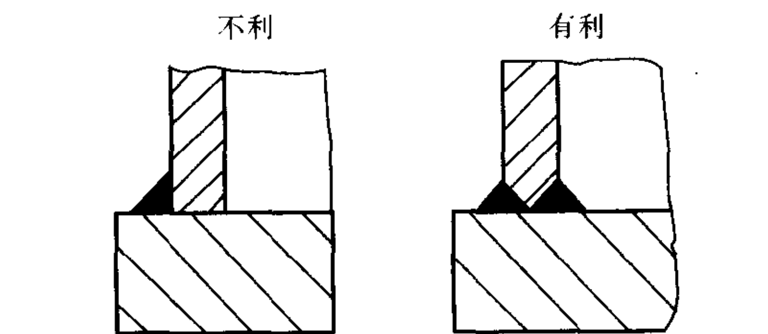


图 5.5-18　K 型焊接

贯穿式接管如果没有焊透就会出现缝隙，引起缝隙腐蚀。又因焊接后，由于接管收缩受到限制，在接管内产生很高的对称分布的径向拉应力，引起应力腐蚀断裂，所以必须焊实，不留焊缝。对于厚壁件采用 K 型焊接为好，焊接后还应退火消除应力，如图 5.5-19 所示。

对于不同厚度的板材焊接时，应使厚度尽量相等，否则较薄的板材会发生变形产生应力。

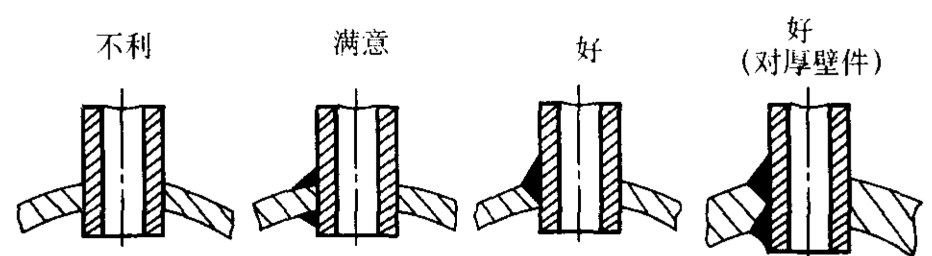


图 5.5-19　接管焊接

设计构件外形时应使弯曲度尽可能的小，以免产生较大应力，弯曲半径大，出现的拉应力小，弯曲半径应大于板材厚度。

（3）减小结构偏心

偏心的害处主要是诱发出二次应力，所以应当尽量避免。

（4）避免复合应力集中

复合应力集中是指两个应力集中区域靠在一起或靠得很近的情况。它所产生的总效果趋近于两个应力集中系数的乘积，因此造成断裂的危险。复合应力集中常出现在骑缝螺钉处。

4　可靠性设计技术和方法

4.1　可靠性设计基本概念

为了确保产品（零件、设备或系统）的可靠性，必须制订完整的可靠性计划。从产品的生产计划开始，经过设计、研制、制造、装配、试验、使用、维修、故障反馈，一直到报废，所有的阶段都应当有可靠性保证措施。其中，设计正确，确保设计阶段的可靠性是最重要的一环。因为生产是按照设计进行的，设计决定了产品的固有可靠性，而在使用阶段达到的使用可靠性永远不可能超过设计阶段决定了的固有可靠性。

4.1.1　可靠性设计的基本特点

（1）可靠性设计的概念

为了达到产品的可靠性指标而进行的设计称为可靠性设计。可靠性设计的目的是在达到产品的可靠性指标的前提下，配合产品的价值工程设计，尽可能降低产品在寿命周期内的总费用。通过可靠性设计可以在设计过程中挖掘和确定隐患和薄弱环节，并采取设计预防和设计改进措施有效地消除隐患和薄弱环节。

可靠性设计的特点。

1）与传统的设计方法中将安全系数作为惟一的评价项目和度量完全不同，可靠性设计要求对不同的产品、场合采取不同的可靠性指标，如可靠度、失效率、MTBF、维修度、有效度等。

2）可靠性设计强调在设计阶段要把可靠度直接设计到零件中去。即由设计决定固有可靠度，由制造保证固有可靠度。

3）可靠性设计必须考虑环境的影响。高温、低温、冲击、振动、潮湿、盐雾等环境条件对应力有很大影响，而研究表明，应力分布的尾部比强度分布的尾部对可靠度的影响大得多。

4）对维修性的考虑。在浴盆曲线的耗损失效期及当有效度是主要可靠性指标时，都必须考虑维修性。设计决定了机械产品的固有可靠性，而环境、使用、维修等则影响了使用可靠性。

5）可靠性设计承认在设计阶段及其以后的阶段都需要可靠性增长。即随着产品设计、研制、生产各阶段工作的逐步进行，产品的可靠性特征量值也逐步提高。

（2）可靠性设计与传统设计的异同点

1）可靠性设计与传统设计方法相同点　传统的机械设计采用确定的许用应力法和安全系数法研究、设计机械零件和简单的机械系统。这是广大工程技术人员很熟悉的设计方法。而机械可靠性设计，又称机械概率设计是以非确定性的随机方法研究、设计机械零件和机械系统。它们共同的核心内容都是针对所研究对象的失效与预防问题，建立起一整套的设计计算理论和方法。在机械设计中，不论是传统设计还是可靠性设计，判断一个零件是否安全都是将引起失效的一方，如零件中的载荷、应力或变形等，与抵抗失效能力的一方，如零件的许用载荷、许用应力或许用变形等，加以对比来判断。

如果引起零件失效的一方用“应力”s表示，可用一多元函数来描述，即

$$s=f\ (s_1,\ s_2,\ \cdots,\ s_n) \tag{5.5-20}$$

式中，s_1，$s_2\cdots$，s_n 表示影响失效的各项因素，如力的大小、力的作用位置、应力集中、环境因素等。

若抵抗失效能力的一方用“强度”r表示，也可用一多元函数表示，即

$$r=g\ (r_1,\ r_2,\ \cdots,\ r_n) \tag{5.5-21}$$

式中，r_1，$r_2\cdots$，r_n 表示影响零件强度的各项因素，如材料性能、表面质量、零件尺寸等。

上面提到的“应力”和“强度”实际是广义的，当 $r-s>0$，表示零件处于安全状态，当 $r-s<0$，表示零件处于失效状态，当 $r-s=0$ 时，零件处于极限状态。因此，传统的机械设计和机械可靠性设计的共同设计原理可表示为

$$f\ (s_1,\ s_2,\ \cdots,\ s_n)\leqslant g\ (r_1,\ r_2,\ \cdots,\ r_n) \tag{5.5-22}$$

上式表示了零件完成预期功能所处的状态，因此称为状态方程，或称为工作能力方程。不论传统的机械设计或机械可靠性设计，都是以式（5.5-22）所表示的零件或系统各种功能要求的极限状态和安全状态作为设计依据，以保证零件在预期的寿命内正常运行。

2）可靠性设计与传统设计方法不同点

① 设计变量处理方法不同　传统的机械设计，把影响零件工作状态的设计变量，如应力、强度、安全系数、载荷、零件尺寸、环境因素等，都处理成确定性的单值变量。而描述状态的数学模型，即变量与变量之间的关系，可通过确定性的函数进行单值变换，这种把设计变量处理成单一确定值的方法，称为确定性设计法，图5.5-20表示这种确定性设计法的模型。

可靠性设计把设计中所涉及的变量，都处理成多值的随机变量，它们都服从一定的概率分布，这些变量间的关系，可通过概率函数进行多值变换，得到“应力”和“强度”的概率分布，这种运用随机方法对设计变量进行描述和运算的方法，称为非确定性概率设计法，图5.5-21表示了这种非确定性概率设计模型。

② 设计变量运算方法的不同　在传统的机械设计中，有一受拉力作用的杆件，则横断面上的正应力为

$$s=\frac{F}{A} \tag{5.5-23}$$

上式表示了拉力 F、横断面积 A 和应力 s 之间确定性的函数关系，变量之间通过实数代数运算，可得到确定性的单值变换。

在机械可靠性设计中，由于设计变量是非确定性的随机变量，因此，它们均服从一定的分布规律，用概率函数及分布参数（如随机变量的均值和标准差）来表征。于是式（5.5-23）可写成

$$s\ (\mu_s,\ \sigma_s)=\frac{F\ (\mu_F,\ \sigma_F)}{A\ (\mu_A,\ \sigma_A)} \tag{5.5-24}$$

式中，μ_s、σ_s 为应力 s 的均值和标准差；μ_F、σ_F 为力 F 的均值和标准差；μ_A、σ_A 为面积 A 的均值和标准差。

式（5.5-24）表示非确定性随机变量的数字特征之间的函数关系，可运用随机变量的组合运算规则，得到变量与函数间的多值变换。

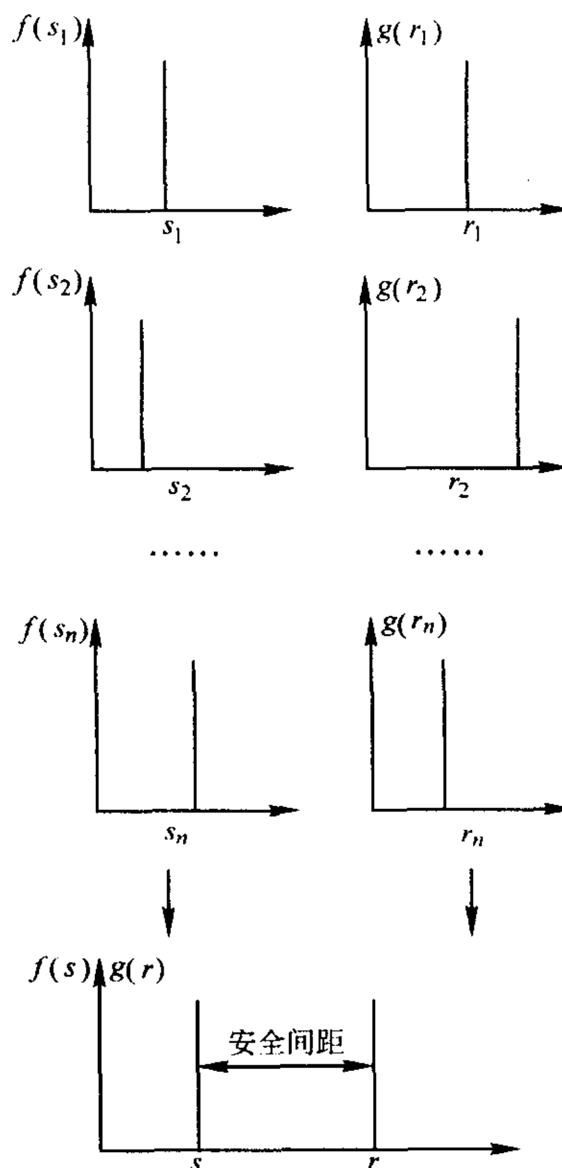


图5.5-20　确定性设计方法

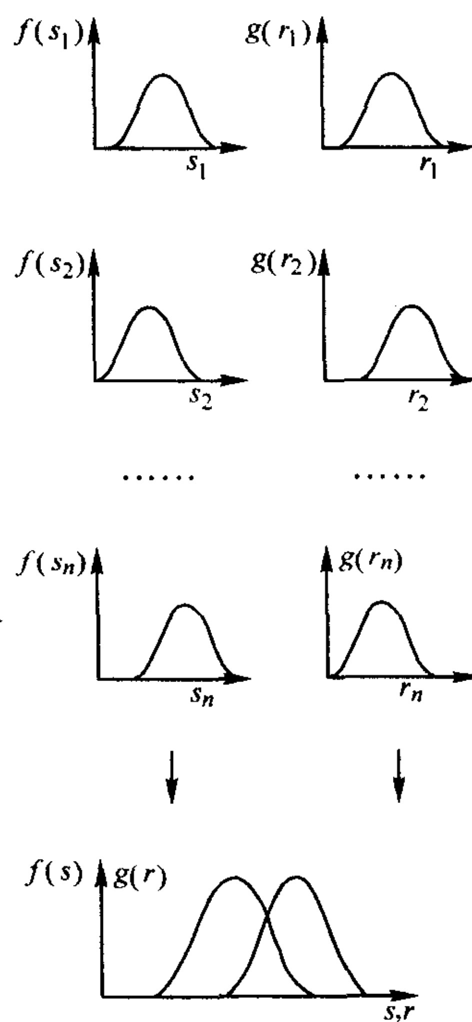


图5.5-21　非确定性概率设计法

③ 设计准则含义不同　在传统的机械设计中，判断一

个零件是否安全，是以危险断面的计算应力 σ_{ca} 是否小于许用应力［σ］，计算安全系数 n 是否大于许用安全系数［n］来决定，相应的设计准则为

$$\begin{cases}\sigma_{ca}\leqslant[\sigma]\\ n\geqslant[n]\end{cases}\tag{5.5-25}$$

上式表示零件的强度和安全程度，是一个确定不变的量，未能定量反映零件强度的许多非确定因素，因而不能回答零件在运行中有多大可靠程度。

在可靠性设计中，由于应力 s 和强度 r 都是随机变量，因此，判断一个零件是否安全可行，是以强度 r 大于应力 s 所发生的概率来表示。其设计准则为

$$R(t)=P(r>s)\geqslant[R]\tag{5.5-26}$$

式中，$R(t)$ 表示零件在运行中的安全概率，即可靠度。它是指零件在工作时间 t 内的一种能力，这种能力是以“强度” r 超过“应力” s 的概率来度量，显然它是零件工作时间 t 的函数。式中［R］称为零件的许用可靠度，它表示零件在规定的时间内、规定的条件下实现设计要求的一种能力，即许用安全概率。上式不仅能定量地回答零件在运行中的安全、可靠程度，而且可以预测零件的寿命。

传统的设计方法，一直延用至今，积累了大量数据，其设计准则，表达形式简单、直观明确，应用方便，因此为广大工程技术人员所熟练地采用。但是，由于它停留在确定性的概念上，没有考虑事物的不确定性质，因而不能真正反映客观实际情况，而且计算中只要安全系数大于某一实际使用经验规定的数值，就认为是安全的，这个规定的安全系数，与系数无法定量表示的因素有关，实际上仍是一个“未知”系数。因此，这种计算有较大的经验性与盲目性，有时取强度和应力的平均值算得的安全系数，虽然大于 1，但是，由于强度和应力数值是离散的，在应力 s 和强度 r 分布曲线下面的干涉区内，有时会出现应力大于强度的情况，实际上并不能保证在任何情况下都安全。为了追求安全，设计中有时则盲目取用优质材料或加大零件尺寸，形成不必要的浪费。而可靠性设计考虑到工程中发生的现象及其表征参数的不确定性（随机性），这是符合客观实际的。

从传统的设计准则变换到可靠性设计准则，这是设计理论的发展、设计概念的深化。可靠性设计以随机方法（概率论和数理统计）分析研究系数和零件在运行状态下的随机规律和可靠性，不仅更能揭示事物的本来面貌，而且能较全面地提供设计信息，这是传统设计方法无法做到的。理论分析和实践表明，可靠性设计比传统设计，能有效地处理设计中的一些问题，提高产品质量，减小零件尺寸，从而节约原材料，降低成本，带来较大的经济效益。

由于传统的机械设计与机械可靠性设计，都是以零件或机械系统的安全与失效，作为其主要研究内容，因此，这两种设计方法又是密切联系着的，可以说机械可靠性设计是在传统的机械设计基础上，补充了可靠性特殊技术的一种新型设计，也是处在传统设计延长线上的一种新的、靠得住的设计。由于我国可靠性设计的数据还比较缺乏，而它的开展又是一项长期、费钱的工作。因此，应该将传统的机械设计和机械可靠性设计有机地结合起来，以丰富发展机械设计理论，提高机械产品的设计水平。

4.1.2　可靠性设计的主要内容

（1）研究产品的故障物理和故障模型

在本篇第 2 章中已经介绍了材料常见的故障模式和分类，本处不再详细介绍。

产品的故障形成过程就是产品运动特性输出参数偏离规定值的过程。偏离规定值越远，故障就越严重，一旦运动无输出，则就完全丧失了规定功能。

在输出参数变化过程中，发生故障的可能性与实际参数接近极限状态的程度有关。产品输出参数极限值，一般是根据使用需要人为地规定的，而产品在工作或试验过程中输出参数逼近这个极限值的经过是不同的，大致可以分成三种情况。

图 5.5-22a 是在考察时间 $t=T$ 内，输出参数随时间的变化 $x(t)$ 局部地达到了极限状态，其故障发生的概率，可用分布 $f(t=T)$ 表示。可靠性要求越高，则允许的 $t=T$ 所对应的 $x(t=T)$ 值应越小于极限值 X_{max}。

图 5.5-22b 是工作一段时间后输出参数的变化趋向某一稳定值。$x(t)$ 不再随时间变化或变化的速度不大，条件 $x\leqslant X_{max}$ 始终可以满足，表明产品有足够的可靠性储备，不会发生故障。

图 5.5-22c 为介于上述两种情况之间的状态，随着使用时间增加，输出参数已趋稳定，但有的局部值有可能超过极限值，继续工作又能回到许可值。这种情况，在局部超过极限值时可能引起故障，此时需采取措施来恢复其工作能力。

图 5.5-22a 属于渐进性故障模型，图 5.5-22c 是突发性故障模型，而图 5.5-22b 是不允许发生故障，应以可靠性储备作为可靠性特征来确定的产品可靠性。

正态分布可用于解决磨损或其他渐进性故障，在遇到非对称分布时也可用对数正态分布、威布尔分布、伽玛分布等。在处理突发性故障时，可选用指数分布或均匀分布。

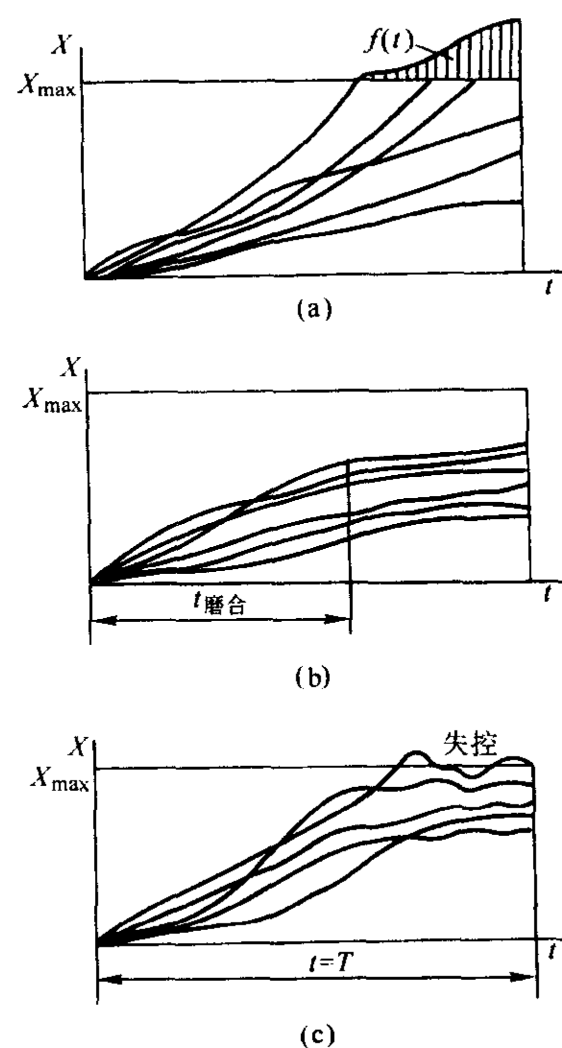


图 5.5-22　输出参数随时间变化曲线

（2）研究产品的可靠性指标及其量值的确定

反映产品可靠性水平高低的可靠性指标有许多个，例如可靠度（R）、失效率（λ）、平均无故障工作时间（MTBF）和有效度（A）等，这些单个的指标只是从不同侧面反映产品的可靠性水平，如可靠度（R）是产品在某一时间内不失效的概率，而 MTBF 是产品在寿命期内平均无故障工作时间，如果两产品的平均无故障工作时间相等，其可靠度不一定相同。

采用什么可靠性指标，取决于产品的设计要求，而可靠性指标的大小则取决于产品的重要性。这里，要重视过去的

经验、服务用户的要求及市场调查。一般来讲，对于不可修复或难修复的产品，常采用可靠度为其可靠性指标。

选择好产品的可靠性指标后，需要确定这些指标的量值。可靠性指标定得过低，则不能满足使用要求及完全失去使用价值，甚至会造成严重后果。如果可靠性指标定得过高，从使用角度来说，虽然是有利的，但会造成额外的经济损失，还会延长工程周期。

确定产品可靠性指标的量值，可以采用以下三种方法。

1）参照同类产品的可靠性指标来确定。对于工程机械，通常规定其有效度 $A(t)=0.90$，对于机床数控系统，一般可取 MTBF = 3 000 h。

2）由产品的可靠性分析模型和可靠性预测方法来预测产品的可靠性指标值，再由预测值来确定产品的可靠性指标的量值。在确定指标量值时，一定要留有一定的裕量。这是因为被预测的可靠性指标要经过系统研究、设计阶段一系列可靠性增长试验，不断排除事先难以预见或考虑不周的各种设计、工艺缺陷，都能逐步实现。一开始就把预测量作为产品的可靠性指标值，这样的指标一般是达不到的。

3）根据全寿命周期费用模型来确定产品的可靠性指标量值。产品的可靠性、维修性与寿命期总费用的关系可以由图 5.5-23 来描述。从图 5.5-23 可以看出，为提高产品可靠性而花的费用随着产品可靠性量值的增大而增大。相反，产品的使用和维修费用则相应降低。因此存在一个最佳的产品可靠性量值 R_{opt}，使该产品的寿命期总费用最小。把 R_{opt} 规定为产品的可靠性指标量值，使产品具有较好的经济性。这个方法的缺点是所需要的信息量大。

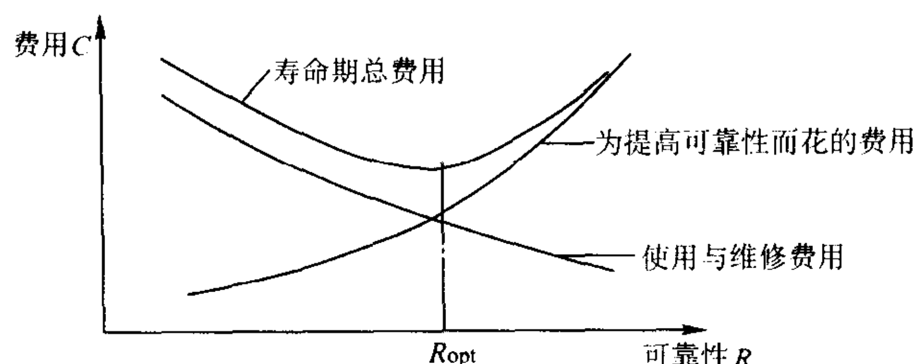


图 5.5-23 可靠性、维修性与寿命期总费用的关系

(3) 合理分配产品的可靠性指标值

可靠性分配是将规定的产品可靠性指标合理地分配给部件、零件的方法。这是一个由整体到局部的分解过程。可靠性分配的目的是建立每个零部件的可靠性指标，使设计和研制人员知道零部件所需的可靠性指标。可靠性指标与费用、零件的功能、复杂程度、体积、质量、研制进度等有关。这些都是对可靠性指标的约束条件。设计和研制人员的任务是，在规定的约束条件下求得最优的可靠性目标值。

随着目标函数和约束条件的不同，可靠性分配的方法也会不同。有时以可靠度指标为约束条件，给出系统要求达到的可靠度值，而以在这一限制下，使重量、成本等其他的系统参数尽可能小，作为目标函数；有时则给出质量、成本等的界限值，要求作出使系统可靠度尽可能高的分配。一般应根据系统的用途，根据优先考虑的参数来选定分配方法。

可靠性分配方法有多种，如代数法、拉格朗日乘数法、动态规划法等。

1）代数法分配可靠度

① 串联系统的可靠性指标分配　按相对失效概率进行可靠性指标分配方法的基本出发点是：每个单元的容许失效概率正比于预计的失效概率。

分配的具体方法和步骤如下。

a）从适当的失效率表查出单元的基本失效率 λ_c，定出失效率的修正系数 K_F，从而求出单元的应用失效率 $\lambda=\lambda_c\times K_F$。

b）估计单元在系统工作期中的实际工作时间 t_i，计算出单元的预计可靠度

$$R'_i=e^{-\lambda_i t_i} \tag{5.5-27}$$

及预计失效概率

$$F'_i=1-R'_i \tag{5.5-28}$$

c）算出单元的相对失效概率

$$w_i=F'_i/\sum F'_i \tag{5.5-29}$$

d）根据系统要求的可靠度指标 R_s，算出系统容许的失效概率 $F_s=1-R_s$。

e）求出单元的容许失效概率

$$F_i=w_iF_s=\frac{F'_i}{\sum F'_i}F_s$$

f）最后得到单元分配到的可靠度值 $R_i=1-F_i$。

这种分配方式是普遍提高了单元现有预计水平的可靠度。较好的做法是，当预计的失效概率超过了容许失效概率时，就对关键单元进行改进，或改进系统设计方案，以提高系统的可靠度来达到要求的可靠度。

② 冗余系统的可靠度分配　系统复杂时，把系统可靠性指标分配给各单元，不仅费时，而且得不出一个简单、直接的答案，需要一个迭代过程。

分配过程中的第一步是把每组并联单元适当组合成单个单元，然后将此单个单元视为一个串联单元方框。

并联单元组合好后，系统成了串联系统，便可以用上面的方法将系统的容许失效概率分配给各个串联单元。然后确定并联分支中每个单元的容许失效概率。

若由 n 个并联单元组成的串联方框分配到容许失效概率为 F_A，则

$$F_A=F_1F_2\cdots F_n=\prod_{i=1}^{n}F_i \tag{5.5-30}$$

式中，F_i 表示各个并联单元的容许失效概率。

利用各并联单元已求得的预计失效概率 F'_i，可建立 $n-1$ 个相对关系式：

$$\left.\begin{aligned}F_2&=\frac{F'_2}{F'_1}F_1\\F_3&=\frac{F'_3}{F'_2}F_1\\&\vdots\\F_n&=\frac{F'_n}{F'_1}F_1\end{aligned}\right\} \tag{5.5-31}$$

解式 (5.5-30) 和式 (5.5-31)，可求得各并联单元应分配到的容许失效概率值 F_i。

如果冗余单元在一级以上，则此方法要一级一级地继续使用，直到各个单元得到容许失效概率正比于它的预计失效概率为止。这一方法要一级一级地重复。

2）拉格朗日乘数法分配可靠度　拉格朗日乘数法是求多元函数条件极值的一种数学，用它来求解多元函数自变量有附加条件的极值问题。

拉格朗日乘数法分则为：

欲求 n 元函数 $F(x_1, x_2, \cdots x_n)$，在 m 个 $(m<n)$ 附加条件（式 5.5-32)

$$\left.\begin{aligned}G_1(x_1, x_2, \cdots, x_n)&=0\\G_2(x_1, x_2, \cdots, x_n)&=0\\&\vdots\\G_m(x_1, x_2, \cdots, x_n)&=0\end{aligned}\right\} \tag{5.5-32}$$

下的可能极值点，可以用常数 1，λ_1，λ_2，…，λ_m 顺序乘 F，G_1，G_2，…，G_m，把结果加起来得函数：

$$H(x_1, x_2, \cdots, x_n) = F + \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_m G_m \quad (5.5\text{-}33)$$

然后写出 $H(x_1, x_2, \cdots, x_n)$ 无附加条件时具有极值的必要条件：

$$\left.\begin{aligned}
\frac{\partial H}{\partial x_1} &= \frac{\partial F}{\partial x_1} + \lambda_1 \frac{\partial G_1}{\partial x_1} + \lambda_2 \frac{\partial G_2}{\partial x_1} + \cdots + \lambda_m \frac{\partial G_m}{\partial x_1} = 0 \\
\frac{\partial H}{\partial x_2} &= \frac{\partial F}{\partial x_2} + \lambda_1 \frac{\partial G_1}{\partial x_2} + \lambda_2 \frac{\partial G_2}{\partial x_2} + \cdots + \lambda_m \frac{\partial G_m}{\partial x_2} = 0 \\
&\vdots \\
\frac{\partial H}{\partial x_n} &= \frac{\partial F}{\partial x_n} + \lambda_1 \frac{\partial G_1}{\partial x_n} + \lambda_2 \frac{\partial G_2}{\partial x_n} + \cdots + \lambda_m \frac{\partial G_m}{\partial x_n} = 0
\end{aligned}\right\} \quad (5.5\text{-}34)$$

式（5.5-32）由 n 个方程组成，式（5.5-34）由 m 个方程组成，所以联立可解出 $n+m$ 个未知数，即 $x_1, x_2, \cdots, x_n$ 及 $\lambda_1, \lambda_2, \cdots, \lambda_m$，而 $x_1, x_2, \cdots, x_n$ 就是可能为极值点的坐标。

利用上述拉格朗日乘数法，可以处理有约束条件下的可靠度分配问题。

3）用动态规划法分配冗余度

若 R 是 x 的函数，并可以分解为：

$$R(x) = f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) \quad (5.5\text{-}35)$$

那么，在 $x_1 + x_2 + \cdots + x_n = x$ 的条件下，使 $R(x)$ 为最大的问题就称为规划。这里，x_i 为任意正数，n 是整数。

因为 $R(x)$ 的最大值是由 n 和 x 来决定的，所以可以把它写成

$$\varphi_n(x) = \max_{x \in \Omega} R(x_1, x_2, \cdots, x_n) \quad (5.5\text{-}36)$$

Ω 是满足式（5.5-36）的解的集合。

如果在第 n 次活动中分配到的 x 量为 x_n（$0 \leqslant x_n \leqslant x$），由 x_n 得到的函数值为 $f_n(x_n)$，则根据式（5.5-36），由 x 的其余部分 $x - x_n$ 所能得到的函数值的最大值为

$$\varphi_{n-1}(x - x_n) \quad (5.5\text{-}37)$$

所以，在第 n 次活动中分到 x_n，在其他活动中分到 $x - x_n$ 时的总函数值为

$$f_n(x_n) + \varphi_{n-1}(x - x_n) \quad (5.5\text{-}38)$$

因为求使这一总函数值为最大的 x_n 是与使 $\varphi_n(x)$ 为最大有关的，所以

$$\varphi_n(x) = \max_{0 \leqslant x_n \leqslant x} [f_n(x_n) + \varphi_{n-1}(x - x_n)] \quad (5.5\text{-}39)$$

也就是说，虽然要对 1，2，…，n 一共 n 个进行分配，但没有必要同时对所有组合进行研究；在 $\varphi_{n-1}(x - x_n)$ 已最佳分配之后来考虑总体的函数值，就只需注意 x_n 值就行了。另外不管怎样选择 x_n 值，若要使总体的函数值为最大，也必须作使 $x - x_n$ 的值成为最大那样分配。这种方法通常称为最佳性原理。

对于复杂系统的可靠性分配问题，采用动态规划法可以大大减少计算次数，较快获得最佳分配方案。且由于动态规划是使用递推式，计算逻辑较简单，适用于电子计算机的计算程序。因此，可靠性工程中将得到进一步应用。

（4）以规定的可靠性指标值为依据对零件进行可靠性设计

把规定的可靠性指标直接设计到零部件中去。零件的可靠性设计，一般是先规定可靠性指标的目标值，再求零件的主要断面尺寸。如果是对现有的零部件、设备或系统进行可靠性特征量计算，则属于可靠性分析。

1）静强度分布数据　只承受静载的零件进行可靠性设计时需要考虑静强度指标，静强度主要包括抗拉强度和屈服强度。大量的试验研究表明，抗拉强度通常是正态分布或近似正态分布，而屈服强度通常为近似正态分布。

2）有限寿命下的强度分布数据　将规定寿命和应力水平下的试验数据画成强度直方图，然后可用常规的统计方法对其进行分布拟合，并确定其分布参数，这样就可得到有规定寿命下的强度分布。

3）长寿命下的强度分布数据　长寿命下的疲劳强度通常为正态分布。

4.2 可靠性设计方法

4.2.1 失效树分析法

失效树分析（Fault Tree Analysis）是 1961 年美国贝尔电话研究所的 H.A.Watson 首先提出的，用于对“民兵”导弹的发射控制系统的可靠性分析。后来，波音飞机公司进一步将其用于飞机的可靠性和安全性分析。1965 年，D.F.Haasl 等在波音飞机公司的系统安全学术年会上首次提出了关于失效树分析的论文，引起产业界和学术界的重视。此外，失效树分析的理论和应用发展得很快。上个世纪 70 年代初，发展了计算机辅助建树和分析程序。1974 年，美国原子能管理委员会（AEC）批准公布了“反应堆安全性研究”（WASH－1 400）报告。这个长达 3 000 页的报告是以 N.Rusmussen 教授为首的 60 名专家耗资 300 万美元、用两年时间对 100 所核反应堆进行安全性和风险分析的结果。报告被认为是失效树分析发展史上的重要里程碑。之后，FTA 作为系统可靠性和安全性分析的有效方法不仅在许多工业领域内得到推广，而且也应用于社会经济管理问题。

失效树分析法是“在系统设计过程中，通过对可能造成系统失效的各种因素（包括硬件、软件、环境、人为因素）进行分析，画出框图（即失效树），从而确定系统失效原因的各种可能组合方式或其发生概率，以计算系统失效概率，采取相应的纠正措施，以提高系统可靠性的一种设计分析方法”（GB/T 3187—1994）。

失效树分析的特点是：

1）失效树分析是一种图形演绎方法，是失效事件在一定条件下的推理方法。它可以围绕特定的氯化镁状态作层层深入的分析，因而在清晰的失效树图形下，表达了系统的内在联系，并指出零部件失效与系统失效之间的关系，从而可找出系统的全部故障谱，找出系统的薄弱环节。

2）它能考虑可能造成系统失效的各种因素，不仅可以分析某些零部件失效对系统的影响，还可考虑软件的、环境的和人的行为等因素。

3）失效树建成后，对不曾参与系统设计的管理和维修人员来说相当于一个形象的管理、维修指南，因此可用于培训使用系统的人员和用于检查事故发生的原因。

4）通过失效树分析，可以定量地计算复杂系统的失效概率和有关的可靠性参数，为改善和评估系统可靠性提供定量数据。

5）失效树分析比较复杂，工作量大，必须使用计算机。

失效树分析的缺点是：建树时需要考虑的因素太多，耗费人力和时间，而且容易失察和出错；缺乏数据。

失效树分析的主要步骤是：选择和确定项事件；自上而下地建造失效树；失效树的定性分析；失效树的定量计算。

建树的方法有以下几种：①演绎法；②合成法；③决策表法。第一种是靠人工来建树，后两种是用计算机辅助建树。

失效树建立方法请参考本篇第 2 章内容。

4.2.2 失效模式、影响及致命度分析法

失效模式、影响及致命度分析（Failure Mode，Effect and Criticality Analysis，FMECA）是“在系统设计过程中，通过对

系统各组成单元潜在的各种失效模式及其对系统功能的影响，与产生后果的严重程度进行分析，提出可能采取的预防改进措施，以提高产品可靠性的一种设计分析方法”。

FMECA起源于1950年，当时美国的格鲁曼（Grumman）飞机公司首先把失效分析模式及影响分析（FMEA）用于战斗机操纵系统的设计分析上，以后又逐步用于不同的机械产品上，这种方法易于掌握、使用方便、不需要很多的数据。FMECA是FMEA的发展。一般，FMEA只进行定性分析，而FMECA可以进行定量分析。美国军用标准MIL-STD-785B（1980）中指出：“这个分析应当与设计努力同时安排和完成，使设计反映分析的结果和建议。FMECA的结果和现状应当用来作为设计综合评定、安全性工程、维修性工程、后勤支援分析……的输人信息。”

以前认为FMECA主要用于设计评审之前，而近来年认为应当反复进行几次。

FMEA和FMECA的内容及实施步骤如下。

1）按照系统功能图画出可靠性框图，弄清系统的所有零件及其功能。

2）确定分析的范围，列出每个零件明显的和潜在的失效模式及影响。

3）研究如何检测各种失效模式的方法。

4）针对各种失效模式，找出失效原因，提出可能的预防措施。

5）估计各种失效模式的发生概率，并计算致命度。

小松制作所采用的估计致命度的公式如下：

$$C_B = F_1 F_2 F_3 F_4 F_5 \tag{5.5-40}$$

式中，C_B为致命度评分值；F_1为对系统损害程度系数；F_2为故障对系统影响程度系数；F_3为故障发生频率系数；F_4为预防故障难易程度系数；F_5为是否重新设计系数。

F_1、F_2、F_3、F_4、F_5的荐用值列于表5.5-13中。

表5.5-13　F_1、F_2、F_3、F_4、F_5的荐用值

项目	含　义	系数F值
F_1（损害程度）	有致命损伤的故障 造成相当大损失的故障 不失掉功能的故障	5.0 3.0 1.0
F_2（对系统影响程度）	对系统有两个以上的重大故障 对系统有一个以上的重大故障 对系统无重大影响	2.0 1.0 0.5
F_3（发生频率）	故障发生的频率较高 故障有可能发生 故障发生的可能很小	1.5 1.0 0.7
F_4（预防难易程度）	不能预防 能够预防 预防很容易	1.3 1.0 0.7
F_5（是否需要更改设计）	需要大修改 可用类似的设计 用同一类的设计	1.2 1.0 0.8

故障频率分类及发生概率列于表5.5-14中。

表5.5-14　故障频率分类及发生概率

编码	故障分类	发生概率
5	常常发生的故障模式	75%以上
4	极普通的故障模式	40%～75%
3	普通的故障模式	15%～40%
2	少见的故障模式	5%～15%
1	极少发生的故障模式	5%以下

当按式（5.5-40）算得的致命度$C_B>1$，说明所分析的零件失效时对系统造成的危险程度较大，因此需要改变原有的设计，并提出预防失效的措施。如果$C_B \approx 0.5 \sim 1$，说明有一定程度的危险，这时需要修改原有设计，并采取其他措施，防止发生危险。如果$C_B<0.5$，说明没有危险，因此不需要修改原设计。

6）作FMEA或FMECA表格。

4.2.3　概率设计方法

从可靠性的角度考虑，影响机械产品故障的各种因素可分为“应力”与“强度”两类，“应力”大于“强度”时故障发生。“应力”不仅仅指外力在微元面积上产生的内力与微元面积比值的极限，而且包括各种环境因素，如温度、湿度的变化，腐蚀作用，粒子辐射等。“强度”是力学结构承受应力的能力，因此，凡是能阻止结构或零件故障的因素，统称为强度，如材料力学性能、加工精度、表面粗糙度等。

“应力”是一个受多种因素影响的随机变量，具有一定的分布规律。同样，“强度”也是一个有一定分布规律的随机变量。在这种情况下，研究产品的可靠性问题就是概率设计方法所做的工作。

概率设计方法首先要确定应力分布。

1）确定所有重要的失效模式，并确定适当的失效判据。常见的基本失效模式已在本篇第2章中做了详细介绍。

2）如果零件的主要失效模式是由于复合应力引起的，则应当作应力单元体分析。一般是根据零件的断面尺寸和变化、受力情况，选择几个应力比较高的点取应力单元体。如果不能判别危险的应力点，则应计算二、三个最有可能的点，计算其应力，以确定最有可能导致零件失效的点的位置。

3）计算应力单元体上的6个名义应力分量（3个正应力分量和3个切应力分量）。

4）使用适当的应力修正系数，确定每一应力分量的最大值。这些应力修正系数包括：应力集中系数、载荷系数、温度应力系数、成形或制造应力系数、热处理应力系数、表面处理应力系数、装配应力系数、腐蚀应力系数、直接表面环境应力系数、切痕敏感系数等。但是并非在任何情况下都要考虑这么多修正系数。有些应力修正系数可以看作是强度修正系数，而放在确定强度分布时考虑。

5）计算主应力。根据选用的失效判据，决定是否计算主应力。当以最大主应力或最大切应力作为失效判据时，需要计算主应力。当以变形能作为失效判据时，不需计算主应力。

6）根据失效判据，将上述的应力分量综合为复合应力，复合应力值最大的部位，就是零件的失效概率最大的部位。

7）确定每个名义应力、应力修正系数和设计参数的分布。为了确定名义应力分布，需要知道载荷的分布、尺寸的分布等。载荷分布最好通过试验研究来确定。

8）把上述的那些分布综合成为应力分布。一旦应力公式中的名义应力分布、应力修正系数和参数等的分布得到确定，则应力分布也可以确定。综合这些分布的方法可以采用代数法、矩法和蒙特卡洛模拟法等。

9）对于每一重要的失效模式重复上述过程。为了能够确定零件的可靠度，所有可能出现的重要失效模式都应进行

考虑。因此，需要重复上述过程来确定每一失效模式的应力分布。

确定应力分布后需要确定强度分布。

1）确定强度判据。强度判据应与确定应力分布时所用的失效判据相一致，如最大正应力判据、最大切应力判据、最大变形能判据等。

2）确定名义强度分布。失效模式不同，选用的强度值也不同，可以分别选择屈服强度、抗拉强度、疲劳强度、疲劳极限等值。

3）用适当的强度系数去修正名义强度。试样的强度与零件的强度是有差别的。在确定了试样的名义强度后，需要用适当的强度系数去修正名义强度，把在理想的、标准的试验条件下确定的试样名义强度转换成在实际使用环境下、具有实际的几何尺寸的零件的强度。由于实际使用条件、工作环境和实际几何尺寸反映出的任何偏差，都要求对强度及其分布进行修正。因此，要考虑尺寸系数、表面质量系数和应力集中系数。

4）确定强度公式中名义强度、每一强度修正系数和参数的分布。

5）把上述的这些分布综合成为强度分布。实际零件的强度分布最好通过可靠性试验来获得，如果不通过试验来获得，可以采用与综合应力分布相同的方法来综合强度分布。

在确定应力分布和强度分布以后，便可对这两个分布加以综合（图5.5-24），从而计算零件的可靠度。

从应力分布和强度分布的干涉理论出发，可靠度是"强度大于应力的整个概率"，表示为

$$R(t)=P(r>s)=P(r-s>0)=P\left(\frac{r}{s}>1\right) \tag{5.5-41}$$

当满足上式时，零件不会失效，否则会出现失效。

当 $t=0$ 时，两个分布之间有一定的安全裕度，不会产生失效，但随着时间的推移，由于材料和环境等因素变化，导致在时间 t 时应力分布与强度分布发生干涉，这时将会产生失效。

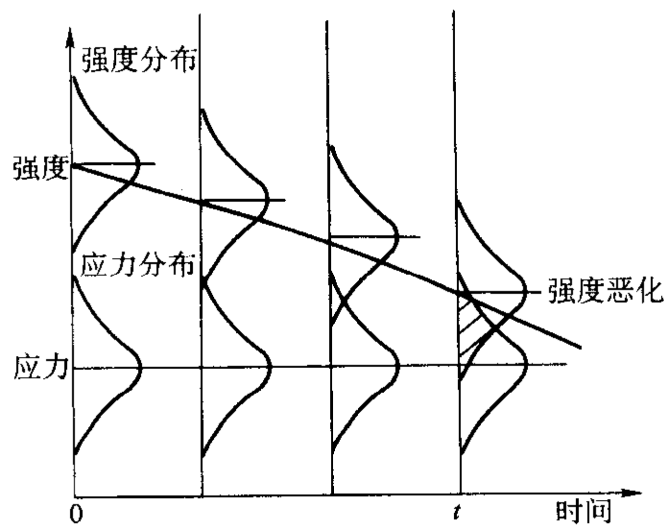


图5.5-24 应力-强度分布与时间的关系

需要研究的是两个分布发生干涉的部分，即尾部发生重叠的部分，阴影部分表示零件的失效概率。

$$R(t)=1-P_f=1-F(t) \tag{5.5-42}$$

式中，P_f 为失效概率；$F(t)$ 为不可靠度。

4.3 可靠性设计的步骤

可靠性设计步骤包括15步。

1）确定设计的问题及任务的轮廓 如果要设计的是一个常用零部件，例如轴，其设计问题简单而明确，目的是为了确定轴的危险断面的直径。但是对于复杂的系统，例如航天飞行器，则必须确定完整的任务轮廓以及相应的环境轮廓，并在此基础上，确定各子系统的目标可靠度。

2）确定有关的设计变量和参数 这些设计变量和参数应当是：①对设计任务是有意义的；②惟一的和不重复的；③在试验前后和进行期间都能度量的。凡设计任务涉及的所有基本议程都能用这些设计变量和参数来进行计算。

例如在设计同时受弯矩和扭矩作用的轴时，需要确定的设计变量和参数有：轴的直径、弯应力、切应力、材料的疲劳极限以及应力集中系数、尺寸系数和表面质量系数等。

3）进行失效模式、影响及致命度分析（FMECA）。

4）确定零件的失效模式是独立的还是相关的 如果所有的失效模式是互相独立的，也就是说一种失效模式不影响任何其他失效模式的性质，则在进行一种失效模式下的应力及强度计算时，不需考虑所有其他失效模式的影响。如果一种失效模式的性质受到另一个可能同时发生的失效模式的影响，则受到影响的应力和强度应当加以修正。这样计算出的每种失效模式的可靠度，将是互相独立的。

5）确定涉及到的每种失效模式的判据 比较常用的判据有：①最大正应力；②最大切应力；③最大变形能；④最大应变能；⑤最大应变；⑥最大变形；⑦疲劳情况下的变形能；⑧疲劳情况下的最大总应变；⑨最大许用腐蚀量；⑩最大许用磨损量；⑪最大许用振幅；⑫最大许用蠕变变形等。

6）确定应力公式 对于每一种失效模式，必须确定载荷、尺寸、温度、时间、物理性质、使用和工作环境等设计变量和参数之间的函数关系，从而得出应力计算公式。

7）确定每个失效模式下的应力分布 根据失效模式、影响及致命度分析（FMECA）鉴别的所有重要的和致命的失效模式，都应当予以考虑，并且按照上述过程确定其应力分布。

8）确定强度公式 对于确定的失效模式而言，强度应当是失效时的应力。因此，一旦强度被工作应力超过，将会导致一定的失效模式，而这种失效模式发生的概率，就是不可靠度。

通常得到的是材料的强度数据，为了得到零件的强度数据，需要用几个系数加以修正。

9）确定每个失效模式下的强度分布 将试样的强度与每个强度修正系数综合起来，便可得到零件的强度分布。

最好用实验方法直接得出零件的强度分布数据。

10）对于每一种致命的失效模式，确定其与应力分布和强度分布相关的可靠度 对于零件的每一种致命失效模式，都应计算其可靠度。这些失效模式可以是：变形，包括弹性变形、塑性变形等；断裂，包括韧性断裂、脆性断裂、疲劳等；腐蚀和磨损等。

有的零件可能只有一种致命的失效模式，这时只需要根据这一种失效模式的判据来计算可靠度。有的零件可能会出现几种致命的失效模式，这时需要计算所有失效模式下相应的可靠度。

11）确定同时考虑到所有致命失效模式的零件的整个可靠度 确定了归因于每一种失效模式的可靠度之后，下一步是确定归因于所有失效模式的零件的整个可靠度。

一种方法是，假设所有的失效模式都是可能的，每个失效模式可能是独立地或相关地产生或扩展，而与每个失效模式有关的应力分布和强度分布，在确定时已经考虑了其他失效模式的影响，所以，当至少一种失效模式出现时，零件即失效。这时零件的整个可靠度为

$$R_0=R_1R_2R_2\cdots R_n=\prod_{i=1}^{n}R_i \tag{5.5-43}$$

式中，R_i 为归因于失效模式 i 的可靠度；n 为可能的失效模式数，$n\geqslant1$。

另一种方法是，假定零件将在最可能发生的一种失效模

式下失效，也就是说，这种失效模式有最高的发生概率或最低的可靠度，表示为 R_{i-min}，于是，零件的整个可靠度为

$$R_0 = R_{i-min} \tag{5.5-44}$$

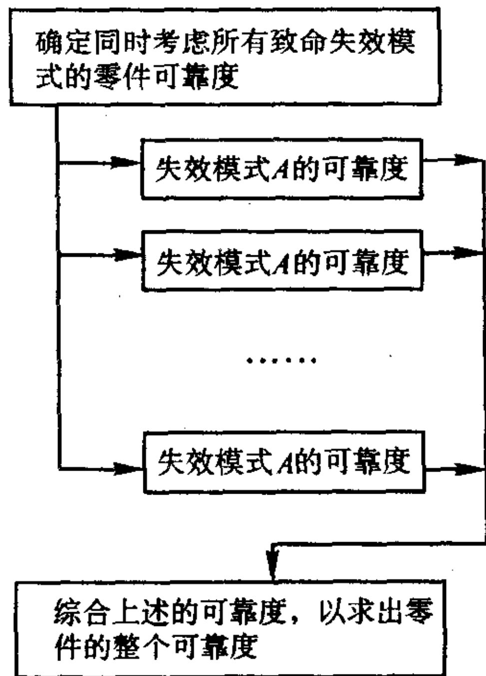


图 5.5-25　确定零件可靠度的流程图

比较一下这两种方法，不难看出，式（5.5-43）给出了最小可靠度，而式（5.5-44）给出了最大可靠度。零件的实际可靠度是在上两式所给出的值之间。如果的确是由于单一的失效模式引起的零件失效，则实际的可靠度将接近于或等于 R_{i-min}。如果是由于多种原因引起零件失效，则零件的实际可靠度将接近于或等于 $\prod_{i=1}^{n} R_i$。

实际上，零件的真实的可靠度只有在进行集中的可靠性试验或获得并分析实际使用数据之后，才能计算出经过验证了的可靠度。确定零件可靠度的流程见图 5.5-25。

12）确定零件可靠度的置信度　在求出零件的可靠度之后，需要计算置信度，以确定预测的可靠度的可信程度。

13）对于系统中所有的关键零件重复上述步骤，求出各自的可靠度。

14）在已知每个零件的可靠度的基础上，计算子系统的以及整个系统的可靠度。然后对设计进行修正，直到系统的可靠度等于或大于事先规定的可靠度目标值为止。

15）如果必要，对整个设计的内容进行优化，包括：①性能；②可靠性；③维修性；④安全；⑤费用；⑥质量；⑦体积；⑧操作性；⑨交货日程表等等。

编写：张峥（北京航空航天大学）

参考文献

1 张栋等编著．机械失效的实用分析．北京：国防工业出版社，1997

2 涂铭旌等编著．机械零件的失效分析与预防．北京：高等教育出版社，1993

3 钟栋梁，姬永兴，陈列等编著．机电气事故（故障）的检查与分析方法．北京：蓝天出版社，1993

4 K.h.G.Schnit，Thomas.Strategy of Failure Analysis and Failure Prevention in Power Generating Equipment，Proc.of Int.Conf.On Failure Analysis and Prevention（ICFAP'95），June，1995，23～26，Beijing，China

5 W.Wallace，J－P.Immarigeor and A.K.Koul.The Function and Failure of Components in Modern Gas Turbine Engines，Proc.of Int.Conf.On Failure Analysis and Prevention（ICFAP'95），June 1995，23～26，Beijing，China

6 刘英杰，成志强编著．磨损失效分析．北京：机械工业出版社，1991

7 肖纪美．韧化原理和工艺——提高材料断裂韧性的途径．国外金属材料，1997，No.2

8 刘正义，吴连生等编著．机械装备失效分析图谱．广州：广东科技出版社，1990

9 航空航天工业部航空装备失效分析中心编著．金属材料断口分析及图谱．北京：科学出版社，1991

10 宋子濂，史建平主编．车轴断裂失效分析图谱．北京：中国铁道出版社，1995

11 洛阳矿山机器厂等编著．重载齿轮损伤失效图谱．北京：中国科学技术出版社，1990

12 宋兆泓主编．航空发动机典型故障分析．北京：北京航空航天大学出版社，1993

13 William R.，Warke.Seeking the Basic Cause of Gas Turbine Hot Section Failure，Proc.of Int.Conf.On Failure Analysis and Prevention（ICFAP'95），June，1995，23～26，Beijing，China

14 Toshiro kobayashi.Evaluation of Dynamic Fracture Toughness by Instrumented Impact Test，Proc.Of Int.Conf.On Failure Analysis and Prevention（ICFAP'95），June 23～26，1995，Beijing,China

15 B.Bems，C.Broeckman，D.Weichet.Fracture Mechanisms in Particle Reinforced Metal Matrix Composites，Proc.Of Int.Conf.On Failure Analysis and Prevention（ICFAP'95），June，1995，23～26，Beijing，China

16 Zhong Qunpeng，Zhang Zheng，et al. The theoretical Derivation and Experimental Demonstration of Ductile—Brittle Transition Model and Mechanism，Proc.of Int.Conf.On Failure Analysis and Prevention（ICFAP'95），June，1995，23～26，Bei-jing，China

17 ［美］S.Suresh著．材料的疲劳．王光中等译．北京：国防工业出版社，1993

18 郑修麟著．金属疲劳的定量理论．西安：西北工业大学出版社，1994

19 H.S.Wu，Q.P.Zhong & Y.P.Ying.Probabilistic Approach to Applications of The R6 Structural Integrity Assessment Method，Int.J.press.Ves & Piping，65，1996

20 H.S.Wu，Q.P.Zhong.Structure Integrity Assessment for a Flawed Ti—tanninum Pipe，Int.J Pres.Ves & Piping，1998，75：151～154

21 Jin Xing，Q.P.Zhong & Y.J.Hong.A Probabilistic Fracture Assessment Method Based on the R6 Procedure，Int.J.Pres.Ves & Piping，1997，73：161～163

22 Jin Xing，Q.P.Zhong & Y.J.Hong.A Simple Log Normal Random Process Approach of the Fatigue Crack Growth Considering，the Distribution of Initial Crack Size and Loading Condition Int.J.Pres.Ves & Piping，74，1997，7～12

23 徐滨士主编．表面工程与维修．北京：机械工业出版社，1996

24 钟群鹏，张峥，武淮生，左尚志．金属疲劳扩展区和瞬断区的物理数学模型．航空学报，2000，Vol21（增刊）：s11～14

25 陈南平，顾守仁，沈万慈．脆断失效分析．北京：机械工业出版社，1993

26 张栋．机械失效的痕迹分析．北京：国防工业出版社，1997

27 胡世炎．机械失效分析．手册．成都：四川科学技术出版社，1999

28 钟群鹏．金属断口宏观特征形态与力学参量之间的定量关系．兵器材料与力学，1984（1）：2

29 钟群鹏，傅国如，张峥，骆红云．机电装备失效研究的内涵及其科学问题探讨．机械工程学报，2003，Vol（10）：23～31

30 吴学仁，陶春虎，张栋，钟培道．失效分析的历史发展与航空装备失效分析的展望．全国第二届航空装备失效分析会议论文集，1997

31 钟群鹏，张峥，武淮生，等．失败（失效）是成功（安全）之母—试论机电装备失效分析预测预防．中国机械工程，2000，11（1）：70～72

32 钟群鹏，王仁智，陈玉民等．机电装备失效分析预测预防进展—失效学体系的形成和发展（一）．理化检验—物理分册，1999，35（2）：76～79

33 钟群鹏，王仁智，陈玉民等．机电装备失效分析预测预防进展—失效学体系的形成和发展（二）．理化检验—物理分册，1999，35（3）：114～118

34 钟群鹏，张峥，田永江．机械装备失效分析诊断技术．北京航空航天大学学报，2002，28（5）：497～502

35 Charlie R.Brooks，Ashok Choudhury.Metallurgical Failure Analysis.Mcgraw－Hill，1993

36 ［法国］G.亨利，［联邦德国］D.豪斯特曼．曾祥华译．宏观断口学及显微断口学．北京：机械工业出版社，1990

37 陶春虎，钟培道，王仁智等．航空发动机转动部件的失效与预防．北京：国防工业出版社，2000

38 李想，张栋，傅国如．疲劳失效件断口定量初探．全国第二届航空装备失效分析会议论文集，1997

39 航空航天材料咨询研究组．航空航天材料咨询报告．国防工业出版社，北京：1999

40 钟秉林，黄仁．机械故障诊断学．北京：机械工业出版社，1997

41 王江萍．机械设备故障诊断技术及应用．陕西：西北工业大学出版社，2001

42 盛兆顺，尹琦岭．设备状态监测与故障诊断技术及应用．北京：化学工业出版社，2003

43 王仁智等编著．疲劳失效分析．北京：机械工业出版

社，1990
44 程育仁等编著．疲劳强度．北京：中国铁道出版社，1990
45 李志安．压力容器断裂理论与缺陷评定．大连：大连理工大学出版社，1994
46 荆树峰，曾广欣，高世国译．国外压力容器缺陷评定标准．北京：劳动人事出版社，1982
47 王军．损伤力学的理论与应用．北京：科学出版社，1997
48 杨光松．损伤力学与复合材料损伤．北京：国防工业出版社，1995
49 中华人民共和国国家标准．在用含缺陷压力容器安全评定（报批稿）
50 丁遂栋．断裂力学．北京：机械工业出版社，1997
51 楼志文．损伤力学基础．西安：西安交通大学出版社，1991
52 [法] J. 靳迈特．损伤力学教程．倪金刚译．北京：科学出版社，1996
53 李灏．损伤力学基础．济南：山东科学技术出版社，1992
54 Rahman, S., Brust, F., Ghadiali, N., Choi, Y.H.Refinement And Evaluation of Crack Opening Area Analyses For Circumferential Through – Wall Cracks In Pipes, NUREG/CR – 6300, 1995
55 Paul, D.D., Ahmad, J., Scott, P.M., Flanigan, L.A.And Wilkowski, G.M., Evaluation And Refinement of Leak – Before – Break Models, NUREG.CR – 5128, Rev.1, June 1995
56 沈士明，周剑秋，朱利洪．压力管道缺陷评定方法的现状与发展．南京航空航天大学学报，1999，1：74～78
57 钟群鹏，武准生，田永江等．我国压力容器安全评定技术的形式和发展．中国机械工程，1997，Vol.8（5）：95～100
58 俞树荣，王志文．压力容器评定的现代方法．化工机械，1995，Vol.22（5）：52～58
59 朱立洪．在役含缺陷压力管道可靠性模糊评定方法研究．南京化工大学博士论文．南京化工大学，2000
60 赵振宇，徐用懋．模糊理论和神经网络的基础与应用．北京：清华大学出版社，南宁：广西科学技术出版社，1995
61 Jovanovic A., Gel S.Overview of some relevant applications in the field of power and structural engineering in Europe an USA.17 MPA – Seminar, 2/47, Sruttart, Germany, 1991.
62 肖纪美．应力作用下的金属腐蚀．北京：化学工业出版社，1990
63 Muhlbauer W.K.Pipeline Risk Management Manual.2nd ed.Gulf Publishing Co.,Houston, Texas, 1996
64 ASME B31.8 S 2001.Supplement to B31.8 on Management System of Gas Pipeline.ASME B31.8 – 2001, New York, 2001
65 Linbens, D., Shetty, N.K.et.al; A Probabilitic Approach to Fracture Assessment of Onshore Gas Transmission Pipelines, Pipes & Pipelines International 1998, 7～8
66 Qualitative Methods for Assessing Risk.Mahn; J.A.ASME PVP, 1995, Vol (296): 53～58
67 ASME Ⅺ I WB – 3640 and Appendix C, Flaw Evaluation Procedures and Acceptance Criteria for Austenic Piping, The ASME, New York, 1995
68 American Petroleum Institute.Risk based Resource Document. APIPR581.2000
69 Michael.J.Zelensky, et al.Public Safety Risk Assessment of Natural Gas Liquid Pipelines, International Pipeline Conference, ASME, 1996 (1)
70 BS7910: 1999 (Incorporating Amendment No.1) Guide on Methods for Assessing the Acceptability of Flaws in Metallic Structures, British Standards Institution, London, 2000
71 API.579, Recommended Practice for Fitness – for – Service Washington, D.C.American Petroleum Institute, 2000
72 BSI PD6493．焊接结构缺陷评定方法．1991．李培宁译．华东理工大学化机所
73 黄祥瑞．可靠性工程．北京：清华大学出版社，1990
74 J.K.Vrijling & P.H.A.J.M.van Gelder.An Analysis of the Valuation of a Human Life.Foresight and Precaution, Cottam, Harvay.Pape&Tait (eds). Balkema Rotterdam.2000: 197～201
75 何庆复，杨萍，王德志．机械工程材料及选用．北京：中国铁道出版社，2001
76 姜伟之，赵时熙，王春生，张峥．工程材料力学性能．北京：北京航空航天大学出版社，2000
77 吴清可．防断裂设计．北京：机械工业出版社，1991
78 马国柱，类柄林．现代质量管理理论与实务．北京：机械工业出版社，1999
79 李新和．机械设备维修工程学．北京：机械工业出版社，1999
80 宋余九．金属材料的设计、选用、预测．北京：机械工业出版社，1998
81 郑修麟．材料的力学性能．陕西：西北工业大学出版社，1991
82 张树松，仝爱莲．钢的强韧化机理与技术途径．北京：兵器工业出版社，1995
83 冯端，师昌绪，刘治国．材料科学导论．北京：化学工业出版社，2002
84 赵少汴．抗疲劳设计．北京：机械工业出版社，1994
85 葛中民．耐磨损设计．北京：机械工业出版社，1991
86 黄祥瑞．可靠性工程．北京：清华大学出版社，1990
87 吴继勋．金属防腐蚀技术．北京：冶金工业出版社，1998
88 孙秋霞．材料腐蚀与防护．北京：冶金工业出版社，2001
89 章葆澄，朱立群，周雅．防腐蚀设计与工程．北京：北京航空航天大学出版社，1998
90 王英杰．金属工艺学．北京：高等教育出版社，2001
91 刘秀晨，安成强．金属腐蚀学．北京：国防工业出版社，2002
92 钱苗根，姚寿山，张少宗．现代表面技术．北京：机械工业出版社，2001
93 陆迁孝，郑鹏洲．可靠性设计与分析．北京：国防工业出版社，1995
94 朱文予．机械可靠性设计．上海：上海交通大学出版社，1992
95 牟致忠，朱文予．机械可靠性设计．北京：机械工业出版社，1993
96 牟致忠．可靠性设计．北京：机械工业出版社，1993

中 国 材 料 工 程 大 典

CHINA MATERIALS ENGINEERING CANON

第 1 卷 材料工程基础

第 *6* 篇

材料强度设计

主 编 谢里阳 王永岩

编 写 谢里阳 王永岩 何雪浤

王 雷 李 佳

审 稿 中国材料工程大典编委会

中国机械工程学会

中国材料研究学会

中国材料工程大典编委会

第1章 概 论

1 强度设计的目的与作用

工程中许多材料的结构破坏或失效都是由于强度不足引起的。随着对设备或结构性能要求的提高，其工况、载荷及环境条件越来越苛刻，涉及到的强度问题也越来越复杂。材料强度设计的目的就是保证设计对象的强度满足要求，或者说保证设计对象在规定的使用条件下不发生强度失效。

材料强度是表征材料承载能力的力学性能指标。影响材料强度的因素除材料自身的属性如化学成分、微观组织结构（在很大程度上由热处理工艺决定）、缺陷尺寸与分布外，还有环境因素如应力状态、载荷性质、加载速率、温度和介质等。

在研究材料强度问题时，根据材料性质、载荷性质和环境条件等的不同，可以做不同的分类。

按材料性质的不同，可分为脆性材料强度、塑性材料强度和带裂纹材料的强度。脆性材料强度研究的是脆性材料的强度问题。如铸铁等脆性材料，受载后几乎没有塑性变形就突然断裂。这种脆性材料的强度计算以抗拉强度 σ_b 为标准。塑性材料强度研究塑性材料的强度问题。如软钢等塑性材料，在断裂前有较大的塑性变形，卸载后不消失，又称残余变形。塑性材料的强度计算以屈服点 σ_s 为标准。对于没有屈服现象的塑性材料，取与0.2%的塑性变形相对应的应力作为名义屈服极限，用 $\sigma_{0.2}$ 表示，以此为强度计算的标准。带裂纹材料的强度是研究含裂纹体材料的强度问题，由断裂力学中的 K_C、K_{IC}、δ_C 或 J_C 作为强度计算的标准。

按载荷性质的不同，材料强度又分为静强度、冲击强度和疲劳强度。静强度指材料在静载荷下的强度，根据材料性质的不同，分别以屈服点 σ_s 或抗拉强度 σ_b 作为强度计算的标准。冲击强度指材料在冲击载荷下的强度，是金属材料抵抗冲击破坏的能力。冲击载荷在零件中产生的冲击应力除与零件的形状、体积和局部弹塑性变形等有关外，还同与其联接的物体有关。因此，冲击载荷作用下的强度计算比静载荷下的强度计算要复杂的多。一般情况下，在引入动载系数后，按静强度情况进行计算。结构钢的强度极限和屈服极限随冲击速度的增大而提高。疲劳强度指材料在循环载荷作用下的强度。在循环载荷作用下，材料产生疲劳失效，通常以材料的疲劳极限 σ_{-1} 作为强度计算的标准。

按环境条件的不同，材料强度又分为高温强度、低温强度、腐蚀强度等等。

材料强度设计需要精确地确定材料的性能和载荷，要用到数学、力学、概率统计及材料力学行为等知识，还要用到应力分析方法、材料破坏理论，等等。强度设计有根据确定性假设进行设计和根据概率论进行设计两种设计思想。传统的设计是把所有设计变量都作为确定性的量对待，设计准则是满足“设计应力小于材料的许用应力”。而概率设计则是将有关设计变量作为随机变量，用统计方法描述材料的强度指标。概率强度设计准则是使“强度大于载荷的概率”满足设计指标。这个貌似简单的设计准则的改变，对设计理念有深远的影响。

概括地讲，强度设计涉及到以下三个方面：

1）结构应力分析；

2）材料力学特性；

3）强度理论。

一般的固体力学计算都着重于解决第一个问题；第二个问题一般根据材料试验和结构条件等确定。联系这两方面的结果进行强度分析时，则需要研究和建立强度理论，即研究材料在复杂应力作用下产生破坏的规律，从而建立相应的理论和强度计算法则。

在固体力学中所研究的力学量，如应力、应变、应变速度、时间等，都是宏观度量，所采用的方法是宏观的连续介质力学方法。因此，固体力学所研究的材料破坏规律性的理论，一般称为宏观强度理论或工程强度理论，它是固体力学的一个分支，也可以说是力学与材料科学相结合的一门边缘科学。

材料强度学的重要任务之一是研究材料（零件）在受载过程中变形和断裂的规律。这方面的重要命题很多，例如：零件受载过程中，应力与应变的分布；各种内应力的形成和变化；位错和各种缺陷以及裂纹附近应力场的分析；位错开动和裂纹扩展推动力的计算；变形和断裂过程中物理模型的建立及数学公式的推导等等。研究这些问题，除了要对其物理现象进行分析外，还需要对其力学过程进行讨论。另外，解决生产实际问题时，还要求对机件的服役条件和失效状况从力学角度进行详细可靠的分析。

研究材料强度问题所需要的力学基础知识，相当一部分是属于材料力学、弹性力学和塑性力学的范畴，另一部分是属于材料的变形和断裂力学范畴。

为了分析和处理问题方便，力学中对所研究的对象通常作如下假定。

1）连续。认为物体是由一种连续性的介质所构成，没有一点空隙。这样就可把描述物体变形过程的参量（如应力、应变等）用物体所占空间的连续函数来表达。

2）均匀。认为所研究物体的各个部分都有相同的成分、组织和与之相应的机械性质，则物体的强度、塑性、弹性模量、泊松比等都是常数，而不是坐标位置的函数。

3）各向同性。假定所讨论物体的各种性质在物体的各个方向上都是相同的，不依方向而改变。

4）如果所讨论的物体处于弹性变形阶段，则假设物体是完全弹性的，即加载卸载过程中，物体的应力应变关系是可逆的，其应力应变互为单值函数。当载荷卸到零时，变形也恢复到零，无任何残余变形。并且，一般情况下，应力应变符合正比直线关系，即胡克定律。

当然，工程用的实际金属材料，并不严格符合上述这些假设。实际金属材料中不可避免地含有各种宏观尺度或微观尺度的孔洞和裂纹，例如大大小小的夹杂和其他缺陷，因而事实上是不连续、不均匀的。材料在冶炼、浇铸、压力加工、焊接和热处理等过程中，由于温度不均匀、受载不均匀、冷却速度不均匀，在机件内部，宏观各部位之间或微观区域之间，会形成成分、组织和内应力以及相应的力学性能的不均匀。另外，在凝固过程中，有可能形成结晶构造的方向性、成分偏析；在锻压过程中形成金属流动的方向性；相组织初缺陷分布的方向性、内应力分布的方向性等，从而产生材料宏观上的各向异性。特别是新兴的复合材料，各向异性更是其主要特点。因此实际材料总是不同程度地具有不连续性、不均匀性和各向异性。所谓的“完全弹性”也是有条件的而不是绝对的。

5）初应力为零。

2　材料强度设计方法

2.1　常规强度理论

随着科学和技术的不断发展和进步，人们对机械强度的认识也在不断深入。人们最早认为影响强度的外在因素就是外载荷，后来又提出了应力的概念。早期对强度的认识是材料抵抗破坏的能力仅取决于材料本身的力学性能，并且只限于静强度破坏这一现象，相应地发展了静载荷条件下的材料强度理论，形成了传统的常规的强度理论体系。

与现代强度理论比较而言，传统机械强度理论具有两个明显的特点：一是假设材料性能是均匀的、各向同性的、连续的；二是只考虑静载荷作用。

常规强度设计的计算步骤是，根据设计要求，由理论力学方法确定零构件所受外力，用材料力学、弹性力学或塑性力学计算其内力，再根据机械设计的知识确定其结构尺寸和形状，最后计算设计对象的工作应力或安全系数。强度设计准则可用公式表示为

许用应力准则：　$\sigma \leqslant [\sigma]$　(6.1-1)

安全系数准则：　$n \geqslant [n]$　(6.1-2)

式中，$[\sigma]$、$[n]$ 分别称为许用应力、许用安全系数，安全系数定义为强度与应力之比。

满足式 (6.1-1) 和式 (6.1-2) 的设计是安全的；否则是不安全的。

对于塑性材料，一般要用屈服强度这个指标：

$$[\sigma] = \sigma_s / [n]_s \qquad (6.1\text{-}3)$$

式中，$[n]_s$ 为以屈服极限为基准的许用安全系数。

对于脆性材料　$[\sigma] = \sigma_b / [n]_b$　(6.1-4)

式中，$[n]_b$ 为以强度极限为基准的许用安全系数。

安全系数 n 是考虑到实际结构中可能有的缺陷和其他意想不到的或难以控制的因素（如计算方法的不准确性、载荷估计的不准确性等），用来保证所设计的机械零构件有足够的强度安全储备量，保证在最大工作载荷下，其工作应力不超过制造零构件材料的极限应力。

确定安全系数时，一般遵循如下原则：对重要零构件，$[n]$ 取大值；对非重要零构件，$[n]$ 取小值，但必须满足 $n>1.0$ 要求。

这种传统的常规强度设计方法，虽然不适用于含裂纹和缺陷材料及复合材料等制造的零构件，以及循环随机载荷作用下的零构件，但由于其经过长期的发展已形成一套较为完整的体系，比较实用、简便，因而长期以来，在工程中得到了广泛的应用，到现在仍然是一种应用广泛的工程计算方法，也是现代机械强度设计计算的基础。

2.2　现代机械强度理论

工程中绝大多数设备或结构是在动载荷作用下工作的，因而疲劳断裂是工程中常见的失效形式。19世纪40年代，人们从火车轴大量断裂的事故中，认识到了在交变应力作用下的疲劳破坏现象。德国工程师 Wöhler 通过大量试验研究，奠定了疲劳研究的基础。疲劳强度理论已经成为现代机械强度理论的主要内容，成为每个机械设计人员必须具备的基础知识。

20世纪20年代，动力机械开始应用于高压、高温蒸汽等恶劣环境中，材料蠕变成为这些机械零构件的主要破坏形式。从此，蠕变以及蠕变与疲劳的交互作用成了强度问题中的一个重要研究领域。

随着对疲劳和蠕变研究的深入进展，人们又发现零件抵抗破坏的能力和时间有密切关系。因此，强度问题又直接地和寿命的概念连接在一起。故一般情况下的强度计算同时也包括了寿命计算。

在工程实践中，寿命计算结果分散性很大。这是因为，表征材料强度的参数都是由试验确定的，如强度极限、疲劳极限、表面状况、尺寸大小等都具有离散性，数值是一个范围，故计算结果误差必然很大。为了在强度和寿命计算中反映出这一特性，人们又引入了疲劳强度和寿命的可靠性概念，即出现了疲劳强度可靠性和疲劳寿命可靠性的学科分支。

随着疲劳强度研究的深入，人们发现零件应力分布不均匀对疲劳强度影响很大，因此在应力应变分析领域内发展起来局部应力应变的研究分支。这方面的研究在接触应力和零件几何形状不连续处的应力应变集中两个方向上深入和发展。

20世纪40年代，尤其是二战中，飞机零构件的脆性断裂事故不断发生，导致了断裂力学这门新学科的建立。目前，断裂力学在零构件的脆性断裂和疲劳裂纹形成与扩展寿命方面有着广泛的应用。

由于现代的机械零构件工作环境越来越恶劣，如高温、高压、腐蚀环境，工作载荷大，变化频繁，且多数是随机载荷，制造零构件的材料也由过去主要是钢铁，发展到用高强度钢、超高强度钢、复合材料、陶瓷材料及非金属聚合物等，因此常规强度设计的理论和方法已远远不能满足现代机械使用的材料、工作条件及环境的要求，必须加以改进、发展和完善，故而形成了现代机械强度的设计理论和方法。

现代机械强度理论除了仍然要用到弹塑性理论之外，还需要应用疲劳和断裂理论，同时需要利用现代测试技术手段及计算机技术对机械结构进行综合分析与计算，最终给出科学的强度设计计算指标，以满足工程的要求。

与传统的材料强度设计相比，现代设计有如下一些具体的特点。

1) 传统的设计中主观和经验的成分占有很大的比重，过去的产品一般都是在实践中通过经验积累不断改进、逐步完善的。当代的设计过程已从基于经验转变为基于设计科学，成为人们主动的、按思维规律有意识地向目标接近的创造过程。

2) 传统的强度设计主要依据材料力学或测试手段，偏重于静强度；而现代的有限单元方法、断裂力学等领域的研究成果，能够处理动强度与裂纹扩展等问题，进一步强化了强度设计的能力。

3) 传统的机械设计过程往往是根据任务和目标，先作出第一方案，甚至造出样机，然后通过评定与考核，进行修改，形成第二轮方案，如此反复，直到满意为止。这一设计过程历时长，耗费大。现代机械设计则可根据各种给定的条件，运用优化设计理论和方法，借助电子计算机求得最佳设计参数和方案，因此，设计的耗费低、速度快，而且科学地反映设计的最优状态。

4) 传统的设计建立在手工操作的基础上，机器的动态效果只能通过简化、抽象的运动学、动力学模型来反映，而今天的计算机仿真技术能对未来机器的运转状态清晰地加以描述。

编写：谢里阳（东 北 大 学）
王永岩（青岛科技大学）

第 2 章　材料静强度设计

1　材料的静载拉伸特性

1.1　材料的力学特性和软性系数

在温度、环境介质和加载速度确定的条件下，材料的力学性能受加载方式（即应力状态）的影响。切应力主要引起塑性变形和导致韧性断裂，正应力容易导致脆性断裂。因此，研究金属的变形与断裂特征，需要研究在不同加载形式下，切应力和正应力的相对大小。在一般复杂应力状态下，最大切应力 $\tau_{max}=\frac{1}{2}(\sigma_1-\sigma_3)$，而按最大正应变条件计算得到的等效最大应力为 $\sigma_{max}=\sigma_1-\nu(\sigma_2+\sigma_3)$，定义二者的比值 α 为软性系数。

$$\alpha=\frac{\tau_{max}}{\sigma_{max}}=\frac{\sigma_1-\sigma_3}{2[\sigma_1-\nu(\sigma_2+\sigma_3)]} \tag{6.2-1}$$

α 越大，最大切应力分量越大，表示应力状态越软，金属易于先产生塑性变形。反之，α 值越小，表示应力状态越硬，则金属易于产生脆性断裂。几种典型加载方式的应力状态软性系数 α 值如表 6.2-1 所示。

表 6.2-1　不同加载方式下的软性系数 α 值（$\mu=0.25$）

加载方式	主应力			软性系数 α
	σ_1	σ_2	σ_3	
三向等拉伸	σ	σ	σ	0
单向拉伸	σ	0	0	0.5
扭　转	σ	0	$-\sigma$	0.8
二向压缩	0	$-\sigma$	$-\sigma$	1
单向压缩	0	0	$-\sigma$	2
三向压缩	$-\sigma$	-2σ	-2σ	∞

由表 6.2-1 可知，三向等拉伸时应力状态最硬，因为其切应力分量为零。扭转时 α 接近 1，属于较软的应力状态。单向压缩或多向压缩时，α 都大于 1，因此属于软性应力状态。当用某一特定的硬度头在工件表面施压测定硬度时，其应力状态相当于三向不等压缩。因此，硬度试验时的加载方式属于很软的应力状态。

单轴拉伸应力状态的软性系数（$\alpha=0.5$）适中，所以只要材料固有的塑性较好，强度不是太高，利用这种试验方法就能够较全面地显示材料的力学行为，即既能显示弹性变形，又能充分显示塑性和最后断裂，因而相应可标定出一系列对应的基本力学性能（强度和塑性）指标，为结构零部件的设计、选材以及制定材料的合理加工工艺提供必要的性能数据。因此，单轴拉伸是工程材料力学性能中最基本的试验方法。

1.2　静载拉伸试验

如不特别注明，拉伸试验是指采用光滑试样，在室温（10～35℃）、大气环境中，缓慢施加（应变速率≤10^{-1}/s）单向拉伸载荷，测定材料力学性能的方法。俗称静拉伸试验。

拉伸试验一般采用圆形或板状试样，如图 6.2-1 所示，l_0 为试样的标距长度。圆形试样便于测量径向应变，试样加工简便；板状试样用于测定板材和带材的拉伸性能。GB/T 228—2002规定了不同试样的尺寸范围，以实现轴向均匀加载。同时，国标中还规定了可使用非比例试样。

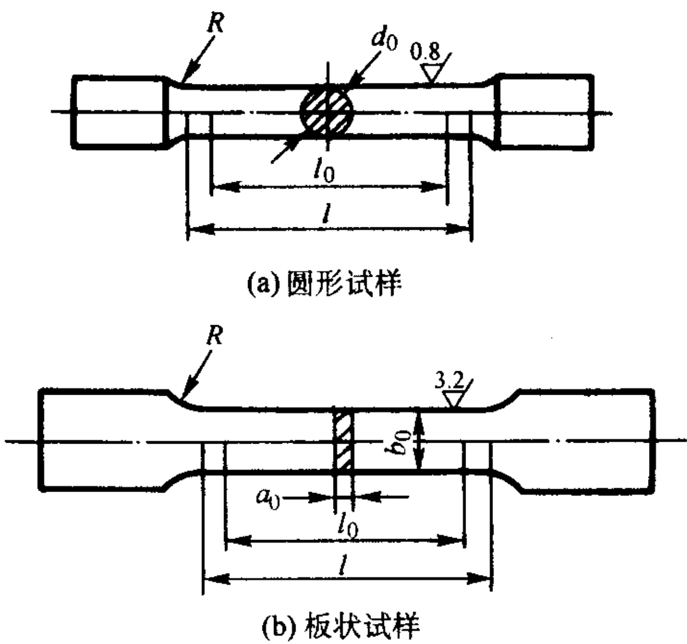


图 6.2-1　常用的拉伸试样

1.3　拉伸曲线

1）拉伸图　拉伸试验机通常带有自动记录或绘图装置，以记录或绘制试样所受的载荷 F 和伸长量 Δl 之间的关系曲线，$\Delta l=l-l_0$，其中 l 为加载后标距间的长度。这种曲线通常称为拉伸图，图 6.2-2a 即为退火低碳钢的拉伸图。

2）应力－应变曲线（拉伸曲线）　图 6.2-2b 为静拉伸试验的应力－应变曲线（拉伸曲线），又称工程应力－工程应变曲线。

工程应力由载荷除以试样的原始截面积得到，即

$$\sigma=F/A_0 \tag{6.2-2}$$

工程应变为伸长量除以试样的原始标距长度，即

$$\varepsilon=\Delta l/l_0 \tag{6.2-3}$$

比较图 6.2-2a 和图 6.2-2b，两者具有相同或相似的形状，但坐标刻度不同，意义不同。

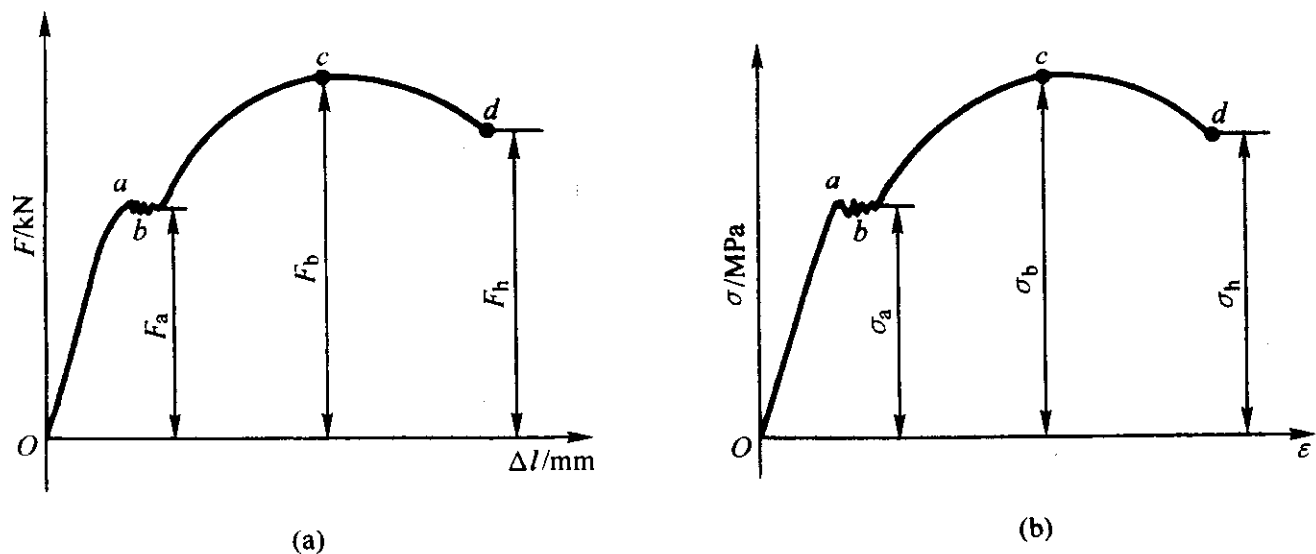


图 6.2-2　低碳钢的拉伸图和应力－应变曲线

3）真应力-真应变曲线　应当指出，上述应力-应变曲线只是在弹性及小塑性变形范围内才近似真实地反映了材料对外加载荷的响应，而其不真实性随塑性变形量的增大而增大，一旦由宏观均匀塑性变形过渡到局部集中塑性变形（如在金属材料中的缩颈和高聚物中的细颈）现象后，则尤其不真实了。此时，用真应力-真应变曲线表征材料对外力的响应更合理。

真应力为用真实瞬间截面积 A 除相应的载荷 F，即

$$\sigma_{真} = F/A \tag{6.2-4}$$

真应变为各瞬时应变的总和，即

$$\varepsilon_{真} = \sum\left(\frac{l_1 - l_0}{l_0} + \frac{l_2 - l_1}{l_1} + \cdots\right) = \int_0^l \frac{\mathrm{d}l}{l} = \ln\frac{l}{l_0} \tag{6.2-5}$$

由真应力和真应变表示的应力-应变曲线称为真应力-真应变曲线，如图6.2-3所示。由图可见，在弹性变形阶段，由于应变较小，一般均低于1%，且横向收缩小，因而真应力-真应变曲线与工程（或称标称）应力-应变曲线基本重合。随着应变的增大，两者的差增大，在试样标距长度的某处出现颈缩后，塑性变形集中在颈缩区，试样的截面积急剧减小，虽然工程应力随应变增加而减小，但真应力仍然增大，此时真应力和工程应力的差别才明显增大。

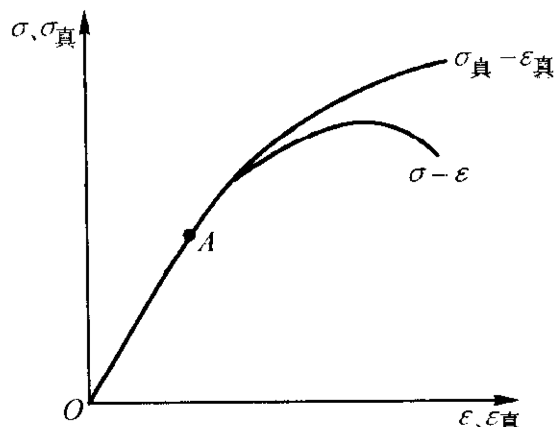


图6.2-3　真应力-真应变曲线

1.4　几种典型的应力-应变曲线

由于不同材料具有不同的化学成分和微观组织，在相同的试验条件下，会显示出不同的应力-应变响应。一般，在金属、陶瓷和高聚物材料中，拉伸条件下的应力-应变曲线大致有五种类型（见图6.2-4）。

1）纯弹性型　见图6.2-4a。有这种 $\sigma-\varepsilon$ 曲线的材料主要是大多数玻璃、陶瓷、岩石、横向交联很好的聚合物以及一些低温下的金属。

2）弹性-均匀塑性型　见图6.2-4b。有这种 $\sigma-\varepsilon$ 曲线的材料主要是许多金属及合金、部分陶瓷和非晶态高聚物。应该指出，把部分非晶态高聚物归入此类只是按应力-应变曲线的形式划分的。对于高聚物，尽管表观弹性变形和塑性变形与金属有相仿 $\sigma-\varepsilon$ 曲线形式，但在变形本质上是有一定区别的。

3）弹性-不均匀塑性型　见图6.2-4c。有这种 $\sigma-\varepsilon$ 曲线的材料主要是低温和高应变速率下的面心立方金属，其塑性变形常常不是通过滑移而是孪生。当孪生应变速率超过试验机夹头运动速度时，载荷会突然松弛而呈现记录到的锯齿形 $\sigma-\varepsilon$ 曲线。某些含碳原子的体心立方铁合金以及铝合金低溶质固溶体也有类似的 $\sigma-\varepsilon$ 曲线。

4）弹性-均匀塑性-不均匀塑性型　见图6.2-4d。有这种 $\sigma-\varepsilon$ 曲线的材料主要是一些体心立方的铁基合金和若干有色合金。它与图6.2-4b的不同仅在于中间增加了一段不均匀塑性屈服区（应变值大致为1%～3%）。

5）弹性-不均匀塑性-均匀塑性型　见图6.2-4e。有这种 $\sigma-\varepsilon$ 曲线的材料主要是一些结晶态高聚物和未经拉伸的线形非晶态高聚物。受拉结晶高聚物出现这种情况是因为有两个因素相互制约的结果。开始变形时，结晶高聚物中原有的结晶结构被破坏，随之发生细颈屈服，从而载荷下降，继续增加应变可促使变形最剧烈的区域重新组合成新的、方向性好和强度高的结晶结构。随着这种新结构的增多，应力-应变曲线再次上升，直至断裂。线形非晶态高聚物受拉伸在形式上呈现与此十分相似的 $\sigma-\varepsilon$ 曲线，但细颈的发生是由于线形大分子链段的取向而不是结晶结构的变化。

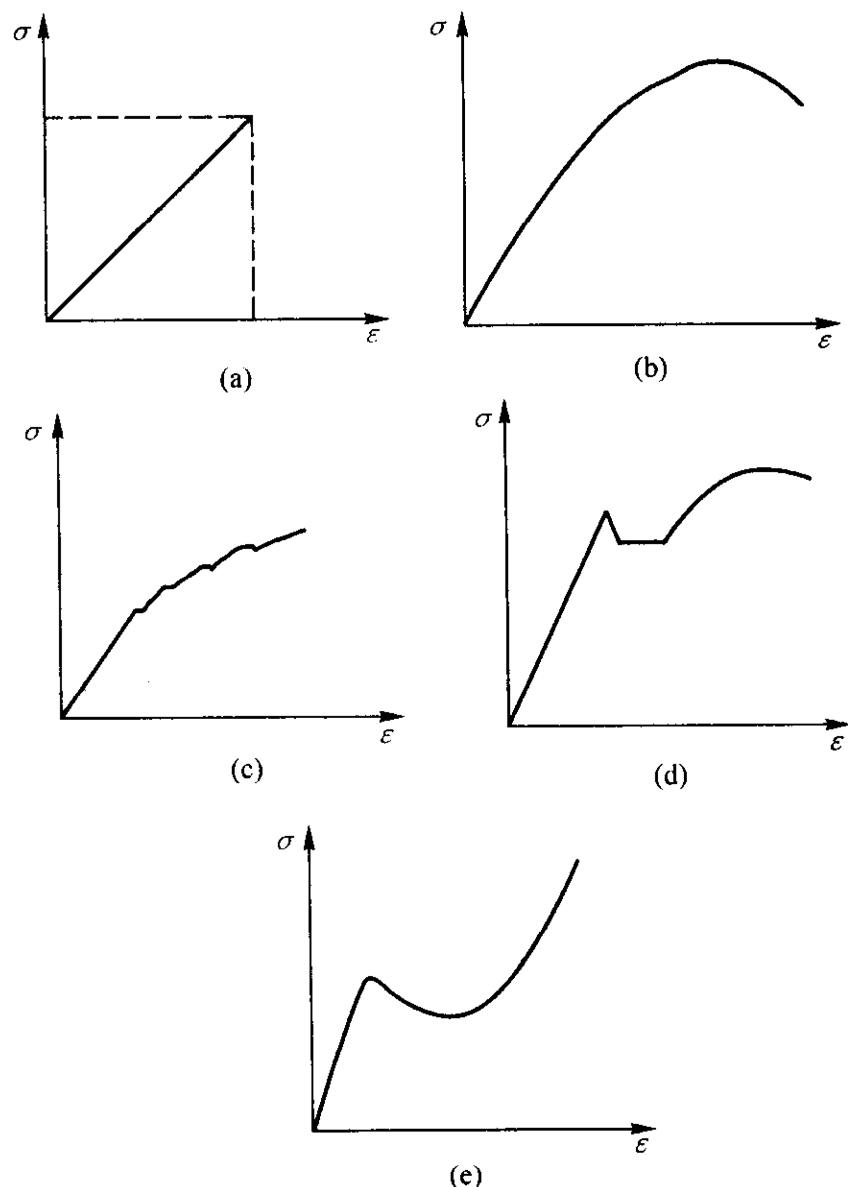


图6.2-4　五种类型的应力-应变曲线

2　材料的力学性能指标及其影响因素

材料的拉伸性能指标，又称力学性能指标，用应力-应变曲线上反映变形过程性质发生变化的临界值表示。力学性能指标可分为两类：反映材料对塑性变形和断裂的抗力的指标，称为材料的强度；反映材料塑性变形能力的指标，称为材料的塑性。

2.1　强度指标

2.1.1　屈服强度

原则上，材料的屈服强度应理解为开始塑性变形时的应力值。但实际上，对于连续屈服的材料，这很难作为判定材料屈服的准则，因为工程中的多晶体材料，其各晶粒的位向不同，不可能同时开始塑性变形。当只有少数晶粒发生塑性变形时，在应力-应变曲线上反映不出来。当较多晶粒发生塑性变形时，会造成宏观塑性变形的效果，从而在应力-应变曲线上反映出来。工程上采用规定一定的残留变形量的方法，确定屈服强度。

工程上常用的屈服标准有三种：

1）比例极限　应力-应变曲线上符合线性关系的最高应力值，用 σ_p 表示，超过 σ_p 时，即认为材料开始屈服。

2）弹性极限 试样加载后再卸载，以不出现残留的永久变形为标准，材料能够完全弹性恢复的最高应力值，用 σ_{el}表示，超过 σ_{el}时，即认为材料开始屈服。

上述二定义并非完全等同，有的材料，如高强度晶须，可以超出应力 - 应变的线性范围，发生较大的弹性变形。一般材料中弹性极限稍高于比例极限。工程上之所以要区分它们，是因为有些设计，如火炮筒材料，要求有高的比例极限，而另一些情况，如弹簧材料，要求有高的弹性极限。

3）屈服强度 以规定发生一定的残留变形为标准，如通常以0.2%残留变形的应力作为屈服强度，用 $\sigma_{0.2}$或 σ_{ys}表示。

上述定义都是以残留变形为依据的，彼此区别在于规定的残留变形量不同。现行国家标准将屈服强度规范为三种情况。

1）规定非比例伸长应力（σ_p） 试样在加载过程中，标距长度内的非比例伸长量达到规定值（以%表示）的应力，如 $\sigma_{p0.02}$，$\sigma_{p0.05}$等。

σ_p 通常用图解法测定，对有明显弹性直线段的材料，可利用自动记录的载荷 - 伸长（$P-\Delta L$）曲线，如图 6.2-5，自弹性直线段与伸长轴的交点 O 起，截取一相应于规定非比例伸长的线段 OC（$\overline{OC}=nL_e\varepsilon_p$，其中，$n$ 为拉伸图放大倍数，L_e 为引伸计标距，ε_p 为规定的非比例伸长率），过 C 点作弹性直线段的平行线 CA，交曲线于 A 点，A 点对应的载荷 P_p 即为所测定的非比例伸长载荷，规定非比例伸长应力由下式计算

$$\sigma_p = P_p/A_0 \tag{6.2-6}$$

2）规定残余伸长应力（σ_r） 试样卸载后，其标距部分的残余伸长达到规定比例时的应力，常用的为 $\sigma_{r0.2}$，即规定残余伸长率为 0.2%时的应力值。

测定 σ_r 通常用卸载法（如图 6.2-6），即当卸载后所得残余伸长为规定残余伸长载荷 P_r，规定残余伸长应力由下式计算

$$\sigma_r = P_r/A_0 \tag{6.2-7}$$

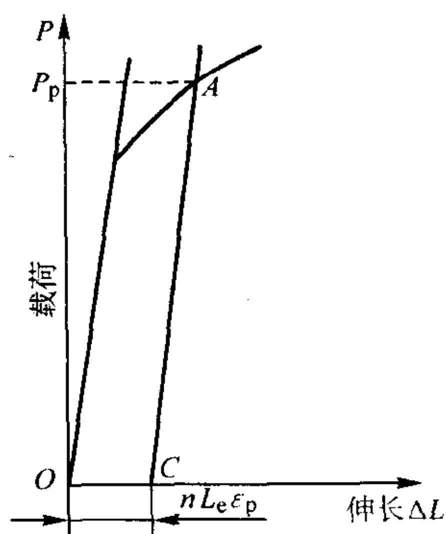


图 6.2-5 图解法测 σ_p

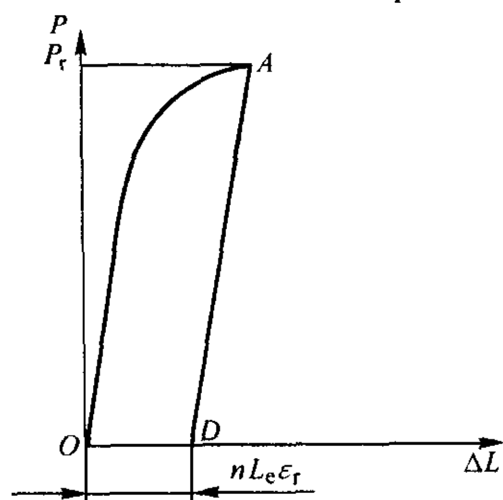


图 6.2-6 卸载法测 σ_r

3）规定总伸长应力（σ_t） 试样标距部分的总伸长（弹性伸长与塑性伸长之和）达到规定比例时的应力。应用较多的规定总伸长率为 0.5%、0.6%和 0.7%，相应地，规定总伸长应力分别记为 $\sigma_{t0.5}$、$\sigma_{t0.6}$和 $\sigma_{t0.7}$。

测定 σ_t 也用图解法，操作与测定 σ_p 相同，拉伸图横轴放大倍数不小于 50 倍。如图 6.2-7 所示，在 $P-\Delta L$ 曲线上，自曲线原点 O 起，截取相应于规定总伸长的线段 OE（$\overline{OE}=nL_e\varepsilon_t$，式中 ε_t 为规定总伸长率），过 E 点作纵轴平行线 EA 交曲线于 A 点，A 点对应的载荷即为规定总伸长的载荷，规定总伸长应力由下式计算

$$\sigma_t = P_t/A_0 \tag{6.2-8}$$

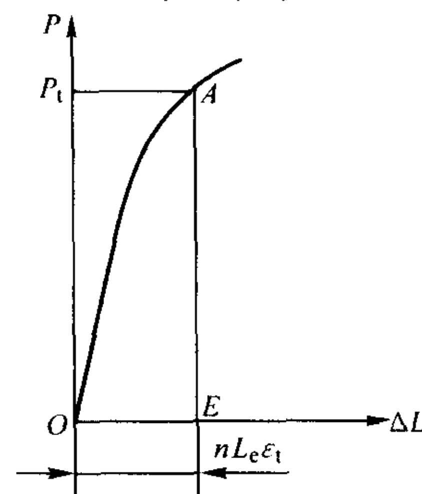


图 6.2-7 图解法测 σ_t

在上述屈服强度的测定中，σ_p 和 σ_t 是在试样加载时直接从应力 - 应变（载荷 - 位移）曲线上测量的，而 σ_r 则要求卸载测量。由于卸载法测定规定残余伸长应力 σ_r 比较困难，而且效率低，所以，在材料屈服抗力评定中，更趋于采用 σ_p 和 σ_t。σ_t 在测试上又比 σ_p 方便，而且不失 σ_p 表征材料屈服特征的能力，所以，可以用 σ_t 代替 σ_p，尤其在大规模工业生产中，采用 σ_t 的测定方法，可以提高效率。

对于不连续屈服即具有明显屈服点的材料，其应力 - 应变曲线上的屈服平台就是材料屈服变形的标志，因此，屈服平台对应的应力值就是这类材料的屈服强度，记作 σ_{ys}，按下式计算

$$\sigma_{ys} = P_y/A_0 \tag{6.2-9}$$

式中，P_y 为物理屈服时的载荷或下屈服点对应的载荷。

屈服强度是应用最广的一个性能指标。因为任何机械零件在工作过程中，都不允许发生过量的塑性变形，所以，机械设计中，把屈服强度作为强度设计和选材的依据。

2.1.2 抗拉强度

材料的极限承载能力用抗拉强度表示。拉伸试验时，与最高载荷 P_b 对应的应力值 σ_b 即为抗拉强度

$$\sigma_b = P_b/A_0 \tag{6.2-10}$$

对于脆性材料和不形成颈缩的塑性材料，其拉伸最高载荷就是断裂载荷，因此，其抗拉强度也代表断裂抗力。对于形成颈缩的塑性材料，其抗拉强度代表产生最大均匀变形的抗力，也表示材料在静拉伸条件下的极限承载能力。对于钢丝绳等零件来说，抗拉强度是一个比较有意义的性能指标。抗拉强度很容易测定，而且重现性好，与其他力学性能指标如疲劳极限和硬度等存在一定关系，因此，也作为材料的常规力学性能指标之一用于评价产品质量和工艺规范等。

2.1.3 实际断裂强度

拉伸断裂时的载荷 P_K 除以断口处的真实截面积 A_K 所得的应力值称为实际断裂强度 S_K

$$S_K = P_K/A_K \tag{6.2-11}$$

S_K 是真实应力，其意义是表征材料对断裂的抗力，因此，有时也称为断裂真应力。

2.2 塑性指标

材料的塑性变形能力，即塑性用伸长率和断面收缩率表示。

2.2.1　伸长率

试样断裂后的总伸长率称为极限伸长率，用 δ_K 表示，其条件应变值用式（6.2-12）计算

$$\delta_K = \frac{l_K - l_0}{l_0} \times 100\% \quad (6.2\text{-}12)$$

式中，l_K 为断裂后的标长。

对于形成颈缩的材料，其伸长量 $\Delta l_K = l_K - l_0$，包括颈缩前的均匀伸长 Δl_b 和颈缩后的集中伸长 Δl_c，即 $\Delta l_K = \Delta l_b + \Delta l_c$。因此，伸长率也相应地由均匀伸长率 δ_b 和集中伸长率 δ_c 组成，即

$$\delta_K = \delta_b + \delta_c \quad (6.2\text{-}13)$$

研究表明，均匀伸长率取决于材料的冶金因素，而集中伸长率与试样几何尺寸有关，即

$$\delta_c = \beta\sqrt{A_0}/l_0 \quad (6.2\text{-}14)$$

可以看出试样 l_0 越大，集中变形对总伸长率的贡献越小。为了使同一材料的试验结果具有可比性，必须排除试样几何尺寸的影响，为此必须对试样尺寸进行规范化，这只要使 $\sqrt{A_0}/l_0$ 为一常数即可。工程上规定了两种标准拉伸试样，$\sqrt{A_0}/l_0 = 11.3$ 和 5.65。对于圆形截面拉伸试样，相应于 $l_0 = 10d_0$ 和 $l_0 = 5d_0$，分别称为10倍和5倍试样。相应地，伸长率分别用 δ_{10} 和 δ_5 表示。由上述可见，$\delta_5 > \delta_{10}$。

2.2.2　断面收缩率

试样断裂后所得总断面收缩率称为材料的极限断面收缩率，用 ψ_K 表示，其条件应变值由下式计算

$$\psi_K = \frac{A_0 - A_K}{A_0} \times 100\% \quad (6.2\text{-}15)$$

式中，A_K 为试样断口处的最小截面积。与延伸率一样，断面收缩率 ψ_K 也由两部分组成，均匀变形阶段的断面收缩率和集中变形阶段的断面收缩率，但与延伸率不同的是，断面收缩率与试样尺寸无关，只决定于材料性质。

2.2.3　塑性指标间的关系

对于颈缩前，根据体积不变条件

$$l_0 A_0 = lA$$

$$l = l_0 + \Delta l = l_0\left(1 + \frac{\Delta l}{l_0}\right) = l_0\ (1 + \delta)$$

$$A = A_0 - \Delta A = A_0\left(1 - \frac{\Delta A}{A_0}\right) = A_0\ (1 - \psi)$$

于是得条件塑性指标间的关系

$$1 + \delta = \frac{1}{1 - \psi}$$

或

$$\delta = \frac{\psi}{1 - \psi} \quad (6.2\text{-}16)$$

上式表明，在均匀变形阶段 δ 恒大于 ψ。

上面的讨论，塑性指标都采用工程应变，为条件塑性指标。在拉伸过程中每一时刻的真应变 $d\varepsilon$ 为

$$d\varepsilon = dl/l$$

试样从 l_0 拉伸至 l 时，完成的真应变为

$$\varepsilon = \int d\varepsilon = \int_{l_0}^{l} \frac{dl}{l} = \ln\frac{l}{l_0} \quad (6.2\text{-}17)$$

于是真应变与条件塑性指标的关系为

$$\varepsilon = \ln\ (1 + \delta)\ = \ln\left(\frac{1}{1 - \psi}\right) \quad (6.2\text{-}18)$$

在颈缩开始以后，颈部的变形是非常复杂的，此时，条件塑性指标之间已不存在上式关系，但真实塑性应变与条件断面收缩率之间尚有如下关系

$$\varepsilon = \ln\frac{l}{l_0} = \ln\frac{A_0}{A} = 2\ln\frac{d_0}{d} = \ln\left(\frac{1}{1 - \psi}\right) \quad (6.2\text{-}19)$$

因此，试样断裂后，可通过测量断面收缩率 ψ_K，求得真实极限塑性 ε_f

$$\varepsilon_f = \ln\left(\frac{1}{1 - \psi_K}\right) \quad (6.2\text{-}20)$$

2.2.4　塑性指标的意义

伸长率和断面收缩率是工程材料的重要性能指标。设计零件时，不但要选择材料，提出强度要求，进行强度计算，同时还要提出对材料塑性的要求。因为在零件工作过程中，难免偶然过载，或者应力集中部位的应力水平超过材料屈服强度，这时，材料如果具有一定的塑性，则可用局部塑性变形松弛或缓冲集中应力，避免断裂，保证安全。如汽车齿轮箱的传动轴，选用中碳钢调质处理，要求 $\sigma_{0.2}$ 为 400～500 MPa，同时还要求 δ 不小于 6%～7%。这里对塑性的要求便是出于安全考虑。

另外，金属塑性变形能力是压力加工和冷成形工艺的基础。冷成形过程中，如冷弯、冲压等，为保证金属的流动性，必须具有足够的塑性，尤其材料均匀塑性变形能力的大小十分重要。

2.3　材料强度、塑性与韧性的合理配合

传统的零件设计方法首先是以材料的强度指标 σ_s、σ_b 或 σ_{-1} 为依据进行强度计算，然后考虑零件在油孔、键槽、尖角等处有应力集中和工作时会遇到难以预料的过载或偶然的冲击等情况，再凭经验对材料的塑性、韧性提出一定要求。通常材料的强度与塑性、韧性是相互矛盾的，强度高则塑性、韧性低。

传统的设计思想过分追求塑性、韧性，使零件普遍存在塑性、韧性有余而强度不足，故寿命不高。若适当降低对塑性、韧性的要求而提高强度，则会使零件寿命大幅度提高。但也不能走向另一个极端，认为强度愈高愈好。对于含裂纹的零构件，应适当降低强度，提高塑性、韧性，例如火箭发动机壳体，过分追求强度、反而使其寿命降低。

因此，材料的强度、塑性、韧性必须合理配合。对于以高周疲劳断裂为主要危险的零件，在 $\sigma_b < 1\,400$ MPa 范围内，材料的强度愈高，其疲劳强度亦愈高，则零件的寿命愈高，因此提高材料强度，适当降低塑性、韧性，对提高零件寿命有利。而且这类中低强度材料的断裂韧度较高，除工作在低温或尺寸较大的零件外，一般不易发生低应力脆断，故可以用工作应力 $\sigma \leqslant \sigma_s/k$ 来计算和选材，提高屈服强度可以提高零件的允许工作应力和减小零件的质量；若 $\sigma_b > 1\,400$ MPa，由于这类材料的强度对缺口、表面加工质量、热加工缺陷、冶金质量等都很敏感，随强度增加，其疲劳寿命反而降低；对于以低应力脆断为主要危险的零件，如中低强度钢制造的汽轮机转子、发电机转子、大型轧辊、低温或高压化工容器以及高强度钢制造的火箭发动机壳体等，这时材料的韧性比强度更重要，应该用断裂韧度来选材，即 $K_I = Y\sigma a^{1/2} < K_{IC}$。适当增加材料的塑性、韧性，牺性强度对提高零件寿命有利。

3　材料的弯曲特性

3.1　静弯曲试验的特点

对圆形或矩形横截面试样施加弯曲应力，一般直至断裂，测定其弯曲力学性能的方法，即为弯曲试验。GB/T 14452—1993 规定了有关的试验方法和数据处理规范。

弯曲试验的加载方式有两种：三点弯曲和四点弯曲（见图6.2-8）。后者有足够的均匀加载段，因而断裂的具体位置将由均匀加载段内相对较弱的组织来决定，可较好地反映材

料全面的品质；前者的断裂位置被固定在集中加载截面处，这种试验方法较简单，故常采用。

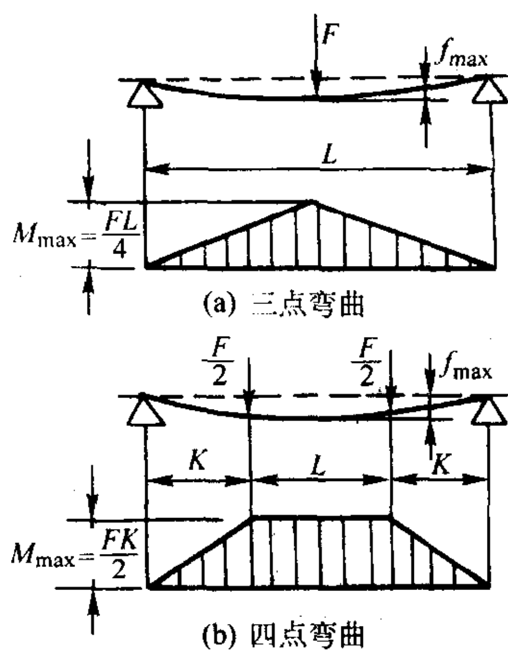


图 6.2-8 弯曲加载方式及其弯矩图

弯曲试验通常用矩形或圆形试样，通过记录载荷和最大挠度（$F-f_{max}$）之间的曲线来确定有关力学性能。图 6.2-9 表示三种不同材料的弯矩图。从试样的拉伸侧看，其应力状态与拉伸试验相仿，但从整体上说又存在着不同的特点。

对于高塑性材料，弯曲试验常达不到破坏的程度，其 $F-f_{max}$ 图最后部分可以很长（见图 6.2-9a)。因此，这类材料的弯曲试验难以测得材料的强度，且试验结果的分析也较复杂，故宜改用拉伸试验来测定其有关断裂的抗力。

对于脆性材料，弯曲图（见图 6.2-9c）比拉伸试验的几何外形简单，所以适用于测定加工不方便的脆性材料，如铸铁、工具钢、硬质合金乃至陶瓷材料的断裂强度和塑性。

对于高分子材料，也常用于测定其弯曲强度及模量，用以筛选配方或控制产品质量。

和扭转试验类似，弯曲试验中，其应力分布是不均匀的，表面处应力最大，故可较灵敏地反映材料的表面缺陷，以鉴别诸如渗碳、表面淬火等热处理零件的质量。

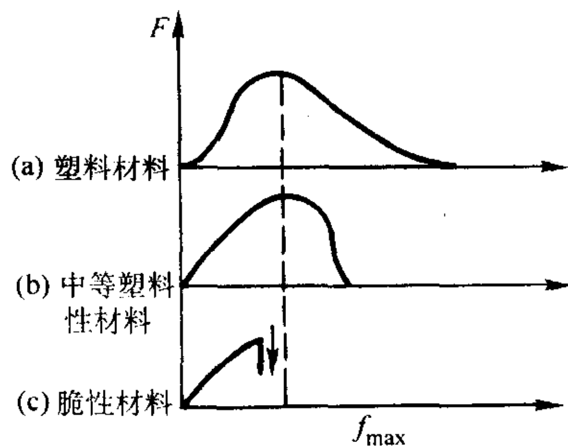


图 6.2-9 典型弯曲图

3.2 材料弯曲的力学性能指标

弯曲试样拉伸侧表面的最大正应力 σ_{max} 由下式给出：

$$\sigma_{max}=\frac{M_{max}}{W} \qquad (6.2\text{-}21)$$

式中，对三点弯曲试验，$M_{max}=\frac{FL}{4}$。对四点弯曲试验，$M_{max}=\frac{FK}{2}$（见图 6.2-8)。一般对脆性材料只测定断裂时的抗弯强度 σ_{bb}。由式（6.2-21）有

$$\sigma_{bb}=\frac{M_b}{W} \qquad (6.2\text{-}22)$$

式中，M_b 为试样断裂弯矩，由 $F-f_{max}$ 图中的 F_b 代入 M_{max} 表达式算出；W 为试样的弯曲截面系数，对圆柱试样，$W=\frac{\pi d^3}{32}$，对矩形试样，$W=\frac{bh^2}{6}$。与抗弯强度 σ_{bb} 相应的最大挠度 f_{max} 值可用百分表或其他位移传感器读出。

由 $F-f_{max}$ 曲线的直线部分可计算弯曲模量，对矩形试样，其弯曲模量为

$$E_b=\frac{ml^3}{4bh^3} \qquad (6.2\text{-}23)$$

式中，m 为 $F-f_{max}$ 直线段的斜率；l 为试样的跨距。

3.3 弯曲试验的应用

弯曲试验可用于测定灰铸铁的抗弯强度。灰铸铁的弯曲试件一般采用铸态毛坯圆柱试件，试验时加载速度不大于 0.1 mm/s。若试件的断裂位置不在跨距的中点，而在距中点 x 处（见图 6.2-8)，则抗弯强度应按式（6.2-22）计算。

弯曲试验可用于测定硬质合金的抗弯强度。硬质合金由于硬度高，难以加工成拉伸试件，故常做弯曲试验以评价其性能和质量，而且由于硬质合金价格昂贵，故常采用方形或矩形截面的小尺寸试件。常用的规格是 5 mm × 5 mm × 30 mm，跨距为 24 mm。

弯曲试验还可用于陶瓷材料的抗弯强度测定。由于陶瓷材料脆性大，测定抗拉强度很困难，故目前主要是测定其抗弯强度作为评价陶瓷材料性能的指标。

陶瓷材料的弯曲试件常采用方形或矩形截面的试件。考虑到实验结果的分散性，试件应从同一块或同质坯料上切出尽可能多的小试件，以便对实验结果进行统计分析。还应指出，试件的表面粗糙度对陶瓷材料的抗弯强度有很大的影响。表面愈粗糙，抗弯强度愈低。再则，磨削方向与试件表面的拉应力垂直，也会降低陶瓷材料的抗弯强度。

铸铁、工具钢、硬质合金以及陶瓷材料之所以常用弯曲试验测定其强度性能，是因为试件加工方便，试验操作简单，且不会出现拉伸试验时试件装夹偏斜对实验结果造成的影响，而且在一定程度上还能比较材料的变形能力。再则，拉伸试验时试件表面的正应力大，对表面缺陷敏感，故弯曲试验常被用于检验和比较渗碳层和表面淬火层的质量和性能。

4 材料的扭转特性

4.1 静扭转试验的特点

对试样施加扭矩，测量扭矩及其相应的扭角，一般扭至断裂，以便测定材料扭转力学性能的方法称为扭转试验，它是重要的力学性能试验方法之一。扭转试验采用圆柱形（实心或空心）试件，在扭转试验机上进行。扭转试样如图 6.2-10 所示。GB/T 10128—1988 规定了金属室温扭转试验的一些基本规范。

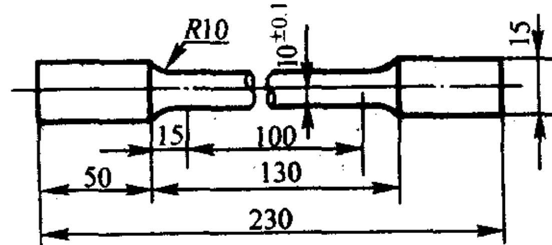


图 6.2-10 扭转试样

等直圆杆受到扭矩作用时，其中的应力应变分布如图 6.2-11 所示：在横截面上无正应力而只有切应力作用。在弹性变形阶段，横截面上各点的切应力与半径方向垂直，其大小与该点距中心的距离成正比；中心处切应力为零，表面处切应力最大（见图 6.2-11b)。当表层产生塑性变形后，各点的切应变仍与该点距中心的距离成正比，但切应力则因塑性

变化而降低，如图6.2-11c所示。在圆杆表面上，在切线和平行于轴线的方向上切应力最大，在与轴线成45°的方向上正应力最大，正应力等于切应力（见图6.2-11a）。

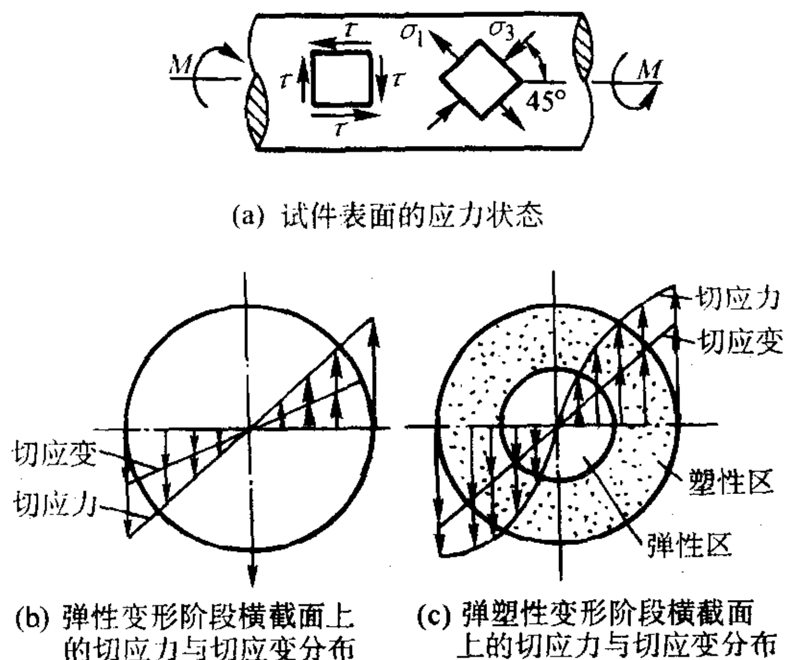


图6.2-11　扭转试样中的应力与应变

扭转试验具有如下的特点。

1）扭转时，应力状态的软性系数 $\alpha=0.8$，所以它可以为拉伸试验时表现为脆性的材料测定有关塑性变形的抗力指标，如淬火低温回火钢的塑性。

2）用圆柱形试样进行扭转试验时，整个长度上塑性变形始终是均匀的，其截面及标距长度基本保持不变，没有颈缩现象。因此，可用以精确评定拉伸时出现颈缩的高塑性金属材料的变形抗力和变形能力（见图6.2-12）。

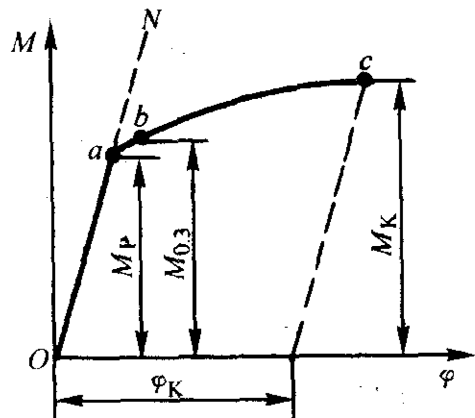


图6.2-12　扭转负荷-变形曲线

3）扭转试验中的正应力和切应力数值上大体相等，而常用金属结构材料的 $\sigma_{K真}>\tau_K$，所以扭转试验是测定这类材料切断抗力 τ_K 的最可靠方法。

4）扭转试验可明确地区分金属材料最终断裂的方式：正断和切断。对于塑性材料，断口平整垂直于试样轴向，通常有回旋状的塑性变形痕迹，如图6.2-13a所示，这是由切应力作用的切断。对于脆性材料，断口呈螺旋状，与试样轴约成45°角，如图6.2-13b所示，为正应力作用下的正断。图6.2-13c所示为另一种切断形式，断口出现层状或木片状形式。当存在较多非金属夹杂物或偏析的金属材料经过轧制、锻造或拉拔后，顺轧制、拉拔方向进行扭转试验时，常出现这种断口。这是因为材料的轴向切断抗力比横向的低而造成的。

5）扭转试验时，圆柱试样截面上的应力分布不均匀，表面处应力最大，心部应力最小。所以，它不能显示材料的体积缺陷，但对表面缺陷以及表面硬化层的性能都很敏感。因此，可利用扭转试验研究或检验工件热处理的表面质量和各种表面强化工艺的效果。

6）扭转试验时，试件受到较大的切应力，因而被广泛应用于研究有关初始塑性变形的非同时性的问题，如弹性后

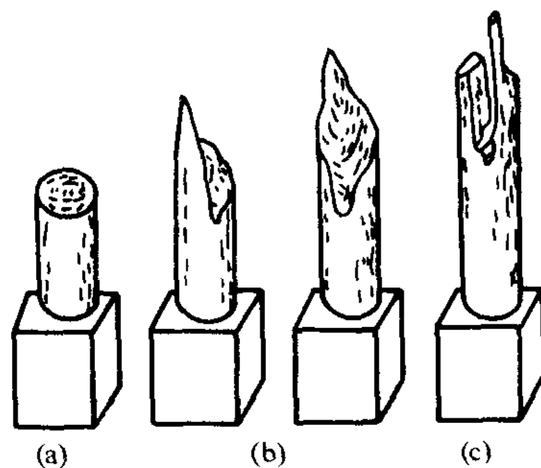


图6.2-13　扭转试样的断口形式

效、弹性滞后以及内耗等。

综上所述，扭转试验可用于测定塑性材料和脆性材料的剪切变形和断裂的全部力学性能指标，并且还有着其他力学性能试验方法所无法比拟的优点。因此，扭转试验在科研和生产检验中得到较广泛地应用。然而，扭转试验的特点和优点在某些情况下也会变为缺点。例如，由于扭转试件中表面切应力大，越往心部切应力越小，当表层发生塑性变形时，心部仍处于弹性状态（见图6.2-11c）。因此，很难精确地测定表层开始塑性变形的时刻，故用扭转试验难以精确地测定材料的微量塑性变形抗力。

4.2　材料扭转的力学性能指标

图6.2-12所示的扭转（$M-\varphi$ 曲线）图是用圆柱试样（直径 d_0、标距长度 l_0）测定的。其中 φ 为相距 l_0 的两截面间的相对扭转角，M 为扭矩。显然，该图与拉伸试验测定的真应力-真应变曲线很相似。

试样在弹性范围内每一瞬间的表面切应力 τ 和表面切应变 γ 由下式给出

$$\tau=\frac{M}{W}\tag{6.2-24}$$

$$\gamma=\frac{\varphi d_0}{2l_0}\tag{6.2-25}$$

式中，W 为试样的截面系数，对于圆柱试样，$W=\frac{\pi d_0^3}{16}$。已知 τ 和 γ，就可求出切变模量

$$G=\frac{\tau}{\gamma}=\frac{32Ml_0}{\pi\varphi d_0^4}\tag{6.2-26}$$

和拉伸图（$F-\Delta l$）类似，由已知的扭转图（$M-\varphi$）中，还可以求出扭转比例极限 τ_P、扭转屈服强度 $\tau_{0.3}$ 和扭转条件强度极限 τ_K，即

$$\begin{cases}\tau_P=\dfrac{M_P}{W}\\ \tau_{0.3}=\dfrac{M_{0.3}}{W}\\ \tau_K=\dfrac{M_K}{W}\end{cases}\tag{6.2-27}$$

静强度设计中常用到如下三种扭转抗力。

1）计算扭转比例极限用的扭矩 M_P　$M-\varphi$ 图上某一点对 M 轴的正切比直线的正切值大50%时的扭矩。因此，这种极限实质上是一种规定条件的比例极限。

2）扭转屈服强度　残留变形规定值为0.3%。这是为了能与拉伸时的屈服强度的残留变形规定值0.2%相当。因为真实主切应变 γ_{max} 和真实主正应变 ε_1、ε_2，ε_3 之间的关系是 $\gamma_{max}=\varepsilon_1-\varepsilon_3$。按胡克定律，$\varepsilon_1=\frac{\sigma_1}{E}-\nu\left(\frac{\sigma_2+\sigma_3}{E}\right)$，$\varepsilon_3=\frac{\sigma_3}{E}-\nu\left(\frac{\sigma_1+\sigma_2}{E}\right)$。单向拉伸时，$\sigma_2=\sigma_3=0$，若取 $\nu=0.5$

（考虑已进入弹塑性状态），则 $\varepsilon_1=\frac{\sigma_1}{E}$，$\varepsilon_3=-0.5\frac{\sigma_1}{E}$，于是得 $\gamma_{max}=1.5\varepsilon_1$。这就是单向拉伸时真实切应变与真实正应变之间的关系。所以扭转时规定残留应变为0.3%就可以理解了。严格说，上述计算中的弹性模量 E 应取形变硬化模数 D 代替，所以这种对应关系也还是一种近似估计。

3）扭转强度极限　$\tau=M/W$ 只是弹性变形过程中的应力表达式（见图6.2-11b），在扭转强度极限 τ_K 中应用则失真较大。因为塑性变形后，切应变虽能保持直线分布，但在塑性区的切应力已重新分布，表层有所下降而呈曲线分布（见图6.2-11c），所以按 $\tau_b=\frac{M_b}{W}$ 计算的 τ_b 是一种条件扭转强度极限。只有陶瓷等很脆的材料，在扭断时也没有明显塑性变形时，上述计算的 τ_b 值才比较真实。

5　材料的压缩特性

单向压缩应力状态的软性系数 $\alpha=2$，故用于测定脆性材料诸如铸铁、铸铝合金、轴承合金和建筑材料的力学性能。对于塑性材料，只能压扁不能压破，所以试验只能测定弹性模量、比例极限、弹性极限和屈服强度等指标而得不到压缩强度极限。金属压缩试验方法由GB/T 7314—1987规定。

试样通常用圆柱形，图6.2-14a为短圆柱形试样（$d_0=10\sim20$ mm，$h_0=2.5d_0\sim3.5d_0$），供破坏试验用；图6.2-14b为长圆柱试样（$d_0=10\sim20$ mm），供测定弹性及微量塑性变形抗力用。试样端面加工要求很高，两端面平行并和轴线垂直，表面粗糙度 R_a 为0.8～0.2 μm，以减少端面的摩擦对测定性能的影响。高分子材料的试样除采用图6.2-14a所示的圆柱试样外，还采用长方体和薄板条两种类型，后者试验时需要设置侧向支撑夹具以防屈曲。

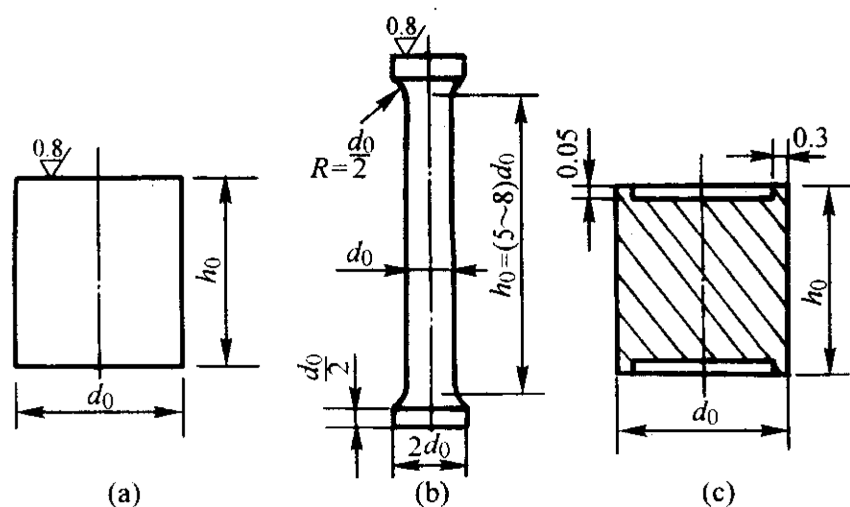


图6.2-14　压缩试样

压缩试验通常测出压力和变形的关系，即 $F-\Delta h$ 曲线。图6.2-15分别给出了铸铁和低碳钢两类典型 $F-\Delta h$ 曲线。由图可见，塑性好的材料只能压缩变形，不能破坏，如图中的曲线2最后一直上升。所以，压缩试验主要用于脆性材料和低塑性材料，以显示在静拉伸、扭转、弯曲试验中不能显示的材料在韧性状态下工作的行为，如图中的曲线1所示。和拉伸试验相似，当试样进入弹塑性状态后，所给出的强度指标，特别是压缩强度极限也有条件值 $\sigma_{bc}=\frac{F_{bc}}{A_0}$，和真实值 $\sigma_{bc真}=\frac{F_{bc}}{A_K}$ 之分，而且 $\sigma_{bc}\geqslant\sigma_{bc真}$。两者的关系是

$$\sigma_{bc}=(1+\psi_c)\ \sigma_{bc真}\tag{6.2-28}$$

$$\sigma_{bc真}=(1+\varepsilon_c)\ \sigma_{bc}\tag{6.2-29}$$

式中，相对压缩率 $\varepsilon_c=\frac{h_0-h_K}{h_0}\times100\%$，相对断面扩展率 $\psi=\frac{A_K-A_0}{A_0}\times100\%$。

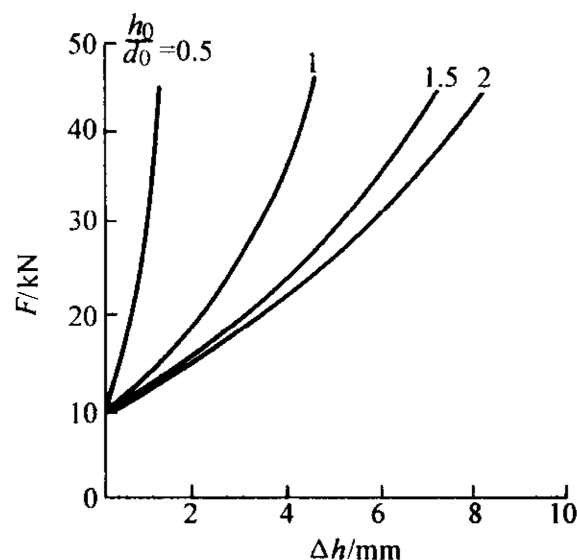


图6.2-15　压缩曲线

压缩时，端面存在很大的摩擦力，影响试验结果。首先，这种摩擦力阻碍试样端面的横向变形，造成上下端面小而中间凸的腰鼓形；其次，端面摩擦力引起附加变形阻力会提高总的变形抗力，降低了变形能力，如图6.2-16所示。由图可见，高径比（h_0/d_0）$\geqslant1.5$，端面摩擦力的影响就可以大大下降。端面摩擦力还影响到破坏的形式，图6.2-17中分别给出了脆性材料有端面摩擦（见图a）和没有端面摩擦（见图b）时的破坏情况。因此，压缩试验中应尽量减小端面摩擦，从而使试验结果稳定。通常采用一定高度的试样，当然也不能太长，以免造成纵向失稳。通常规定 $h_0/d_0=1\sim3$，并且要求压头和试样端面加工光洁度高，试验时端面要涂润滑油脂，必要时可用图6.2-14c所示的带蓄油槽的试样，试验效果很好。

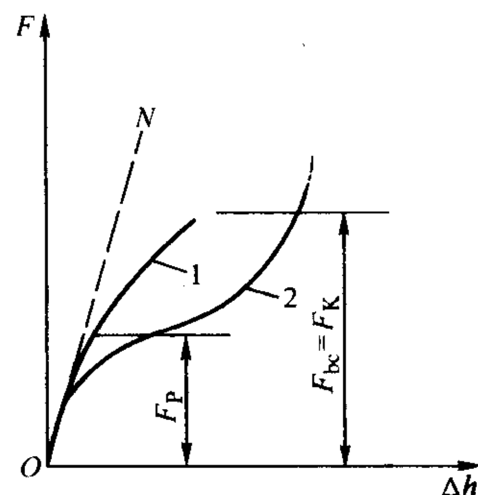


图6.2-16　h_0/d_0 对紫铜试验的影响

1—$h_0/d_0=1$；2—$h_0/d_0=2$

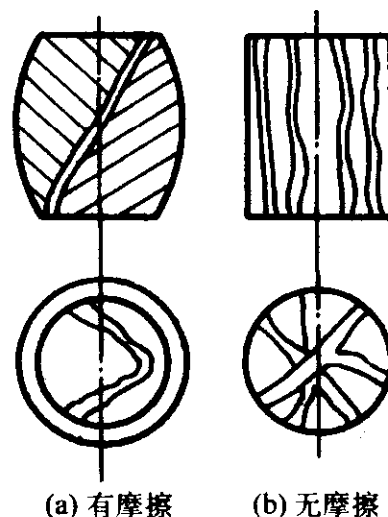


图6.2-17　端面摩擦力对压缩破坏的影响

6 缺口对静加载下力学性能的影响

由于实际的机器结构总是由多个零部件组成，而且它们之间必有某种联接以传递载荷或完成某种预定的工作，所以这些零部件上也就必有各种螺纹、台阶、孔洞、键槽等几何不均匀性，从一般意义上说它们都是缺口体，而缺口的引入会引起一系列力学效应，因此有必要研究这些效应对材料力学性能的影响。

6.1 缺口效应

以光滑试样为例，缺口的存在将引起应力集中、多轴应力状态、应变集中和局部应变速率增大等缺口效应。

图 6.2-18 所示为圆截面缺口试样单轴静拉伸时在弹性变形范围内的应力分布。轴向应力 σ_l 和切向应力 σ_t 都是在缺口顶端表面处最大，但径向应力 σ_r 的最大值却位于离表面一定距离处，而在缺口表面处等于零。通常用最大轴向应力 σ_{lmax} 与该截面上的平均应力 σ_m 之比给出理论应力集中系数 K_t 值：

$$K_t = \frac{\sigma_{lmax}}{\sigma_m} \tag{6.2-30}$$

对于椭圆缺口，$K_t = 1 + 2\sqrt{a/\rho}$，式中 a 为椭圆长轴之半，ρ 为椭圆长轴端点处的曲率半径 $\rho = b^2/a$；对于圆缺口，由于 $a = \rho$，则 $K_t = 3$。

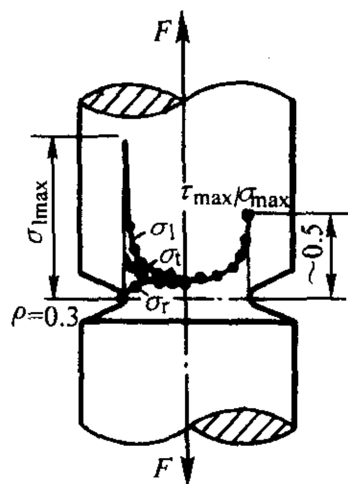
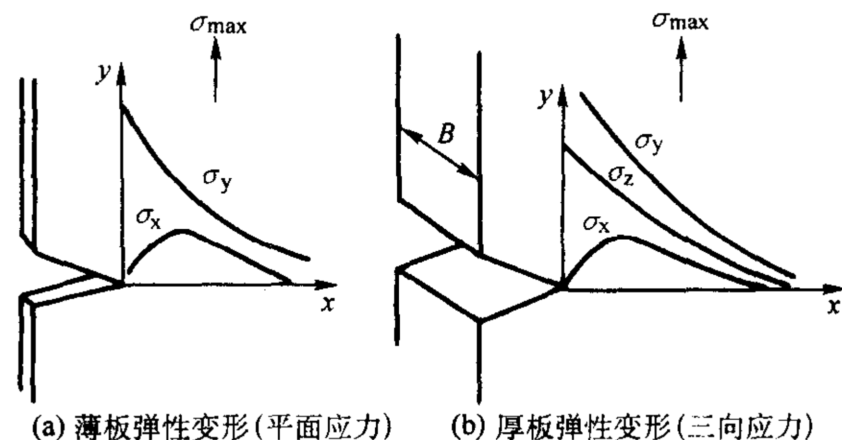
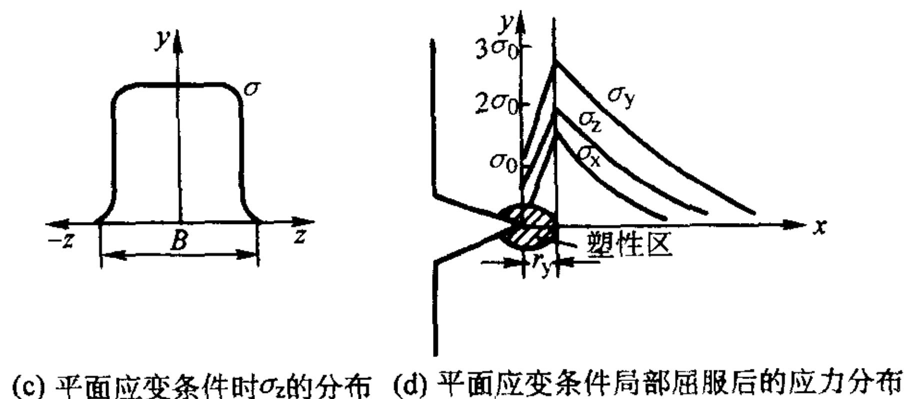


图 6.2-18　缺口试样弹性变形时应力分布示意图

缺口更重要的作用是在缺口根部出现多轴应力状态，对材料屈服和塑性变形起到约束作用。讨论较简单的薄板缺口试样如图 6.2-19 的应力分布。薄板试样受 y 方向拉力作用后，当 $\sigma_y < \sigma_s$ 时，试样产生的弹性变形应力分布如图 6.2-19a，由于纵向伸长，必然引起横向（板宽方向）收缩。假设在缺口附近的不同距离内，取若干同样大小的微单元，离缺口最近的微单元受 σ_y 最大，产生的纵向伸长也最大，相应的横向收缩也应最大，与其相邻的微单元因 σ_y 相对较小，其横向收缩也较小。即这些微单元在 x 方向的收缩量各不相等。各微单元的横向收缩，将引起相邻微单元间的分离。但实际上，材料在变形中是保持连续的，各微单元被联在一个整体内，不能自由收缩，即受到约束。也就是在 x 方向存在一个拉力，将它们拉在一起，这个力就是 σ_x。在缺口自由表面，因不存在对 x 方向的约束，可以自由收缩，故 $\sigma_x = 0$。在离缺口不远处，σ_x 有一极大值，这是因为在缺口附近 σ_y 的应力梯度很大，相应的微单元间的横向收缩差也大，所以在很小的 x 距离内，σ_x 便升高到最大值，随着距离 x 增加，σ_y 逐渐平缓，各单元间横向收缩差也减小，σ_x 又逐渐减小。于是带有缺口的薄板试样在弹性变形阶段的应力分布为双向应力 σ_x 和 σ_y。板厚较大时，如图 6.2-19b 所示，缺口顶端的变形除受 x 方向约束外，还受到板厚度方向（z 方向）的约束。在试样前后表面，z 方向无约束，为平面应力状态。但在缺口截面内部均存在 z 方向的约束，使 $\sigma_y \neq 0$，$\sigma_x \neq 0$，$\sigma_z \neq 0$，成三向应力状态。缺口截面上，由表面向内部，应力状态是变化的，由平面应力状态过渡为三向应力状态。



(a) 薄板弹性变形(平面应力)　(b) 厚板弹性变形(三向应力)



(c) 平面应变条件时σz的分布　(d) 平面应变条件局部屈服后的应力分布

图 6.2-19　侧面带有缺口的薄板和厚板受拉伸时的应力分布

6.2 应变集中和缺口截面应力分布的变化

缺口处的弹性应力分布表明，表面有最大轴向应力 σ_{lmax}，所以会在缺口根部发生屈服。屈服以后，材料的应力-应变呈非线性关系，通常应力集中的效果有所下降，而应变集中效果则相应提高。在弹塑性变形条件下，应力集中系数 K_σ 和应变集中系数 K_ε 分别定义为

$$\begin{cases} K_\sigma = \dfrac{\sigma_N}{\sigma_m} \\ K_\varepsilon = \dfrac{\varepsilon_N}{\varepsilon_m} \end{cases} \tag{6.2-31}$$

式中，σ_N 和 ε_N 为缺口根部的真实应力和应变；σ_m 和 ε_m 为缺口试样的名义应力和名义应变。为找出已进入弹塑性状态的缺口应力和应变，可求助于 Neüber 提出的关系式，即

$$K_\sigma K_\varepsilon = K_t^2 \tag{6.2-32}$$

把式（6.2-31）代入式（6.2-32），得

$$\sigma_N \varepsilon_N = K_t^2 \sigma_m \varepsilon_m \tag{6.2-33}$$

在给定几何形状和给定外力的情况下，上式右端为常数。另一方面，不管 σ_m 和 ε_m 还是 σ_N 和 ε_N，它们都应服从材料的应力—应变关系 $\sigma = f(\varepsilon)$。所以存在

$$\sigma_N = f(\varepsilon_N) \tag{6.2-34}$$

联解式（6.2-33）和式（6.2-34），则缺口根部表面上弹塑性应力 σ_N 和应变 ε_N 值即为图 6.2-20 所示两条实线的交点。由 Neüber 关系虽然只能给出缺口根部表面的应力和应变，但它有助于说明图 6.2-21 所示的缺口根部截面由表面到中心逐渐进入弹塑性状态后，轴向应力 σ_l 分布随外力增大的演变过程。

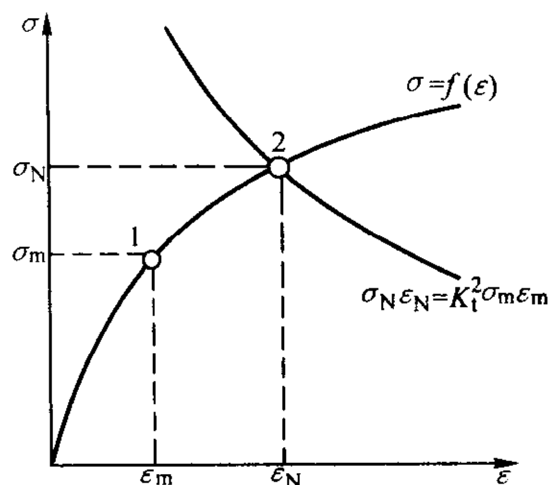


图 6.2-20　Neüber 关系的图解

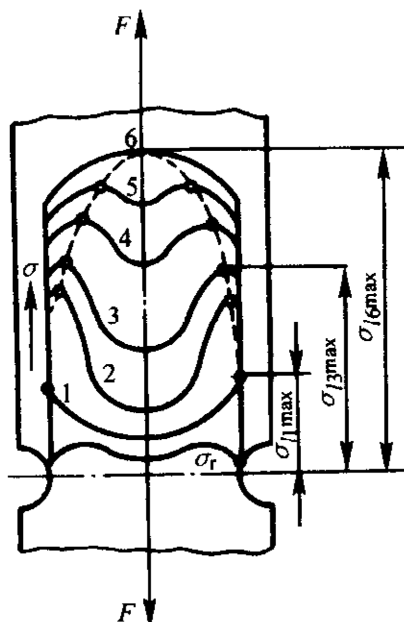


图 6.2-21　试样缺口根部截面由表及里逐步进入塑性状态后的 σ_1 应力分布的变化

由图 6.2-20 可见，在给定外力和缺口几何条件下，σ_N，ε_N 和 $K_t\sigma_m$，$K_t\varepsilon_m$ 都处于 $\sigma_N\varepsilon_N = K_t^2\sigma_m\varepsilon_m$ 所决定的双曲线上，而且随外力增大，σ_N 和 ε_N 与 $K_t\sigma_m$ 和 $K_t\varepsilon_m$ 的差别愈来愈大。所以，在缺口根部截面上，由表面开始部分进入弹塑性状态后，表层应力就比按弹性集中应力所预计的应力有所下降，而且随表层屈服深度加大而下降更大。因而由弹性状态进入了弹塑性状态的给定外力条件下，由缺口根部表面算起会出现一个一定深度的塑性区，而且 σ_1 的弹性分布（见图 6.2-21 曲线 1）过渡到弹塑性分布（见图 6.2-21 曲线 2），σ_{lmax}也由表面进入弹性区与弹塑性区交界的内部某点（$\sigma_{l_{1max}} \rightarrow \sigma_{l_{2max}}$）。随着外力的增大，塑性区不断扩大，$\sigma_{lmax}$点也就不断向内移动，只要材料的断裂抗力水平仍大于 σ_{lmax}，则此过程可以一直演变到图 6.2-21 中曲线 6 所示的分布，σ_{lmax}点最后到达缺口截面的中心（如图 6.2-21 所示的 σ_{l6max}）。

由于缺口应变集中的存在，在外加应变速率给定时，缺口根部的应变速率可能提高几个数量级，提高的程度随缺口根部的曲率半径减小而增大。

6.3　缺口静拉伸及斜拉伸试验

缺口对静拉伸力学性能的影响要视材料原来在光滑试样静拉伸时的塑性变形能力而定。

对原来呈脆性状态的材料，缺口根部截面在弹性应力集中的情况下，表面最大轴向应力 σ_{lmax}（见图 6.2-21 的曲线 1）就可以在外载荷增大时达到材料的断裂抗力，并萌生裂纹而引起过早的脆性断裂。其断口呈完全的放射状。此时，缺口试样的断裂强度 σ_{bN}将低于光滑试样的断裂强度 σ_b。对于原先有一定塑性（断裂抗力高于屈服应力）的材料，引入缺口则会有缺口根部出现塑性区和区内应力峰值被削减的过程，而且 σ_{lmax}向内部移动。当外加载荷增大到使 σ_{lmax}达到材料的断裂抗力时，则在此弹塑性区的交界处产生微裂纹。微观上属于微孔聚合机制，有韧窝形貌；宏观上则出现纤维区，然后继续向中心的弹性区呈放射状扩展（正好与光滑试样中颈缩截面的放射区扩展方向相反），中心为最后断裂区。在这种情况下的 σ_{bN}既可稍高于 σ_b，也可略低于 σ_b，这要看材料原来的塑性能力而定。

对原先有足够塑性即断裂抗力远大于屈服抗力的材料，当缺口较钝（$K_t < 2$）时，随载荷增大，塑性区不断向内扩展，σ_{lmax}也不断增大并向中心移动（见图 6.2-21），甚至可出现 σ_{lmax}达到中心，试样全面屈服。这时缺口试样也会出现颈缩，结果断裂及断口情况就会呈现与光滑试样有颈缩时类似的杯锥形断口。对尖锐缺口（$K_t > 6$），此时表层应力集中及三向应力状态严重，裂纹源仍可在表层发生，并向内扩展。不过，对中等塑性材料，向内的放射形扩展被抑制，代之以一层一层地纤维状向内扩展，最后在中心断裂。对于这两类缺口试样，都有 $\sigma_{bN} > \sigma_b$。

应当指出，缺口的应力集中和三向应力状态与缺口几何形状有关，其中缺口曲率半径 ρ 对应力集中很敏感，而缺口深度则影响三向应力状态。因此，用缺口试样来鉴别缺口对静拉伸力学性能的影响时，试样的形状要有所规定，以使数据有可比性。常用的缺口试样如图 6.2-22 所示，其中缺口角 $\omega = 45° \sim 60°$，缺口处曲率半径 $\rho \leqslant 0.2$ mm，缺口直径 $d_N = 7 \sim 20$ mm，$d_N/d_0 = 0.7$。对比用的光滑试样直径应等于 d_N。此外，应防止缺口处发生脱碳而造成 σ_{bN} 提高（可达 25% ~ 30%），所以试样加工时宜在热处理后再开缺口。

通常用缺口试样的强度 σ_{bN}对光滑试样的强度 σ_b 的比值——缺口强度比（NSR）来表示静拉伸强度对缺口的敏感度。

$$NSR = \frac{\sigma_{bN}}{\sigma_b} \qquad (6.2\text{-}35)$$

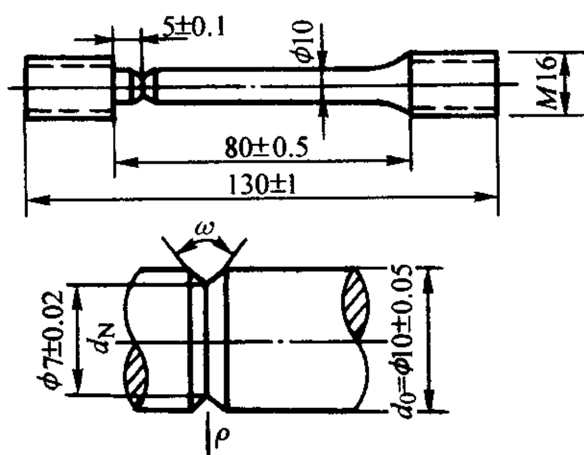


图 6.2-22　缺口拉伸试样

试验表明，NSR < 1 的材料主要是一些明显的脆性材料，如铸铁、低温回火的高碳工具钢，而绝大多数金属材料均表现为 NSR > 1。但后者同样表现出对缺口的敏感性。

表 6.2-2 给出了结构钢 40Cr 在典型热处理状态下的有关数据，其中尽管 NSR = 1.23 ~ 1.55，但 $\psi_N << \psi$。由此可见，NSR 只能表示材料对缺口敏感的趋势，比 1 愈大，材料愈不敏感，不能把 NSR = 1 作为敏感与否的分界。还要指出，当 NSR > 1，$\sigma_{bN} > \sigma_b$，尤其不可得出引入缺口可使材料强化的结论，因为实际上是缺口引起的三向应力状态抑制了材料的塑性变形，部分或全部抑制了光滑试样生产的颈缩现象，从而使缺口拉伸试样的 σ_{bN}接近于光滑试样的 $\sigma_{K真}$（见表 6.2-2）。由此可见，$\sigma_{bN} > \sigma_b$ 本质上并没有提高材料的强度，而且这种表观现象是以大大牺牲材料的塑性变形能力为代价的。

表 6.2-2　40Cr 钢的光滑、缺口拉伸和缺口偏斜拉伸试验的数据

热处理	光滑试样			缺口试样				$\frac{\sigma_{bN}}{\sigma_b}$ (NSR)	$\frac{\sigma_{bN}}{\sigma_{K真}}$
	σ_b/MPa	ψ/%	$\sigma_{K真}$/MPa	σ_{bN}/MPa	ψ_N/%	σ_{bN}^4/MPa	σ_{bN}^8/MPa		
300℃等温淬火	1 334	52.0	2 059	1 834	5.0	1 422	1 010	1.37	0.89
345℃等温淬火	1 089	61.5	1 844	1 687	8.6	1 657	1 285	1.55	0.91
淬火+420℃回火	1 491	50.5	2 246	1 837	1.5	1 226	441	1.23	0.82
淬火+510℃回火	1 260	51.5	1 961	1 893	7.3	1 510	940	1.50	0.97

鉴于上述情况，可以认为，如果引入比缺口拉伸更硬的应力状态，则有可能较真实地反映材料对硬性应力状态的敏感性。所以，对重要的高强度螺钉类紧固件用的材料，在缺口拉伸的基础上还要进行偏斜拉伸试验。图 6.2-23 所示为其试验装置，即在试样头部与试验机夹头之间垫一块有倾斜角 α 的垫圈，α 有 0°，4°和 8°三种，以得到不同角度偏斜试验的结果，相应的拉伸强度以 σ_{bN}^{α} 表示。表 6.2-2 所列的 σ_{bN}^{8} 数据表明，等温淬火处理显著优于调质处理，而在缺口拉伸中，它们对缺口敏感性上差异则未能充分显示出来。

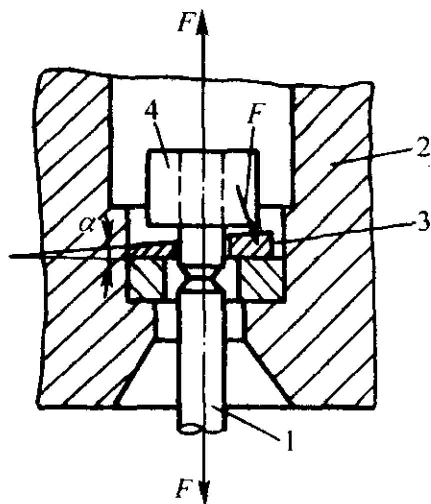


图 6.2-23　缺口拉伸偏斜试验

1—试样；2—试验机夹头；3—垫圈；4—螺纹夹头

7　材料静强度设计的一般方法

7.1　强度判据

常规机械静强度设计中，假设制造机械零构件的材料性能是均匀的、各向同性的、连续的实体。其静强度设计的计算步骤是：由理论力学确定零构件所受外力，材料力学（有时采用弹性力学或塑性力学）计算其内力，再由机械原理和机械零件确定其结构尺寸和形状，最后计算该零构件的工作应力或安全系数。

一般以公式表示，即

$$\sigma \leqslant [\sigma] \tag{6.2-36}$$

满足式（6.2-36），则构件是安全的；反之则不安全。式中，σ 为零部件计算部位的实际应力，可以用理论分析计算方法，如材料力学、弹塑性力学、有限元法求得。也可用实验的方法，如光弹法、云纹法、电阻应变测量技术获得。$[\sigma]$ 为许用应力，它不仅与零部件中应力确定的精确度有关，还与机械的运行工况、零部件材料的选用和检查方法等有关系，其确定是个复杂的问题。

机械零构件中的实际应力，除以零构件所用材料的极限应力（如材料的抗拉强度 σ_b，屈服强度 σ_s 等）的数值，称为工作安全系数，用 n 表示。安全系数 n 是考虑到实际结构中可能有的缺陷和其他意想不到的或难以控制的因素（如计算方法的不准确性、载荷估计的不准确性等），用来保证所设计的机械零构件有足够的强度安全储备量，保证在最大工作载荷下，其工作应力不超过制造零构件材料的极限应力。当计算得到的实际安全系数 n 大于或等于许用安全系数 $[n]$ 时，表示所设计的机械能保证安全运行。即

$$n \geqslant [n] \tag{6.2-37}$$

由于材料不同，其力学性能也不一样，所以各种材料在相同载荷作用下的许用应力值是不同的。如果采用安全系数表示时，各种材料的安全系数只有一个值，故表达形式比许用应力要简便得多。

一般情况下，在设计新机械结构的初算中，常用许用应力；在校核零部件强度时，则常用安全系数。采用过大的安全系数或过小的许用应力，会使所设计的零部件过分笨重而浪费材料；如果安全系数过小，或许用应力过大，会使零部件单薄而损坏。因此，合理选取安全系数和许用应力是机械强度计算中一项基本工作。

不同材料及不同载荷下的静强度判据还可以由以下各式表达：

对于塑性材料　$[\sigma] = \sigma_s / [n]_s$　(6.2-38)

式中，$[n]_s$ 为以屈服极限为基准的许用安全系数。

对于脆性材料　$[\sigma] = \sigma_b / [n]_b$　(6.2-39)

式中，$[n]_b$ 为以强度极限为基准的许用安全系数。

对于轴类零件，要求挠度不能过大，即刚度计算应满足：

$$f \leqslant [f] \tag{6.2-40}$$

式中，f 为零件的计算挠度；$[f]$ 为许用挠度，对于轴，$[f] = (0.000\,1 \sim 0.000\,5)\ L$，$L$ 为轴的跨度。

对于受轴向载荷的柱、杆等零构件，要求工作满足稳定性要求，即

$$n_{sw} = [n]_{sw} \tag{6.2-41}$$

式中，n_{sw} 为零件的弹性失稳安全系数；$[n]_{sw}$ 为许用失稳安全系数。

7.2　静应力下的安全系数

在设计中采用的许用安全系数大多是由经验获得，但也可用部分系数法求得安全系数，即用一系列的部分系数的乘积来确定安全系数。这些部分系数的数值，分别表明零部件的用途、工作状况及制造工艺等系数。如以强度极限为基准的安全系数，可写成

$$[n]_b = K_1 K_2 K_3 K_4 K_5 K_6 K_7 \tag{6.2-42}$$

式中，K_1 为考虑零部件重要程度的系数，见表 6.2-3；K_2 为考虑计算载荷及应力公式的准确度系数，根据三种情况确定：①计算公式准确，所有作用力及应力已知时，则取 $K_2 = 1.0$；②计算公式或图表，使计算所得应力较实际应力高时，则 $K_2 = 1.0$；③若计算应力较实际应力低，根据二者之差异，可选取 $K_2 = 1.05 \sim 1.65$；K_3 为抗拉强度极限与其他失效形式的强度极限之间关系的系数，对于静载荷：塑性材料 $K_3 = \frac{抗拉强度}{屈服强度}$，脆性材料 $K_3 = \frac{抗拉强度}{所考虑的强度极限}$；对于循环变载荷：$K_3 = \frac{抗拉强度}{疲劳极限}$；$K_4$ 为考虑应力集中的系数

(用有效应力集中系数 K_σ)；K_5 为考虑截面尺寸增大的系数，$K_5=1/\varepsilon$，ε 为尺寸系数；K_6 为考虑零部件表面加工状况的系数，$K_6=1/\beta$，β 为表面系数；K_7 为检验质量系数，成批产品抽样试验，$K_7=1.15\sim1.30$；每个零部件都检验，$K_7=1.05\sim1.15$。

表 6.2-3 系数 K_1 的取值

零部件的重要程度	价值低的零部件	价值高的零部件
零部件的破坏不会引起停车	1.0	1.0
零部件的破坏会引起停车	1.1	1.2
零部件的破坏会造成事故	1.2	1.3

7.3 复杂应力状态下的强度理论

7.3.1 最大主应力理论

最大主应力理论认为：最大主应力是引起材料失效的主要因素。即认为无论材料处于何种应力状态，只要最大主应力 σ_1 达到并超过材料单向应力状态下破坏时的最大主应力，材料即发生断裂。按此理论，材料的破坏条件为

$$\sigma_1=\sigma_{\lim} \tag{6.2-43}$$

式中，$\sigma_{\lim}$可以是材料单向应力状态下的抗拉强度 σ_b 或屈服强度 σ_s 等。

以安全系数除极限应力取得许用应力 [σ]，于是按第一强度理论建立的强度破坏条件为

$$\sigma_1-\nu\ (\sigma_2+\sigma_3)\ \leqslant\ [\sigma] \tag{6.2-44}$$

试验表明：脆性材料在二向或三向受拉破坏时，最大主应力理论与试验结果相当接近；而当存在压应力的情况下，则只要最大压应力值不超过最大拉应力值或超过不多，最大主应力理论也是正确的。

7.3.2 最大正应变理论

最大正应变理论认为，最大伸长线应变是引起材料失效的主要因素。即认为无论材料处于何种应力状态，只要最大伸长线应变 ε_1 达到并超过材料单向应力状态下破坏时的最大主应变，材料即发生破坏。设脆性材料直到破坏可认为仍服从胡克定律，则单向拉伸的伸长线应变的极限值为 σ_b/E。按照这一理论，任意应力状态下，只要 ε_1 达到极限值，材料就发生破坏，故断裂准则为

$$\varepsilon_1=\frac{\sigma_b}{E} \tag{6.2-45}$$

由广义胡克定律

$$\varepsilon_1=\frac{1}{E}\ [\sigma_1-\nu\ (\sigma_2+\sigma_3)]$$

代入式 (6.2-45)，得破坏准则为

$$\sigma_1-\nu\ (\sigma_2+\sigma_3)\ =\sigma_b \tag{6.2-46}$$

以安全系数除极限应力取得许用应力 [σ]，于是按第二强度理论建立的强度破坏条件为

$$\sigma_1-\nu\ (\sigma_2+\sigma_3)\ \leqslant\ [\sigma] \tag{6.2-47}$$

石料或混凝土等脆性材料的试块受轴向压缩时，如在试验机与试块的接触面上添加润滑剂以减小摩擦力，试块将沿垂直于压力的方向裂开，这正是 ε_1 的方向。铸铁受拉—压二向应力且压力较大时，试验结果也与这一理论接近。不过，如在试块受压的垂直方向再加压力，使它成为二向受压，按照这一理论，其强度应与单向受压不同。但混凝土、花岗石和沙岩的试验资料表明，两种情况的强度并无明显差别。与此相似，按照这一理论，铸铁在二向拉伸时应比单向拉伸安全，但试验结果并不能证实这一点。对这种情况，还是第一强度理论更接近试验结果。

7.3.3 最大切应力理论

最大切应力理论认为，最大切应力是引起失效的主要因素。即认为在复杂应力状态下，只要最大切应力达到并超过单向应力状态下破坏时的最大切应力水平，材料即发生破坏。

在三向应力状态下，只考虑量值，不考虑方向时，最大切应力可以按下式写出

$$\left.\begin{aligned}|\tau_{12}|&=\left|\frac{1}{2}\ (\sigma_1-\sigma_2)\right|\\|\tau_{23}|&=\left|\frac{1}{2}\ (\sigma_2-\sigma_3)\right|\\|\tau_{31}|&=\left|\frac{1}{2}\ (\sigma_3-\sigma_1)\right|\end{aligned}\right\} \tag{6.2-48}$$

在单向应力条件下达到破坏的最大切应力值为

$$|\tau_{\max}|=|\frac{1}{2}\sigma_{\lim}| \tag{6.2-49}$$

于是破坏条件即为

$$|\tau_{31}|\geqslant|\tau_{\max}|$$

或写成

$$|\sigma_3-\sigma_1|\geqslant2|\tau_{\max}|=\sigma_{\lim} \tag{6.2-50}$$

一般，把$|\sigma_3-\sigma_1|$作为按照最大切应力理论求得的与单向应力等效的正应力，记为 σ_{eq3}，于是

$$\sigma_{eq3}=|\sigma_3-\sigma_1| \tag{6.2-51}$$

试验结果表明，用最大切应力理论来预测塑性材料发生屈服的现象能够取得与试验比较一致的结果。因此，本理论常常用来对塑性材料进行以屈服极限 σ_s 作为极限应力的强度计算，并引入安全系数取得许用应力 [σ]，则按照本理论，不发生屈服的强度条件如下

$$|\sigma_1-\sigma_3|\leqslant\ [\sigma] \tag{6.2-52}$$

对于双向应力状态，取 $\sigma_3=0$。由材料力学已知主应力 σ_1、σ_2 和一般应力 σ_x、σ_y 与 τ_{xy}之间的关系为

$$\sigma_{1,2}=\frac{\sigma_x+\sigma_y}{2}\pm\sqrt{\left(\frac{\sigma_x-\sigma_y}{2}\right)^2+\tau_{xy}^2} \tag{6.2-53}$$

代入式 (6.2-48) 第一式，得双向应力状态时用的由一般的应力分量表达的强度极限条件为

$$2\sqrt{\left(\frac{\sigma_x-\sigma_y}{2}\right)^2+\tau_{xy}^2}\leqslant\ [\sigma]$$

即不发生屈服的条件则为

$$\sqrt{(\sigma_x-\sigma_y)^2+4\tau_{xy}^2}\leqslant\ [\sigma] \tag{6.2-54}$$

对于三向应力状态，此理论没有考虑第三个主应力的影响。另外，当材料的拉伸与压缩的机械性能有差异时，用此理论会导致较大的误差。

最大切应力理论比较圆满地解释了屈服现象。例如，低碳钢拉伸时沿与轴线成45°的方向出现滑移线。这是材料内部沿这一方向相对滑移的痕迹，而沿这个方向的斜截面上切应力也恰为最大值。钢、铝、铜等材料的实验资料表明，塑性变形出现时最大切应力接近于某一常量，这正是屈服准则式 (6.2-52) 表示的关系。这一理论的缺陷是忽略了 σ_2 的影响。在二向应力状态下，与实验资料比较，理论结果偏于安全。

7.3.4 形变能理论

形变能理论认为，形变能是引起失效的主要因素。即认为在复杂应力状态下，只要形变能超过单向应力状态下极限条件的形变能水平，材料即发生破坏。

三向应力状态时的形变能以下式表示

$$U_F=\frac{1+\nu}{6E}[(\sigma_1-\sigma_2)^2+\ (\sigma_2-\sigma_3)^2+\ (\sigma_3-\sigma_1)^2] \tag{6.2-55}$$

单向应力状态时的形变能为

$$U_F = \frac{1+\nu}{3E}\sigma^2 \tag{6.2-56}$$

式（6.2-56）中 $\sigma = \sigma_{lim}$。在工程设计计算中，形变能理论常常用于塑性材料的场合，故常取 $\sigma_{lim} = \sigma_s$。引入安全系数取得许用应力［σ］，则按照本理论，得到形变能理论的当量应力 σ_{eq4} 及不发生屈服的条件如下

$$\sigma_{eq4} = \frac{\sqrt{2}}{2}[(\sigma_{xx}-\sigma_{yy})^2+(\sigma_{yy}-\sigma_{zz})^2+(\sigma_{zz}-\sigma_{xx})^2+6(\tau_{xy}^2+\tau_{yz}^2+\tau_{xz}^2)]^{1/2}$$
$$= \frac{\sqrt{2}}{2}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2]^{1/2} \leqslant [\sigma] \tag{6.2-57}$$

三向应力状态时的八面体切应力为

$$\tau_{oc} = \frac{1}{3}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2]^{1/2} \tag{6.2-58}$$

在单向应力时，$\sigma_2 = \sigma_3 = 0$，并令 $\sigma_1 = \sigma_s$，则有：

$$\tau_{oc} = \frac{\sqrt{2}}{3}\sigma_s \tag{6.2-59}$$

把式（6.2-59）与式（6.2-58）相比较，也可求得式（6.2-57）。因此，形变能理论也叫做八面体切应力理论。

几种塑性材料钢、铝、铜的试验资料与屈服准则式（6.2-57）非常接近。这一理论与实验结果吻合的程度比第三理论更好。在纯剪切的情况下，由屈服准则式（6.2-57）和式（6.2-52）得出的结果相差15%，这是两者相差最大的情况。由于绝大多数工程设计计算均采用了屈服准则，因此第三和第四强度理论在工程设计中是应用最广的。从应用实践来看，形变能理论稍偏于危险，而最大切应力理论总是稍偏于安全。从工程应用的观点来看，它们都可以认为是与试验结果相符合的。因此这两个理论至今为止一直在并行使用着。

编写：何雪浤（东北大学）
谢里阳（东北大学）

第3章　材料的断裂韧度设计

1　材料断裂韧度设计概要

断裂韧度设计是通过对临界裂纹尺寸（或临界剩余强度）、裂纹扩展特性和损伤检测这三个要素的合理组合，保证结构或零部件的使用安全。裂纹长度、剩余强度与检查周期三要素之间的关系如图 6.3-1 所示。

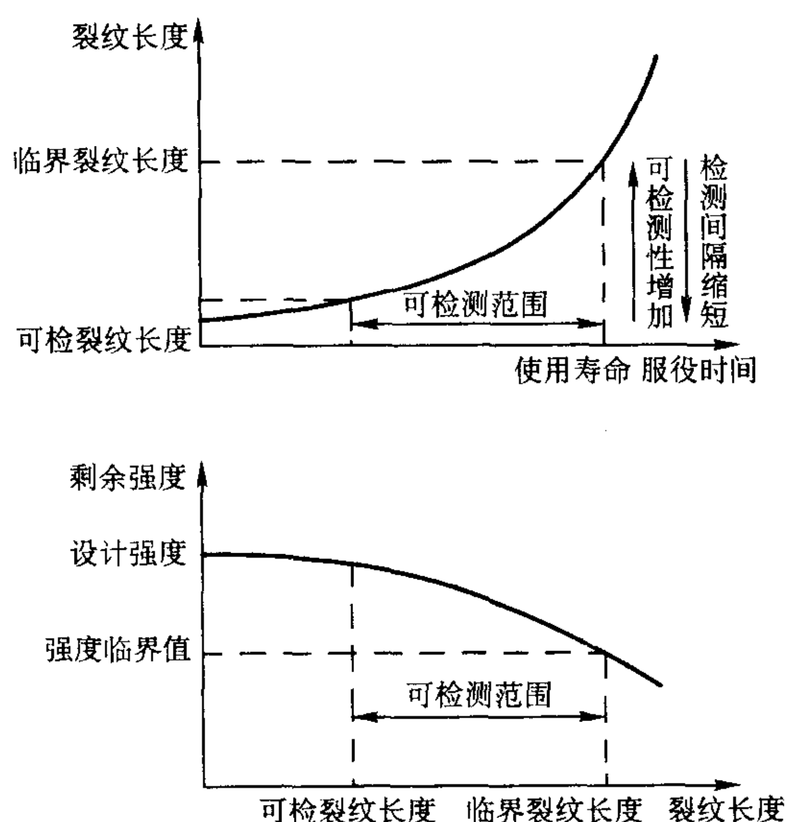


图 6.3-1　裂纹长度、剩余强度及检查周期关系图

采用断裂韧度设计（或称损伤容限设计）准则设计的结构或零件称为损伤容限结构。损伤容限结构对其基本要素有如下要求。

1）结构或零件必须有足够的使用寿命。保证寿命的具体措施包括选用低裂纹扩展速率、采用止裂结构或多途径传力结构。这些措施应保证构件在检修前能够有效地在正常载荷下工作。

2）工作环境中产生的裂纹应能比较容易地在使用中或检修时用简单的方法检测出来。

3）已破损的结构部位或零件可以方便、经济地更换或修复。

4）对于复杂结构，采用多条传力路线。当一个传力路线上的零件破损或失效时，载荷可以分配到其他未受损伤的传力路线上去，并保证结构功能正常。

断裂韧度设计的注意事项。

1）尽量将结构设计成破损安全结构，以便某些部位产生了裂纹后能继续承受载荷。同时，要能通过定期检测发现这些裂纹。要达到此目的，可采用多传力途径的结构布局、静不定结构形式或组合承力结构等，并且使结构具有缓慢裂纹扩展特性。

2）保证结构破损安全的关键是定期地对结构进行检查或考验性试验。因此，对于易产生裂纹的重要构件，要尽量设计成可检结构，有开敞的检查通路，以便日常维护检查、修理和更换。

3）合理地确定检查周期是保证结构破损安全的关键，要确定检查周期必须确定最小可检裂纹尺寸和破损安全载荷下的临界裂纹尺寸。这就要求比较准确地计算对应于每一个裂纹的剩余强度和裂纹扩展速率，从而使裂纹扩展寿命计算更符合实际情况。

4）采用断裂韧度好、抗裂纹扩展能力强的材料，以保证结构具有较高的剩余强度和缓慢裂纹扩展特性。

5）最小可检裂纹尺寸取决于检测手段和检测人员的技术水平，提高检测灵敏度对提高结构的安全性有非常重要的作用。

6）对于尺寸较大的构件应考虑止裂措施，如采用止裂孔、止裂带和结构分段等，以防止裂纹快速扩展。

7）合理地控制结构的设计应力水平。应当综合强度、刚度、断裂韧度、耐久性和可靠性方面的要求，合理确定应力水平和设计指标。

断裂韧度设计需要一系列的分析、计算和试验。在断裂韧度设计中，特别需要注意高强度材料的承力结构、承受多次重复载荷的构件、应力集中程度较高的部位、反复受高温或剧烈振动的部位，以及环境条件恶劣的部位等。对重要的危险部位必须在加工和使用中提出特殊的工艺及检验要求，同时要有相应措施保证危险构件的可跟踪性。

断裂韧度设计的基本内容可概括如下。

1）确定载荷谱。载荷谱是断裂韧度设计的关键，直接关系到结构的安全可靠性。断裂韧度设计中所采用的载荷谱应考虑载荷顺序对裂纹扩展速率的影响。

2）根据使用经验、结构特点以及结构应力分析结果，确定主要结构中需要进行断裂控制的零件及部位，并按可检查程度对结构进行分类，然后按相应的结构类型进行结构设计。

3）兼顾静强度、刚度和疲劳设计的要求，选择抗断裂性能好的材料。

4）确定初始裂纹（缺陷）尺寸。断裂韧度设计中作为裂纹扩展计算起点的初始裂纹尺寸 a_0 应是新构件开始使用时或经检修后再次投入使用时，危险部位可能存在的最大裂纹尺寸。初始裂纹长度在很大程度上影响剩余寿命的长短。精确地估计初始裂纹尺寸目前还是一个比较困难的问题。各种无损检测技术所能发现的最小裂纹尺寸不同。同一种检测手段检测同一尺寸的裂纹，其检出概率也受许多因素的影响。与结构的几何尺寸和材料、检测操作的环境、裂纹位置、方向和尺寸都有关系。因此，需要针对所设计部位给出一定检测概率和一定置信度的初始裂纹尺寸。关键零件、关键部位如未检测出裂纹，为了简单起见，可把 a_0 取为仪器检测精度的最小值。

5）确定材料的断裂韧度和裂纹扩展计算所需要的材料常数。

6）应用断裂力学基本理论确定结构在各种裂纹尺寸下的剩余强度，根据载荷确定临界裂纹尺寸。

7）计算从初始裂纹尺寸扩展到临界裂纹尺寸所需的时间，按照裂纹扩展寿命确定检修周期。

8）对不可检零件，如果裂纹扩展到临界裂纹长度的寿命低于构件的设计寿命，则须调整设计参数以便使其寿命满足设计要求。如果还达不到这个要求，就得规定部件更换周期。

2　断裂力学原理与方法

2.1　裂纹的基本类型

从材料或构件存在的裂纹中，可以抽象出 I 型裂纹、II

型裂纹和Ⅲ型裂纹这三种基本裂纹形式。其他复杂的裂纹形式可以看作是这三种基本裂纹形式的组合。

Ⅰ型裂纹为张开型裂纹，如图 6.3-2 所示。Ⅰ型裂纹的特征是外载荷为垂直于裂纹平面的正应力。在外力作用下，裂纹顶端张开，裂纹扩展方向与外力的方向垂直。这种裂纹形式在工程构件中很常见，也是三种基本裂纹形式中最危险的一种。拉板上垂直于拉力方向的贯穿裂纹、长圆筒形压力容器或受内压的管道壁上的纵向裂纹、飞轮上的径向裂纹（在环向力作用下裂纹将张开并沿径向扩展）等都属于Ⅰ型裂纹。

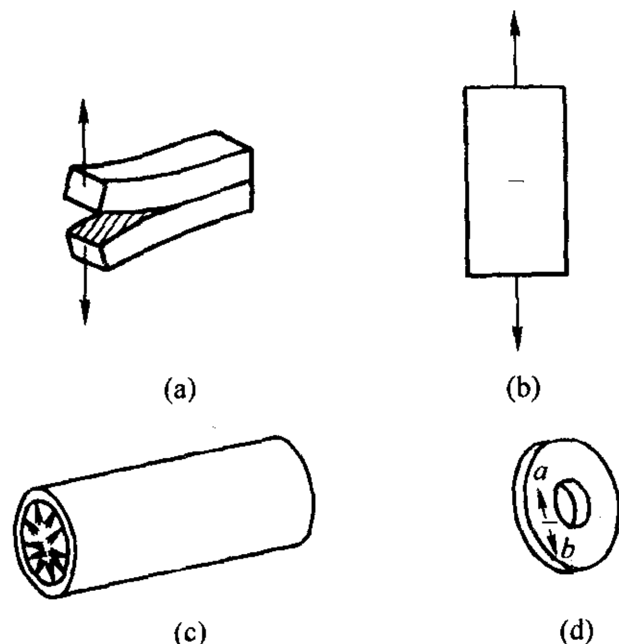


图 6.3-2　Ⅰ型裂纹

Ⅱ型裂纹为滑移型裂纹，如图 6.3-3 所示。Ⅱ型裂纹的特征是外载荷是裂纹面内垂直于裂纹前缘的剪力，裂纹在其自身平面内作垂直于裂纹前缘的滑动。轮齿或花键根部沿切线方向的裂纹、受扭转作用的薄壁圆管上环向贯穿裂纹、受剪力作用的铆钉的剪切面内的裂纹等均属Ⅱ型裂纹。

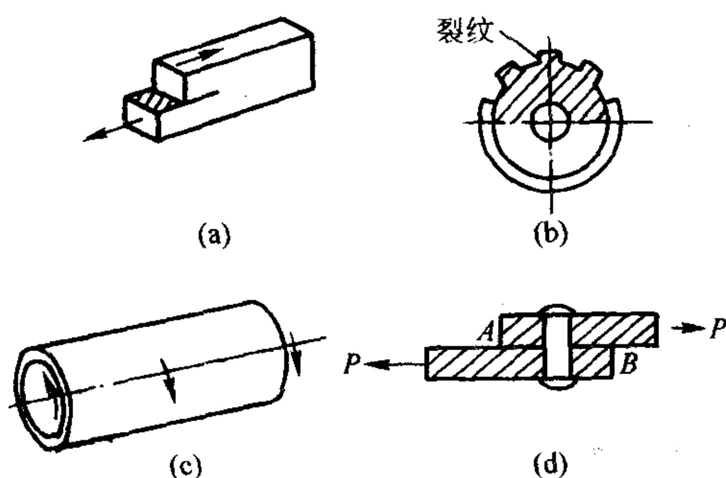


图 6.3-3　Ⅱ型裂纹

Ⅲ型裂纹为撕裂型裂纹，如图 6.3-4 所示。Ⅲ型裂纹的特征是外载荷为平行于裂纹前缘的剪力。受扭转作用圆轴上的环形切槽或表面环型裂纹即为Ⅲ型裂纹。

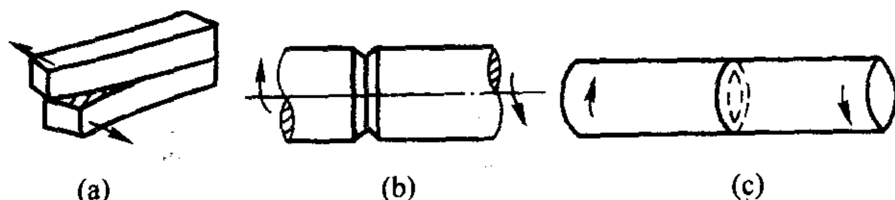


图 6.3-4　Ⅲ型裂纹

在工程实际中，还有各种各样的复合型裂纹。例如，构件内的裂纹同时受到正应力和剪应力的作用，可能同时有张开和滑移或张开和撕裂形式的裂纹扩展，这样的裂纹可以看作是两种或两种以上基本裂纹形式的组合，称为复合型裂纹。

2.2　应力强度因子与断裂判据

线弹性断裂力学可用于分析材料或构件在低应力下的裂纹扩展及脆断问题，也可用于近似地分析小范围屈服（塑性区尺寸远小于裂纹尺寸）情况下的裂纹扩展等问题。在断裂韧度设计中，涉及应力强度因子 K、材料断裂韧度 K_c、断裂判据等概念。

2.2.1　裂纹尖端应力状态与应力强度因子

用断裂力学方法研究裂纹扩展的规律，首先需要了解裂纹顶端的应力状态，求出与应力 σ 和裂纹长度 $2a$ 有关的、称为“应力强度因子 K”的应力场强度特征参数，测定材料的断裂韧度 K_c，然后建立形式为 $K \geqslant K_c$ 的裂纹失稳判据。Ⅰ型裂纹是最危险的裂纹形式，下面以Ⅰ型裂纹为例介绍应力强度因子的概念。

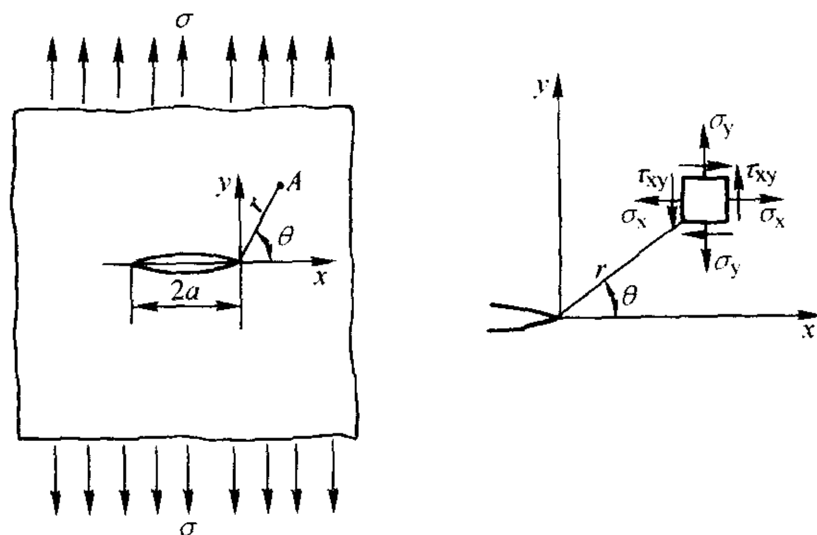


图 6.3-5　Ⅰ型裂纹尖端附近的应力场

图 6.3-5 所示的无限大板含有长度为 $2a$ 中心贯穿裂纹，边界上受有垂直于裂纹的均匀拉应力，这样的裂纹定义为Ⅰ型（张开型）裂纹。裂纹尖端附近的应力状态可以借助弹性理论确定。裂纹尖端附近区域任一点（用极坐标 r，θ 表示）的三个应力分量分别为：

$$\left.\begin{aligned}\sigma_x &= \sigma\sqrt{\pi a}\left[\frac{1}{\sqrt{2\pi r}}\cos\frac{\theta}{2}\left(1-\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right)\right]\\ \sigma_y &= \sigma\sqrt{\pi a}\left[\frac{1}{\sqrt{2\pi r}}\cos\frac{\theta}{2}\left(1+\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right)\right]\\ \tau_{xy} &= \sigma\sqrt{\pi a}\left(\frac{1}{\sqrt{2\pi r}}\cos\frac{\theta}{2}\sin\frac{\theta}{2}\cos\frac{3\theta}{2}\right)\end{aligned}\right\}\qquad(6.3\text{-}1)$$

由此可见，根据线弹性理论，裂纹尖端存在应力奇异性，即裂纹尖端（$r\to 0$）的应力趋于无穷大（与实际情况不符）。裂纹尖端的应力奇异性说明，在有裂纹存在的情况下，传统的应力强度准则不再适用，不能用应力值的大小来衡量含裂纹材料或构件的安全余量。

由上面的应力分量表达式可以看出，各应力分量均含有一个共同的与位置无关的系数 $\sigma\sqrt{\pi a}$，它是一个有限量，仅取决于载荷水平与裂纹尺寸。除 $\sigma\sqrt{\pi a}$ 项之外，各应力分量中的其他各项（于中括号内）都是位置参量。由此可见，这个与位置无关的参量 $\sigma\sqrt{\pi a}$ 是表征裂纹尖端附近区域应力场强度的一个因子，称为应力强度因子。应力强度因子的量纲为 $\mathrm{MPa\cdot m^{1/2}}$。Ⅰ型裂纹的应力强度因子通常记为 K_{I}，其应力强度因子表达式为

$$K_{\mathrm{I}} = \sigma\sqrt{\pi a}\qquad(6.3\text{-}2)$$

相应地，对于Ⅱ型裂纹，即在无穷远处受板面内剪应力 τ 作用的无限大板上的中心穿透裂纹，有

$$K_{\mathrm{II}} = \tau\sqrt{\pi a}\qquad(6.3\text{-}3)$$

对于Ⅲ型裂纹，即在无穷远处受垂直于板面的剪应力 τ_1 作用的无限大板中的中心穿透裂纹，有

$$K_{\mathrm{III}}=\tau_{\mathrm{I}}\sqrt{\pi a} \qquad (6.3\text{-}4)$$

还有一点须说明的是，对于不同的几何形状及载荷情况，应力强度因子的表达式也有所不同。上面的 K_{I} 表达式是含有穿透裂纹的无限大平板在Ⅰ型加载情况下的应力强度因子。对于Ⅰ型加载的一般情况（例如有限宽板），应力强度因子通式为

$$K_{\mathrm{I}}=F\sigma\sqrt{\pi a} \qquad (6.3\text{-}5)$$

式中，F 为几何形状因子，是一个量纲为 1 的系数。它取决于含裂纹体的形状、尺寸，裂纹的形状、尺寸及其在裂纹体中的部位以及载荷形式等因素。

2.2.2 断裂判据

1）基本判据 含裂纹体的承载能力可以用应力强度因子来衡量，当应力强度因子达到某一临界值时裂纹将失稳扩展导致裂纹体断裂。在一定条件下，应力强度因子的这一临界值是材料常数，其大小反映了材料抗断裂的能力，称为材料的断裂韧度，通常记为 K_{IC}。断裂韧度的量纲与应力强度因子相同，也是 $\mathrm{MPa\cdot m^{1/2}}$。由此，断裂判据可表达为

$$K_{\mathrm{I}}\geqslant K_{\mathrm{IC}} \qquad (6.3\text{-}6)$$

即当裂纹尖端附近的应力强度因子达到材料的断裂韧度时就发生断裂。材料的断裂韧度与应力状态和温度等因素有关。断裂韧度还随试件厚度的增加而下降，这是因为薄试件的裂纹尖端处于平面应力状态，裂纹尖端塑性区较大，裂纹扩展需要消耗的能量较多，所以裂纹不易扩展，表现为断裂韧度较高；厚板中裂纹尖端处于平面应变状态，裂纹尖端塑性区较小，裂纹容易扩展，因而表现为断裂韧度较低。

除非特别说明，K_{Ic} 表示平面应变状态下的常温断裂韧度。大多数金属材料的断裂韧度都随着温度的升高而增加，其中存在一个很窄的温度范围，在该温度范围内断裂韧度的变化几乎是跳跃式的。材料断裂韧度发生跳跃的温度称为脆性转变温度，在设计中应予以注意。

对Ⅱ、Ⅲ型裂纹，也有与Ⅰ型裂纹类似的断裂判据。在实际工程结构中，裂纹多属于复合型，即Ⅰ、Ⅱ、Ⅲ三种载荷形式中的两种以上同时存在。在复合型加载的情况下，裂纹扩展方向可能会与原裂纹线方向偏离一个角度 θ_0。这个角度称为开裂角，如图 6.3-6 所示。对复合型裂纹需应用复合型断裂判据，考虑裂纹沿什么方向开裂、裂纹在什么条件下开裂这两方面的问题。目前关于复合型断裂的理论有最大周向应力理论、应变能密度因子理论等。纯Ⅱ型和纯Ⅲ型裂纹的断裂判据可以看做是复合型裂纹的特例。

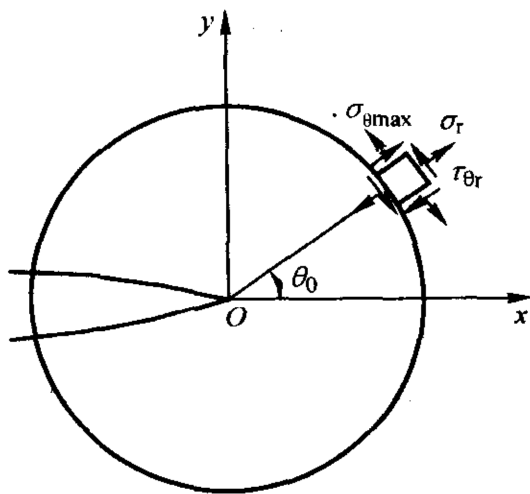


图 6.3-6 最大周向应力与开裂角

2）最大周向应力判据（σ_θ 准则） 最大周向应力判据适用于Ⅰ－Ⅱ型复合裂纹，其基本假设是裂纹沿最大周向应力 $\sigma_{\theta\max}$ 方向开裂，且当此方向的周向应力达到临界值时裂纹失稳。最大周向应力方向与原裂纹方向的夹角即为开裂角 θ_0。开裂角可由下式确定

$$K_{\mathrm{I}}\sin\theta_0+K_{\mathrm{II}}(3\cos\theta_0-1)=0 \qquad (6.3\text{-}7)$$

根据这一判据，裂纹沿着 θ_0 方向扩展，因此这个方向就是新的裂纹方向，应力 $\sigma_\theta(r,\theta_0)$ 与新裂纹线方向垂直，它与Ⅰ型加载下的应力 $\sigma_y(r,0)$ 相当。因此，可以参照Ⅰ型加载下的应力强度因子 K_{I} 定义一个相应的当量应力强度因子 K_e：

$$K_e=\cos\frac{\theta_0}{2}\left(K_{\mathrm{I}}\cos^2\frac{\theta_0}{2}-\frac{3}{2}K_{\mathrm{II}}\sin\theta_0\right) \qquad (6.3\text{-}8)$$

相应地，断裂判据为

$$K_e\geqslant K_{\mathrm{Ic}} \qquad (6.3\text{-}9)$$

这样，复合型裂纹问题就转化为当量Ⅰ型裂纹问题。

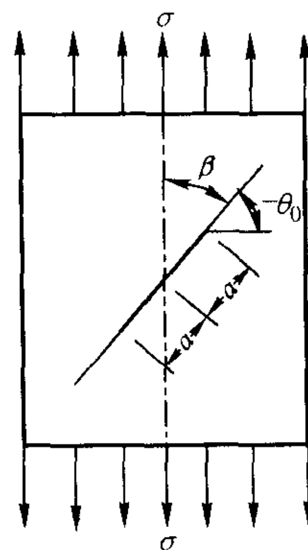


图 6.3-7 Ⅰ－Ⅱ复合裂纹

考虑一受单向拉伸载荷作用的大平板中含有一条斜裂纹的情况（见图 6.3-7）。裂纹长度为 $2a$，裂纹与应力轴之间的夹角为 β。Ⅰ型和Ⅱ型裂纹的应力强度因子分别为

$$K_{\mathrm{I}}=\sigma\sqrt{\pi a}\sin^2\beta$$

$$K_{\mathrm{II}}=\sigma\sqrt{\pi a}\sin\beta\cos\beta$$

将上面的 K_{I}、K_{II} 代入式（6.3-7），可得出开裂角 θ_0 与 β 之间的关系：

$$\sin\theta_0+(3\cos\theta_0-1)\cot\beta=0$$

给定 β，即可由上式求得开裂角 θ_0。须注意的是，当 $0<\beta<\pi/2$ 时，$\theta_0<0$。将上面的应力强度因子表达式代入式（6.3-8），再由断裂判据（6.3-9）即可求出临界断裂应力为

$$\sigma_c=\frac{2K_{\mathrm{IC}}}{\sqrt{\pi a}\cos\frac{\theta_0}{2}\left[(1+\cos\theta_0)\sin^2\beta-3\sin\theta_0\sin\alpha\cos\beta\right]}$$

求出了 σ_c 之后，即可分别算出裂纹失稳时对应的 K_{I} 和 K_{II} 值。对于纯Ⅱ型裂纹，由 $\beta=0$ 可知开裂角 $\theta_0=-70°32'$。此外还可得到以下关系：

$$K_{\mathrm{IIC}}=\frac{\sqrt{3}}{2}K_{\mathrm{IC}}$$

下面给出的是一个应用最大周向应力判据（σ_θ 准则）的具体例子。

例 假设在压力容器筒体上有一长度为 $2a$ 的穿透型斜裂纹，裂纹与筒体轴线的倾角为 $\beta=45°$（如图 6.3-8 所示）。筒体壁厚为 t，半径为 R，要求计算此容器的开裂压力 p_c。

解：根据材料力学的应力状态理论，可知垂直于裂纹线方向的正应力为

$$\sigma_\alpha=\frac{1}{2}(\sigma_1+\sigma_2)-\frac{1}{2}(\sigma_1+\sigma_2)\cos2\beta$$

$$=\frac{pR}{4t}(3-\cos2\beta)$$

相应的切应力为

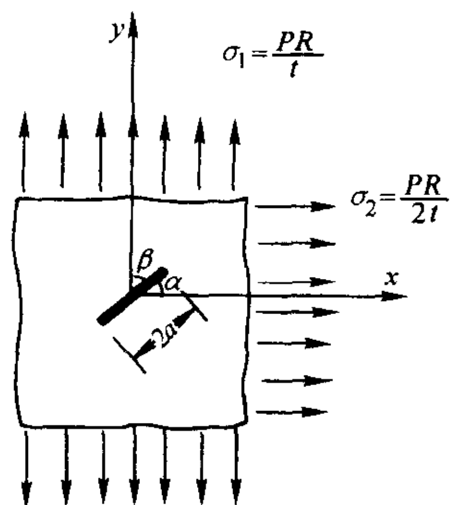


图 6.3-8　压力容器筒壁上的斜裂纹

$$\tau_a=\frac{1}{2}\ (\sigma_1-\sigma_2)\ \sin2\beta=\frac{pR}{4t}\sin2\beta$$

以 $\beta=45°$代入上式，得

$$\sigma_{45}=3\left(\frac{pR}{4t}\right)$$

$$\tau_{45}=\frac{pR}{4t}$$

因此有

$$K_1=\sigma\sqrt{\pi a}=3\left(\frac{pR}{4t}\right)\sqrt{\pi a}$$

$$K_{\rm II}=\tau\sqrt{\pi a}=\frac{pR}{4t}\sqrt{\pi a}$$

把 $K_{\rm I}$、$K_{\rm II}$代入式（6.3-7）可解出 $\theta_0\approx31°22'$。这表明，根据 σ_θ 准则，裂纹将沿着与原裂纹线成 31°22′角的方向扩展。将此 θ_0 值及 $K_{\rm I}$、$K_{\rm II}$值代入式（6.3-8），可得

$$K_e=\cos\frac{\theta_0}{2}\left(K_{\rm I}\cos^2\frac{\theta_0}{2}-\frac{3}{2}K_{\rm II}\sin\theta_0\right)$$
$$=0.85\frac{pR}{t}\sqrt{\pi a}$$

再根据断裂判据 $K_e\geqslant K_{\rm IC}$，即可求得此容器的开裂压力为

$$p_c=1.166\frac{t}{R}\frac{K_{\rm IC}}{\sqrt{\pi a}}$$

3）应变能密度因子准则（S 准则）　对于Ⅰ、Ⅱ、Ⅲ型复合加载下的裂纹问题，由弹性力学可知弹性体内的应变能密度为

$$W=\frac{1}{r}\ (a_{11}k_1^2+2a_{12}k_1k_2+a_{22}k_2^2+a_{33}k_3^2)\quad(6.3\text{-}10)$$

其中

$$\left.\begin{aligned}&k_1=\frac{K_{\rm I}}{\sqrt{\pi}},\ k_2=\frac{K_{\rm II}}{\sqrt{\pi}},\ k_3=\frac{K_{\rm III}}{\sqrt{\pi}}\\&a_{11}=\frac{1}{16G}\ [\ (1+\cos\theta)\ (k-\cos\theta)]\\&a_{12}=\frac{1}{16G}\sin\theta\ [2\cos\theta-\ (k-1)]\\&a_{22}=\frac{1}{16G}\ [\ (k+1)\ (1-\cos\theta)\ +\\&\qquad(1+\cos\theta)\ (3\cos\theta-1)]\\&a_{33}=\frac{1}{4G}\end{aligned}\right\}\quad(6.3\text{-}11)$$

式中，G 为材料的切变模量。在平面应力状态下，$k=\frac{3-\nu}{1+\nu}$；在平面应变状态下，$k=3-4\nu$（ν 为材料的泊松比）。令

$$S=a_{11}k_1^2+2a_{12}k_1k_2+a_{22}k_2^2+a_{33}k_3^2\qquad(6.3\text{-}12)$$

则 $W=S/r$。参量 S 是衡量裂纹尖端附近应变能密度强弱程度的量，称为应变能密度因子。应变能密度因子理论认为，裂纹沿应变能密度因子极小值的方向扩展，开裂角 θ_0 满足以下条件：

$$\left(\frac{\partial S}{\partial\theta}\right)_{\theta=\theta_0}=0;\ \left(\frac{\partial^2S^2}{\partial\theta}\right)_{\theta=\theta_0}>0$$

当应变能密度因子的极小值 $S_{\min}$ 达到某一临界值 S_c 时裂纹开始扩展，断裂判据为

$$S_{\min}\geqslant S_c\qquad(6.3\text{-}13)$$

上式中的 S_c 也是一个材料韧性指标，并可由 $K_{\rm IC}$换算

$$S_c=\frac{k-1}{8\pi G}K_{\rm IC}^2\qquad(6.3\text{-}14)$$

由 S 准则可以很方便地得出Ⅱ型断裂判据

$$\frac{14k-1-k^2}{192G}k_2^2\geqslant S_c\qquad(6.3\text{-}15)$$

或

$$K_{\rm II}\geqslant\sqrt{\frac{24\ (k-1)}{14\kappa-1-k^2}}K_{\rm IC}\qquad(6.3\text{-}16)$$

Ⅲ型加载断裂判据为

$$K_{\rm I}\geqslant\sqrt{1-2\nu}K_{\rm IC}\qquad(6.3\text{-}17)$$

对于平面应变状态下的Ⅰ、Ⅲ型复合加载，裂纹仍旧沿着原裂纹线方向扩展。断裂判据为

$$\sqrt{K_{\rm I}^2+\frac{K_{\rm III}^2}{1-2\nu}}\geqslant K_{\rm IC}\qquad(6.3\text{-}18)$$

下面，用 S 准则求解上例压力容器的开裂压力（按平面应变状态考虑，并取 $\nu=0.3$）。

由

$$k_{\rm I}=\frac{K_{\rm I}}{\sqrt{\pi}}=\sigma\sqrt{a}=3\left(\frac{pR}{4t}\right)\sqrt{a}$$

$$k_2=\frac{K_{\rm II}}{\sqrt{\pi}}=\tau\sqrt{a}=\frac{pR}{4t}\sqrt{a}$$

$$k_3=0$$

可知

$$S=\frac{1}{16G}\left(\frac{pR}{4t}\right)^2a\ (18+6.4\cos\theta_0-$$
$$6\cos^2\theta_0+6\sin2\theta_0-4.8\sin\theta_0)$$

根据开裂条件 $\left(\frac{\partial S}{\partial\theta}\right)_{\theta=\theta_0}=0$，$\left(\frac{\partial^2S^2}{\partial\theta}\right)_{\theta=\theta_0}>0$，经计算可求得开裂角为 $\theta_0\approx-29°30'$。将此 θ_0 值代回上式，即可得

$$S_{\min}=\frac{a}{16G}\left(\frac{pR}{4t}\right)^2\times16.25$$

于是由断裂判据 $S_{\min}\geqslant S_c$即可求得容器的开裂压力为

$$p_c=1.255\frac{t}{R}\frac{K_{\rm IC}}{\sqrt{\pi a}}$$

σ_θ 准则物理意义比较明确，并得到了若干实验支持。但对于单轴拉伸的斜裂纹来说，当裂纹倾角 α 较小，即接近于Ⅱ型加载状态时，理论计算的开裂角与实验值有明显的差别。S 准则的物理意义不很明确，但也得到了若干实验支持。

4）工程判据　对于工程应用来说，可以应用简化的、偏于安全的复合型断裂判据——复合型断裂的工程判据。

Ⅰ、Ⅱ型复合加载的工程断裂判据为

$$K_{\rm I}+K_{\rm II}\geqslant K_{\rm IC}\qquad(6.3\text{-}19)$$

Ⅰ、Ⅲ型复合加载的工程断裂判据为

$$\sqrt{K_{\rm I}^{\rm II}+\frac{K_{\rm III}^2}{1-2\nu}}\geqslant K_{\rm IC}\qquad(6.3\text{-}20)$$

Ⅰ、Ⅱ、Ⅲ型复合加载的工程断裂判据为

$$\sqrt{(K_{\rm I}+K_{\rm II})^2+\frac{K_{\rm III}^2}{1-2\nu}}\geqslant K_{\rm IC}\qquad(6.3\text{-}21)$$

2.2.3 小范围屈服修正

上面关于裂纹尖端附近应力场及应力强度因子的分析都是在完全弹性的假设条件下进行的。对于大多数金属材料，由于应力高度集中，在裂纹尖端附近的材料将发生屈服，形成塑性区。如果裂纹尖端附近的塑性区比裂纹长度及裂纹体其他几何特征尺寸小得多，称为小范围屈服。在大多数情况下，当外载荷低于全塑性屈服载荷（对弹性—理想塑性材料为极限载荷）之半时即可视为小范围屈服。分析表明，对于Ⅰ型裂纹，塑性区形状如图 6.3-9 所示，塑性区尺寸为

$$r_p = \frac{1}{3\pi}\left(\frac{K_I}{\sigma_s}\right)^2 \quad （平面应变） \qquad (6.3\text{-}22)$$

$$r_p = \frac{1}{\pi}\left(\frac{K_I}{\sigma_s}\right)^2 \quad （平面应力） \qquad (6.3\text{-}23)$$

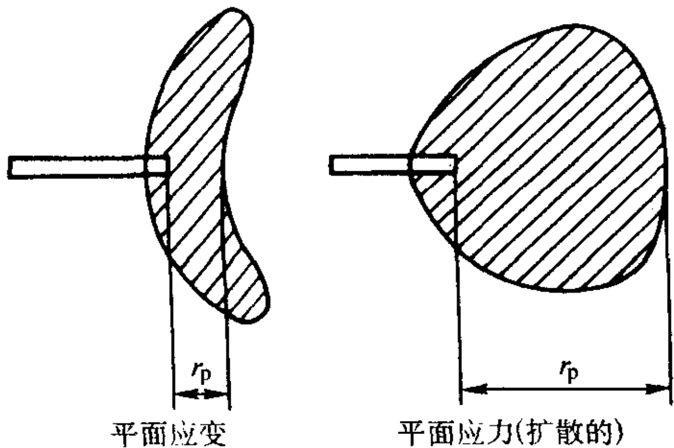


图 6.3-9　裂纹尖端塑性区

在小范围屈服的情况下，应力强度因子 K 仍可作为裂纹尖端附近弹性应力场的参数，但这时须对应力强度因子 K 进行修正。当裂纹尖端应力场为完全弹性时，裂纹截面的法向应力 σ_y 是按照图 6.3-10 中曲线 AB 分布的。由于裂纹尖端出现了尺寸为 r_p 的塑性区，σ_y 按照曲线 $CDEF$ 分布。如果假想裂纹尖端向前移至塑性区中心的 O_1 点，这时假想裂纹的 σ_y 弹性分布曲线 EF 与原裂纹有塑性区时的σ_y 弹性部分的分布曲线 EF 重合。因此，这个假想裂纹的 K 反映了原裂纹的应力场。取假想的裂纹长度 $a_e = a + \frac{1}{2} r_p$ 作为等效裂纹长度，小范围屈服情况下的应力强度因子可按等效裂纹长度计算，即

$$K_I = F\sigma\sqrt{\pi\left(a + \tfrac{1}{2} r_p\right)} \qquad (6.3\text{-}24)$$

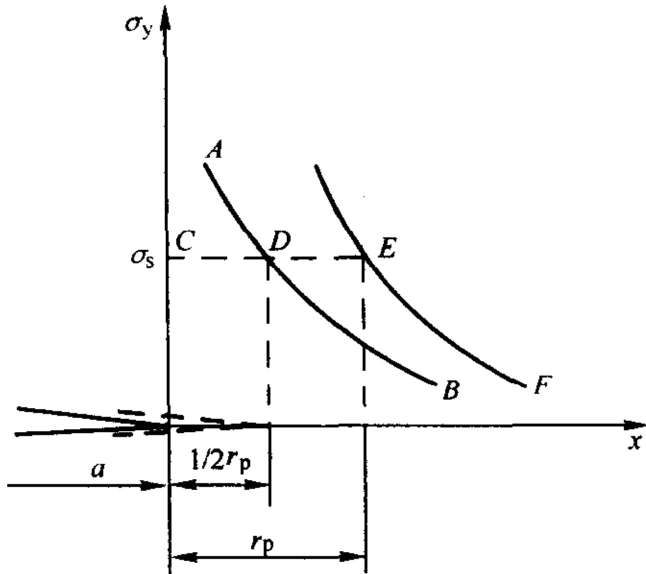


图 6.3-10　裂纹尖端应力分布

2.3 确定应力强度因子的方法

2.3.1 常用应力强度因子公式

应力强度因子是进行断裂韧度设计的前提。确定应力强度因子的方法可分为解析法、数值法和实验法。许多典型的应力强度因子已被计算出来，并汇编成了手册，可供设计者查用。表 6.3-1 列出了工程上常见裂纹形式的应力强度因子。对于手册上不能直接查到的裂纹形式，下面介绍一种工程上常用的应力强度因子计算方法——叠加法，应用叠加法可以在应力强度因子的一些基本解的基础上，求出其他较为复杂情况下的应力强度因子。

表 6.3-1　应力强度因子公式

序号	裂纹及载荷特征	应力强度因子表达式
1	无限大板，中心穿透裂纹，远处作用有均匀应力 σ、τ、τ_1	$K_I = \sigma\sqrt{\pi a}$ $K_{II} = \tau\sqrt{\pi a}$ $K_{III} = \tau_1\sqrt{\pi a}$
2	无限大板，中心穿透裂纹，远处作用有均匀拉应力 σ、$\alpha\sigma$、剪应力 τ	$K_I = \sigma\sqrt{\pi a}$ $K_{II} = \tau\sqrt{\pi a}$ 在此情况下，应力强度因子与 α 无关，亦即与无限远处平行于裂纹线方向的拉伸（或压缩）应力无关。因此，受双向均匀拉伸板（应力分别垂直和平行于裂纹线方向）与受单向均匀拉伸板（应力垂直于裂纹线方向）的应力强度因子相同

续表 6.3-1

序号	裂纹及载荷特征	应力强度因子表达式
3	无限大板，中心穿透裂纹，远处作用有与裂纹成 β 角的均匀应力 σ	$K_{\mathrm{I}}=\sigma\sqrt{\pi a}\sin^2\beta$ $K_{\mathrm{II}}=\sigma\sqrt{\pi a}\sin\beta\cos\beta$
4	垂直悬挂着的具有中心穿透裂纹的无限大板，材料密度为 ρ。假定裂纹线与地面的距离 h 足够大	$K_{\mathrm{I}}=\rho gh\sqrt{\pi a}$ $K_{\mathrm{II}}=0$ g——重力加速度
5	无限大板，中心穿透裂纹，裂纹面上某点作用有法向力 P、切向力 Q 和力矩 M。$+a$ 端的应力强度因子	$K_{\mathrm{I}}=\frac{1}{2\sqrt{\pi a}}\left[P\sqrt{\frac{a+b}{a-b}}+Q\left(\frac{k-1}{k+1}\right)+\frac{Ma}{(a-b)\sqrt{a^2-b^2}}\right]$ $K_{\mathrm{II}}=\frac{1}{2\sqrt{\pi a}}\left[Q\sqrt{\frac{a+b}{a-b}}-P\left(\frac{k-1}{k+1}\right)\right]$ 其中 $k=\frac{3-\nu}{1+\nu}$（平面应力） $k=3-4\nu$（平面应变） ν 为泊松比
6	无限大板，中心穿透裂纹，裂纹面上某点作用有一对法向力 P 和一对切向力 Q	裂纹右端的应力强度因子： $K_{\mathrm{IA}}=\frac{P}{\sqrt{\pi a}}\sqrt{\frac{a+b}{a-b}}$（$b<a$） $K_{\mathrm{II A}}=\frac{Q}{\sqrt{\pi a}}\sqrt{\frac{a+b}{a-b}}$（$b<a$） 裂纹左端的应力强度因子： $K_{\mathrm{IB}}=\frac{P}{\sqrt{\pi a}}\sqrt{\frac{a-b}{a+b}}$（$b<a$） $K_{\mathrm{IIB}}=\frac{Q}{\sqrt{\pi a}}\sqrt{\frac{a-b}{a+b}}$（$b<a$）
7	无限大板，中心穿透裂纹，几个集中法向力作用在裂纹表面上	1）裂纹面上有一对间距为 $2b$ 的单位厚度法向力 P $K_{\mathrm{I}}=\frac{P}{\sqrt{\pi a}}\frac{a}{\sqrt{a^2-b^2}}$ $K_{\mathrm{II}}=\frac{P}{\sqrt{\pi a}}\left(\frac{k-1}{k+1}\right)$ 再加上另一对间距为 $2c$ 的法向力 P $K_{\mathrm{I}}=\frac{P}{\sqrt{\pi a}}\left(\frac{a}{\sqrt{a^2-b^2}}+\frac{a}{\sqrt{a^2-c^2}}\right)$ $K_{\mathrm{II}}=0$ 2）仅裂纹下表面中央作用有集中力 P $K_{\mathrm{I}}=\frac{P}{2\sqrt{\pi a}}$ $K_{\mathrm{II}}=\frac{P}{2\sqrt{\pi a}}\left(\frac{k-1}{k+1}\right)$

续表 6.3-1

序号	裂纹及载荷特征	应力强度因子表达式
8	无限大板，中心穿透裂纹，裂纹面上作用有两对法向力 P	$K_{\mathrm{I}}=\frac{P}{\sqrt{\pi a}}\left(\sqrt{\frac{a+b}{a-b}}+\sqrt{\frac{a-b}{a+b}}\right)\ (b<a)$
9	无限大板，中心穿透裂纹，裂纹面上作用有均布应力 σ	$K_{\mathrm{I}}=\sigma\sqrt{\pi a}$ 在此情况下，对于具有中心穿透裂纹的无限板来说，在无限远处作用均匀的拉伸应力与在裂纹面上作用数值相等的均布应力时裂纹尖端的应力强度因子是相同的
10	无限大板，中心穿透裂纹，裂纹上、下表面上作用有对称的分布应力 $\sigma(x)$	裂纹右端的应力强度因子： $K_{\mathrm{IA}}=\frac{1}{\sqrt{\pi a}}\int_{-a}^{a}\sigma(X)\sqrt{\frac{a+x}{a-x}}\,\mathrm{d}x$ 裂纹左端的应力强度因子： $K_{\mathrm{IA}}=\frac{1}{\sqrt{\pi a}}\int_{-a}^{a}\sigma(X)\sqrt{\frac{a-x}{a+x}}\,\mathrm{d}x$
11	无限大板，在同一轴线上有均匀分布的多条穿透裂纹，远处作用有均匀应力 σ	$K_{\mathrm{I}}=\sigma\sqrt{\pi a}\left(\frac{2b}{\pi a}\tan\frac{\pi a}{2b}\right)^{1/2}$
12	无限大板，中心有半径为 R 的圆孔，沿直径线两端或一端的孔边有穿透裂纹，受双向应力 σ 和 $\alpha\sigma$	$K_{\mathrm{I}}=F\left(\alpha,\frac{a}{R}\right)\sigma\sqrt{\pi a}$

a/R	两侧裂纹 $F(\alpha, a/R)$			单侧裂纹 F	
	$\alpha=-1$	$\alpha=0$	$\alpha=1$	$\alpha=0$	$\alpha=1$
1.02	0.597	0.451	0.306	0.450	0.300
1.06	0.925	0.710	0.496	0.700	0.490
1.10	1.078	0.840	0.603	0.810	0.600
1.20	1.221	0.985	0.749	0.930	0.740
1.30	1.246	1.036	0.826	0.970	0.810
1.50	1.213	1.058	0.903	1.000	0.860
2.00	1.115	1.041	0.967	0.960	0.860
2.50	1.065	1.025	0.986	0.900	0.820
3.00	1.040	1.016	0.993	0.860	0.815
4.00	1.018	1.008	0.998	0.838	0.800
∞	1	1	1	0.707	0.707

续表 6.3-1

序号	裂纹及载荷特征	应力强度因子表达式
13	厚度为 B 的无限大板，具有中心穿透裂纹，对称于裂纹面的某处受集中载荷 P 作用	$K_{\mathrm{I}}=F\sigma\sqrt{\pi a}$ 其中 $\sigma=\dfrac{P}{BS}$ $F=\dfrac{S}{2a}\dfrac{(3+\nu)+2\left(\dfrac{a}{S}\right)^2}{\left[(a/S^2)+1\right]^{3/2}}$ 式中，ν 为泊松比
14	半无限板，边缘穿透裂纹，远处作用有均匀应力 σ	$K_{\mathrm{I}}=1.215\sigma\sqrt{\pi a}$
15	半无限板，边缘穿透裂纹，在裂纹的上、下表面分别作用有集中力 P 和 Q	$K_{\mathrm{I}}=\dfrac{2}{\pi}\dfrac{1+F\left(\dfrac{b}{a}\right)}{\sqrt{a^2-b^2}}P\sqrt{\pi a}$ $K_{\mathrm{II}}=\dfrac{2}{\pi}\dfrac{1+F\left(\dfrac{b}{a}\right)}{\sqrt{a^2-b^2}}Q\sqrt{\pi a}$ 其中　$F\left(\dfrac{b}{a}\right)=\left(1-\dfrac{b}{a}\right)\left[0.2945-0.392\left(\dfrac{b}{a}\right)^2+0.7685\left(\dfrac{b}{a}\right)^4-0.9942\left(\dfrac{b}{a}\right)^6+0.5094\left(\dfrac{b}{a}\right)^8\right]$
16	具有中心穿透裂纹的有限宽板，远处作用有拉伸应力 σ	$K_{\mathrm{I}}=\sigma\sqrt{\pi a}\left(\dfrac{W}{\pi a}\tan\dfrac{\pi a}{W}\right)^{1/2}$
17	具有中心穿透裂纹的有限宽板，裂纹面中心受集中载荷 P 作用	$K_{\mathrm{I}}=F\dfrac{P}{BW}\sqrt{\pi a}$ $F=\dfrac{W\left(1.297-0.297\cos\dfrac{\pi a}{W}\right)}{\pi a}\left[\dfrac{W}{2\pi a}\sin\left(\dfrac{2\pi a}{W}\right)\right]^{-1/2}$

续表 6.3-1

序号	裂纹及载荷特征	应力强度因子表达式
18	宽度为 $2b$ 的无限长板，中心裂纹，远处作用有弯矩 M	$K_{\mathrm{I}} = F\sigma\sqrt{\pi a}$ $\sigma = \frac{3Ma}{2b^3}$ 其中 $F = \frac{\sqrt{1-a/b}}{2\,(1-(a/b))^3}\left[1+\frac{1}{2}\left(\frac{a}{b}\right)+\frac{3}{8}\left(\frac{a}{b}\right)^2-\frac{11}{16}\left(\frac{a}{b}\right)^3+0.464\left(\frac{a}{b}\right)^4\right]$
19	宽度为 b 的无限长板，边缘裂纹，远处作用有应力 σ	$K_{\mathrm{I}} = F\sigma\sqrt{\pi a}$ 其中 $F = 1.12-0.231\left(\frac{a}{b}\right)+10.5\left(\frac{a}{b}\right)^2-21.72\left(\frac{a}{b}\right)^3+30.39\left(\frac{a}{b}\right)^4$
20	宽度为 b 的无限长板，边缘裂纹，远处作用有弯矩 M	$K_{\mathrm{I}} = F\sigma\sqrt{\pi a}$ $\sigma = \frac{6M}{b^2}$ 其中 $F = 1.122-1.4\left(\frac{a}{b}\right)+7.33\left(\frac{a}{b}\right)^2-13.08\left(\frac{a}{b}\right)^3+14.0\left(\frac{a}{b}\right)^4$
21	宽度为 $2b$ 的无限长板，双边裂纹，远处作用有应力 σ	$K_{\mathrm{I}} = F\sigma\sqrt{\pi a}$ 其中 $F = 1.12+0.203\left(\frac{a}{b}\right)-1.197\left(\frac{a}{b}\right)^2+1.93\left(\frac{a}{b}\right)^3$
22	有限宽板，偏心裂纹，远处作用有应力 σ	$K_{\mathrm{I}} = F\sigma\sqrt{\pi a}$ A 端： $F = \sqrt{\sec\frac{\pi}{6}\left(\frac{2.65a}{b_1}+\frac{0.35a}{b_2}\right)}$　$b_1 \leqslant b_2$ B 端： $F = \sqrt{\sec\frac{\pi}{6}\left(\frac{1.45a}{b_1}+\frac{1.55a}{b_2}\right)}$　$b_1 \leqslant b_2$

续表 6.3-1

序号	裂纹及载荷特征	应力强度因子表达式
23	紧凑拉伸试件，厚度为 B，$H=0.6\ W$，$W_1=1.25\ W$	$K_{\mathrm{I}}=F\dfrac{P}{B\sqrt{W}}$ 其中 $F=29.6\left(\dfrac{a}{W}\right)^{1/2}-185.5\left(\dfrac{a}{W}\right)^{3/2}+655.7\left(\dfrac{a}{W}\right)^{5/2}-1\,017.0\left(\dfrac{a}{W}\right)^{7/2}+638.9\left(\dfrac{a}{W}\right)^{9/2}$
24	三点弯曲试件，厚度为 B，$S/W=4$，$0.4\leqslant\dfrac{a}{W}\leqslant0.6$	$K_{\mathrm{I}}=F\dfrac{PS}{BW^{3/2}}$ 其中 $F=2.9\left(\dfrac{a}{W}\right)^{1/2}-4.6\left(\dfrac{a}{W}\right)^{3/2}+21.8\left(\dfrac{a}{W}\right)^{5/2}-37.6\left(\dfrac{a}{W}\right)^{7/2}+38.7\left(\dfrac{a}{W}\right)^{9/2}$ 或 $F=\left[1.88+0.75\left(\dfrac{a}{W}-0.50\right)^2\right]\sec\left(\dfrac{\pi a}{2W}\right)\times\sqrt{\tan\left(\dfrac{\pi a}{2W}\right)}$
25	无限体，内部圆片状裂纹，远处作用 σ 和 τ	$K_{\mathrm{I}A}=\dfrac{2}{\pi}\sigma\sqrt{\pi a}$　$K_{\mathrm{II}A}=\dfrac{4\cos\theta}{\pi(2-\nu)}\tau\sqrt{\pi a}$ $K_{\mathrm{III}A}=\dfrac{4(2-\nu)\sin\theta}{\pi(2-\nu)}\tau\sqrt{\pi a}$
26	无限体，内部椭圆片状裂纹，远处作用 σ、τ 和 τ_1	$K_{\mathrm{I}A}=\dfrac{\sigma\sqrt{\pi a}}{E(k)}(1-k^2\cos^2\theta)^{1/4}$ $K_{\mathrm{I}\max}=K_{\mathrm{I}(\theta=\pi/2)}=\sigma\sqrt{\pi a}/E(k)$ $K_{\mathrm{I}}(\mathrm{ezze})=(2\sigma/\pi)\sqrt{\pi a}$ $K_{\mathrm{I}}(e\rightarrow\infty)=\sigma\sqrt{\pi a}$ $K_{\mathrm{I}A}=\left(\tau_1\dfrac{k^1\cos\theta}{B}+\tau\dfrac{\sin\theta}{C}\right)\dfrac{\sqrt{\pi a}k^2}{(1-k^2\cos^2\theta)^{1/4}}$ $K_{\mathrm{II}A}=\left(\tau_1\dfrac{\sin\theta}{B}-\tau\dfrac{k^1\cos\theta}{D}\right)\dfrac{(1-\nu)\sqrt{\pi a}k^2}{(1-k^2\cos^2\theta)^{1/4}}$ 其中 $k^1=a/c$，$k^2=1-k^{12}=1-a^2/c^2$ θ 为图示 A 点的离心角 $B=(k^2-\nu)E(k)+\nu k^{12}K(k)$ $C=(k^2+\nu k^{12})E(k)-\nu k^2K(k)$ $K(k)=\int_0^{\pi/2}\dfrac{\mathrm{d}\theta}{\sqrt{1-k^2\sin^2\theta}}$（第一类完全椭圆积分） $E(k)=\int_0^{\pi/2}\sqrt{1-k^2\sin^2\theta}\,\mathrm{d}\theta$（第二类完全椭圆积分）

续表 6.3-1

序号	裂纹及载荷特征	应力强度因子表达式
27	厚度为 h 的板，表面半椭圆形裂纹垂直于板面，垂直于裂纹平面方向作用 σ	见下

序号 27　应力强度因子表达式

裂纹前缘各点的应力强度因子

$$K_{\mathrm{I}} = M\sigma \frac{\sqrt{\pi a}}{E(k)}\left[\sin^2\theta + \left(\frac{a}{c}\right)^2\cos^2\theta\right]^{1/4}$$

其中　$E(k)$ 为第二类完全椭圆积分，$k^2 = 1 - a^2/c^2$，a 为椭圆短半轴，c 为椭圆长半轴

裂纹最深点处的应力强度因子

$$K_{\mathrm{I}} = M\sigma \frac{\sqrt{\pi a}}{E(k)}$$

如果考虑小范围屈服修正，裂纹最深点处应力强度因子

$$K_{\mathrm{I}} = M\sigma \sqrt{\frac{\pi a}{Q}}$$

裂纹形状因子 $Q = E^2(k) - 0.212\left(\frac{\sigma}{\sigma_s}\right)^2$

形状因子 Q 数值

σ/σ_s \ $a/2c$	0.10	0.20	0.25	0.30	0.40
1.0	0.88	1.07	1.21	1.38	1.76
0.9	0.91	1.12	1.24	1.41	1.79
0.8	0.95	1.15	1.27	1.45	1.83
0.7	0.98	1.17	1.31	1.48	1.87
0.6	1.02	1.22	1.35	1.52	1.90
<0.6	1.10	1.29	1.43	1.60	1.98

M 数值

$a/2c$ \ a/h	0	0.2	0.4	0.6	0.8
0.1	1.10	1.11	1.14	1.21	1.39
0.2	1.08	1.08	1.10	1.16	1.30
0.3	1.06	1.06	1.08	1.11	1.23
0.4	1.04	1.04	1.04	1.08	1.17
0.5	1.02	1.02	1.02	1.06	1.13

序号	裂纹及载荷特征	应力强度因子表达式
28	板上圆孔内壁有两个半椭圆形裂纹，远处作用拉应力 σ	见下

剖面 A—A

序号 28　应力强度因子表达式

裂纹前缘各点应力强度因子

$$K_{\mathrm{I}} = M\sigma \sqrt{\frac{\pi c}{Q}}\left[\left(\frac{c}{a}\right)^2\sin^2\theta + \cos^2\theta\right]^{1/4} \quad \left(当\frac{c}{a} \leqslant 1\right)$$

$$K_{\mathrm{I}} = M\sigma \sqrt{\frac{\pi a}{Q}}\left[\sin^2\theta + \left(\frac{a}{c}\right)^2\cos^2\theta\right]^{1/4} \quad \left(当\frac{c}{a} > 1\right)$$

其中　Q 为形状因子，同序 27

M 由下表确定

M 数值

c/r \ θ	0	20°	40°	60°	90°
0.1	2.90	2.80	2.70	2.62	2.60
0.2	2.80	2.64	2.47	2.35	2.30
0.4	2.65	2.35	2.13	1.98	1.92
0.6	2.55	2.15	1.90	1.75	1.70
1.0	2.35	1.85	1.63	1.50	1.47
2.0	2.15	1.54	1.34	1.28	1.24
5.0	1.75	1.23	1.14	1.10	1.08
10.0	1.50	1.10	1.07	1.05	1.04

续表 6.3-1

<table>
<tr><th>序号</th><th>裂纹及载荷特征</th><th>应力强度因子表达式</th></tr>
<tr><td>29</td><td>厚度为 h 的板上有一圆孔，孔边两端各有一个四分之一椭圆形裂纹，板承受拉应力 σ</td><td>裂纹前缘各点的应力强度因子
$K_{\rm I} = MM_1M_2\sigma\sqrt{\frac{\pi a}{Q}}\left[\left(\frac{a}{c}\right)^2\sin^2\theta+\cos^2\theta\right]^{1/4}$（当 $\frac{a}{c}\leqslant 1$）
$K_{\rm I} = MM_1M_2\sigma\sqrt{\frac{\pi c}{Q}}\left[\sin^2\theta+\left(\frac{c}{a}\right)^2\cos^2\theta\right]^{1/4}$（当 $\frac{a}{c}>1$）
其中　Q 为形状因子，同序 27
M 值同序 28
$M_1 = 1+0.12\left(1-\frac{a}{2c}\right)^2$
M_2 值由下表确定
M_2 数值
<table>
<tr><th>c/h \ θ</th><th>0°</th><th>30°</th><th>60°</th><th>90°</th></tr>
<tr><td>0.4</td><td>1.00</td><td>1.00</td><td>1.00</td><td>1.00</td></tr>
<tr><td>0.6</td><td>1.025</td><td>1.020</td><td>1.015</td><td>1.010</td></tr>
<tr><td>0.8</td><td>1.100</td><td>1.080</td><td>1.045</td><td>1.020</td></tr>
<tr><td>0.9</td><td>1.200</td><td>1.160</td><td>1.075</td><td>1.030</td></tr>
</table></td></tr>
<tr><td>30</td><td>厚度为 h 的板上有一圆孔，孔边有一个四分之一椭圆形裂纹，板承受拉应力 σ</td><td>裂纹前缘各点的应力强度因子
$K_{\rm I} = MM_1M_2\sqrt{\frac{2r+\pi ac/4h}{2r+2\pi ac/4h}}\sigma\sqrt{\frac{\pi a}{Q}}\times\left[\left(\frac{a}{c}\right)^2\sin^2\theta+\cos^2\theta\right]^{1/4}$
（当 $\frac{a}{c}\leqslant 1$）
$K_{\rm I} = MM_1M_2\sqrt{\frac{2r+\pi ac/4h}{2r+2\pi ac/4h}}\sigma\sqrt{\frac{\pi c}{Q}}\times\left[\sin^2\theta+\left(\frac{c}{a}\right)^2\cos^2\theta\right]^{1/4}$
（当 $\frac{a}{c}>1$）
其中　Q 形状因子，同序 27
M 同序 28
M_1、M_2 同序 29</td></tr>
<tr><td>31</td><td>圆柱的半径为 R，圆颈的半径为 b，承受轴向拉力 P，弯矩 M 和扭矩 T</td><td>$K_{\rm I} = \frac{P}{\pi b^2}\sqrt{\pi b}F_1+\frac{4M}{\pi b^3}\sqrt{\pi b}F_2$
$K_{\rm II} = 0$
$K_{\rm III} = \frac{2T}{\pi b^2}\sqrt{\pi b}F_3$
其中　$F_1 = \frac{1}{2}\left[1+\frac{1}{2}\left(\frac{b}{R}\right)+\frac{3}{8}\left(\frac{b}{R}\right)^2-0.363\left(\frac{b}{R}\right)^3+0.731\left(\frac{b}{R}\right)^4\right]\sqrt{1-\frac{b}{R}}$
$F_2 = \frac{3}{8}\left[1+\frac{1}{2}\left(\frac{b}{R}\right)+\frac{3}{8}\left(\frac{b}{R}\right)^2+\frac{5}{16}\left(\frac{b}{R}\right)^3+\frac{35}{128}\left(\frac{b}{R}\right)^4+0.537\left(\frac{b}{R}\right)^5\right]\sqrt{1-\frac{b}{R}}$
$F_3 = \frac{3}{8}\left[1+\frac{1}{2}\left(\frac{b}{R}\right)+\frac{3}{8}\left(\frac{b}{R}\right)^2+\frac{5}{16}\left(\frac{b}{R}\right)^3+\frac{35}{128}\left(\frac{b}{R}\right)^4+0.208\left(\frac{b}{R}\right)^5\right]\sqrt{1-\frac{b}{R}}$</td></tr>
</table>

续表 6.3-1

<table>
<tr><th>序号</th><th>裂纹及载荷特征</th><th>应力强度因子表达式</th></tr>
<tr><td>32</td><td>圆截面杆，中心圆片状裂纹</td><td>

$K_{\mathrm{I}A}=\frac{P}{\pi(b^2-a^2)}F_1+\frac{4Ma}{\pi(b^4-a^4)}F_2$

$K_{\mathrm{II}}=0$

$K_{\mathrm{III}}=\frac{2Ta}{\pi(b^4-a^4)}F_2$

其中　$F_1=\frac{2}{\pi}[1+\frac{1}{2}(\frac{a}{b})-\frac{5}{8}(\frac{a}{b})^2+0.421(\frac{a}{b})^3]$

$F_2=\frac{4}{3\pi}[1+\frac{1}{2}(\frac{a}{b})+\frac{3}{8}(\frac{a}{b})^2+\frac{5}{16}(\frac{a}{b})^4-\frac{93}{128}(\frac{a}{b})^4+0.483(\frac{a}{b})^5]\sqrt{1-\frac{a}{b}}$

$F_3=\frac{4}{3\pi}[1+\frac{1}{2}(\frac{a}{b})+\frac{3}{8}(\frac{a}{b})^2+\frac{5}{16}(\frac{a}{b})^3-\frac{93}{128}(\frac{a}{b})^4+0.038(\frac{a}{b})^5]\sqrt{1-\frac{a}{b}}$

</td></tr>
<tr><td>33</td><td>长圆筒，内部有一条径向长度为 a 的轴向裂纹，圆筒内半径为 R_1，外半径为 R_2，受内压 P</td><td>

$K_{\mathrm{I}}=\frac{2PR_2^2\sqrt{\pi a}}{R_2^2-R_1^2}F$

其中　F 值由下表确定

<table>
<tr><th>b/R_1 \ a/b</th><th>$\frac{1}{8}$</th><th>$\frac{1}{4}$</th><th>$\frac{1}{2}$</th><th>$\frac{3}{4}$</th></tr>
<tr><td>$\frac{1}{5}$</td><td>1.19</td><td>1.38</td><td>2.10</td><td>3.30</td></tr>
<tr><td>$\frac{1}{10}$</td><td>1.20</td><td>1.44</td><td>2.36</td><td>4.23</td></tr>
<tr><td>$\frac{1}{20}$</td><td>1.20</td><td>1.45</td><td>2.51</td><td>5.25</td></tr>
</table>

</td></tr>
<tr><td>34</td><td>内半径为 R_1，外半径为 R_2 的长圆筒内部有条环向裂纹，圆管端部受拉应力 σ</td><td>

$K_{\mathrm{I}}=\sigma\sqrt{\pi a}F$

加载点位移

$\Delta_c=\frac{4\sigma a}{E'}V$

其中　F、V 值由下表确定

<table>
<tr><th>b/R_1 \ a/b</th><th></th><th>$\frac{1}{8}$</th><th>$\frac{1}{4}$</th><th>$\frac{1}{2}$</th><th>$\frac{3}{4}$</th></tr>
<tr><td rowspan="2">$\frac{1}{5}$</td><td>F</td><td>1.16</td><td>1.26</td><td>1.61</td><td>2.15</td></tr>
<tr><td>V</td><td>0.117</td><td>0.255</td><td>0.743</td><td>1.67</td></tr>
<tr><td rowspan="2">$\frac{1}{10}$</td><td>F</td><td>1.19</td><td>1.32</td><td>1.82</td><td>2.49</td></tr>
<tr><td>V</td><td>0.180</td><td>0.290</td><td>0.885</td><td>2.09</td></tr>
<tr><td rowspan="2">$\frac{1}{20}$</td><td>F</td><td>1.22</td><td>1.36</td><td>2.03</td><td>2.89</td></tr>
<tr><td>V</td><td>0.220</td><td>0.315</td><td>1.04</td><td>2.74</td></tr>
</table>

</td></tr>
</table>

2.3.2 计算应力强度因子的叠加法

叠加法是求解应力强度因子的一种简便方法。其做法是把一个复杂结构看成几个几何形式相同的分离结构的组合，而各分离结构都有较简单的边界条件和已知的应力强度因子。把各分离结构的应力强度因子进行叠加，就可得到所求的解。根据线弹性力学中的叠加原理，几个外载荷同时作用所产生的应力和位移等于各载荷单独作用所产生的应力和位移之和。由线弹性断裂力学可知，对于同一类型的裂纹，裂纹尖端附近应力场与坐标 r、θ 之间的关系是一定的，外加载荷的变化只能引起应力强度因子的线性变化。所以当几种载荷 P_1，P_2，…，P_n 联合作用时的应力强度因子 K_{I} 就等于各载荷单独作用下的应力强度因子的代数和，即

$$K_{\mathrm{I}}=(K_{\mathrm{I}})_1+(K_{\mathrm{I}})_2+\cdots+(K_{\mathrm{I}})_n \qquad (6.3\text{-}25)$$

例如，图 6.3-11a 中所示的载荷可分解为 B、D、E 三种情况的组合。情况 A 的应力强度因子可根据分解后的 3 种载荷情况下的应力强度因子的叠加求得，裂纹体的应力强

度因子 K_A 等于图6.3-11中B、D、E三种情况的叠加，即 $K_A=K_B+K_D-K_E$。

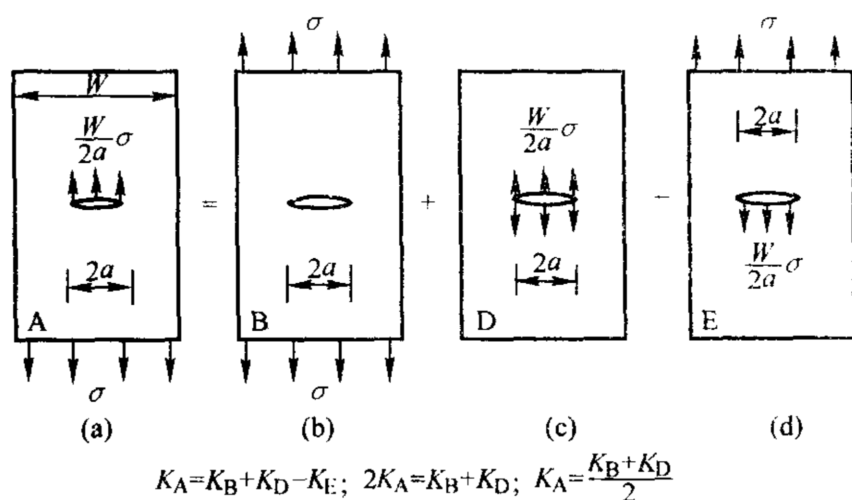


图6.3-11　用叠加法求应力强度因子

由于情况A与情况E的应力强度因子相同，所以有

$$K_A=(K_B+K_D)/2$$

在情况B的应力强度因子 $K_B=\sigma\sqrt{\pi a}$ 和情况D的应力强度因子 $K_D=\frac{W}{2a}\sigma\sqrt{\pi a}$ 为已知的条件下，可求得情况A时的应力强度因子为

$$K_A=0.5\left(1+\frac{W}{2a}\right)\sigma\sqrt{\pi a}$$

2.4　材料断裂韧度

断裂韧度是表征材料抗断裂能力的性能指标。在试样厚度 t 足够大时（$t\geqslant 2.5K_{IC}/\sigma_s^2$，$\sigma_s$ 为材料屈服极限），测得的 K_{Ic} 称为平面应变断裂韧度。断裂韧度与传统的材料强度或塑性指标有两点基本差别。

1）传统的材料强度和塑性指标，是在材料为均匀、连续状态下测得的，而断裂韧度是在材料已带有裂纹的情况下测得的。前者只代表无裂纹时材料的力学性能，断裂韧度则代表具有裂纹时材料的力学性能。

2）传统的性能指标中，屈服强度 σ_s 和极限强度 σ_b 与塑性延伸率 δ 和截面收缩率 ψ 以及冲击韧性是分开的，特别是在设计零部件时无法定量地考虑这些塑性和韧性指标。而断裂韧度是强度和塑性的综合指标。例如，根据断裂判据 $K_I=K_{Ic}$，在知道零件的工作应力和存在的裂纹尺寸的条件下，便可计算出不产生脆断所需要的材料的最低断裂韧度值。

断裂力学为材料建立了一系列指标，其目的都是为了防止带裂纹零部件的断裂，统称为抗断裂指标。其中包括静载荷线弹性条件下的断裂韧度 K_{Ic}、弹塑性情况下的裂纹尖端张开位移临界值 δ_c 和 J 积分的临界值 J_{Ic}；循环载荷下疲劳裂纹扩展速率 da/dN 和疲劳裂纹扩展的门槛值 ΔK_{th}；在腐蚀介质中抗应力腐蚀的界限应力强度因子 K_{Iscc} 和应力腐蚀速率 da/dt、高温条件下高温断裂韧度和蠕变裂纹扩展速率 da/dt 等。

材料的断裂韧度受冶金因素、热处理制度、使用环境等因素的影响，精确的数据须由实验获得。在工程设计中，也可从有关手册上查找适用数据。

根据材料的断裂韧度值，很容易计算含裂纹构件的临界裂纹长度 a_c。对于常规环境下受静载荷的构件来说，只要其初始裂纹尺寸 a_0 不大于临界裂纹尺寸 a_c，该裂纹就不会扩展，不会导致构件断裂。然而，对于腐蚀环境下的含裂纹构件来说，小于临界尺寸的裂纹也会随受载时间的增加而逐渐扩展，直到最终失稳断裂。这种在应力和腐蚀介质共同作用下产生的裂纹扩展和断裂称为应力腐蚀开裂，记为SCC。各种材料在腐蚀环境下也都存在一个临界应力强度因子，Ⅰ型裂纹记为 K_{ISCC}。若外载引起的应力强度因子小于材料在相应环境下的腐蚀临界应力强度因子，即 $K_I<K_{ISCC}$，则不会发生应力腐蚀开裂。可见 K_{ISCC} 是衡量材料在腐蚀介质中抵抗应力腐蚀裂纹扩展能力的指标，称为材料的应力腐蚀断裂韧度。

2.5　弹塑性断裂力学的基本理论与判据

弹塑性断裂力学用于分析裂纹顶端附近材料已产生大范围屈服情况下的裂纹体强度问题。分析弹塑性断裂的方法有裂纹顶端张开位移（COD）法和 J 积分法等，简介如下。

2.5.1　COD法

当裂纹受到垂直于裂纹表面的拉力作用时，上下两个裂纹表面就会张开。裂纹面在裂纹顶端处的张开位移就是所谓的COD，通常用符号 δ 表示（见图6.3-12）。显然，随着载荷的增加，裂纹顶端张开位移量也随之增加。也就是说，一定量的COD对应于一定的受载状态，或者说对应着一定的裂纹顶端的应力应变场的强弱程度。因此可以用其间接地衡量裂纹顶端附近的应力应变场强度。

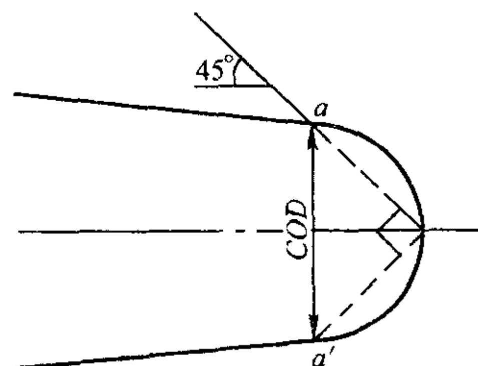


图6.3-12　裂纹顶端处的张开位移

受单向均匀拉应力作用的具有中心穿透裂纹的无限大板，当应力水平 $\sigma<\sigma_s$ 时，其COD值为

$$\delta=\frac{8\sigma_s a}{\pi E}\ln\sec\frac{\pi\sigma}{2\sigma_s} \tag{6.3-26}$$

应力水平 $\sigma\geqslant\sigma_s$ 时，即在全面屈服情况下，COD值为

$$\delta=2\pi ae \tag{6.3-27}$$

式中，e 为名义应变。

与线弹性断裂情况一样，对于弹塑性断裂，当所加的载荷达到某一临界值时，含裂纹体也达到了断裂的临界状态。此时，裂纹顶端的张开位移量 δ 也达到了其临界值 δ_c，此临界值 δ_c 为材料的断裂韧度指标。于是，COD表示的断裂判据可表达为

$$\delta\geqslant\delta_c \tag{6.3-28}$$

2.5.2　J 积分法

J 积分是表示裂纹周围应力应变场强度的量，该参量是一个与路径无关的线积分，定义为

$$J=\int_\Gamma\left(W\mathrm{d}y-\overline{T}\frac{\partial\overline{u}}{\partial x}\mathrm{d}s\right) \tag{6.3-29}$$

式中，W 为应变能密度或称形变功密度，其值为

$$W=\int_0^\varepsilon\sigma_{ij}\mathrm{d}\varepsilon_{ij} \tag{6.3-30}$$

$\overline{T}$ 为积分回路上任意弧线元所对应的面元 $\mathrm{d}s\mathrm{d}z$ 上的应力矢量；$\overline{u}$ 为应力 $\overline{T}$ 作用点的位移矢量；Γ 为从裂纹下自由表面上任一点开始，按逆时针方向环绕裂纹尖端区域，终止于裂纹上自由表面上任一点的任意积分路线。

可以证明，J 积分同样可以作为裂纹尖端附近应力应变场强度的度量。当此 J 积分值达到某一临界值时，裂纹就开始扩展。此临界值即为材料的断裂韧度指标，记为 J_{Ic}。于是断裂判据可表达为

$$J\geqslant J_{Ic} \tag{6.3-31}$$

J_{Ic}与 K_{Ic}有如下关系：

$$J_{Ic}=\frac{1-\nu^2}{E}K_{IC}^2$$

3　裂纹扩展计算

一个含裂纹体在静载荷作用下，只有当应力强度因子大于材料的断裂韧度时，裂纹才会失稳扩展。然而，含裂纹体在循环载荷作用下，即使与最大载荷对应的应力强度因子远小于材料的断裂韧度，裂纹仍可能会随载荷的循环作用逐步扩展，裂纹扩展到临界尺寸时就会导致构件疲劳破坏。

恒幅循环载荷下的实验结果表明，疲劳裂纹扩展速率 da/dN 随应力强度因子幅度 ΔK 的变化而变化，在双对数坐标上大致呈反 S 形，如图 6.3-13 所示。

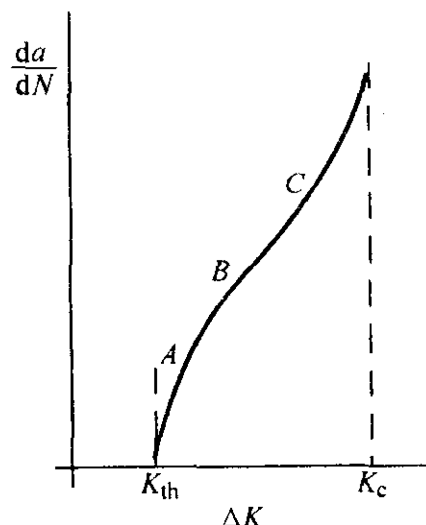


图 6.3-13　$da/dN-\Delta K$ 曲线

图中 A 区为短裂纹扩展区，该区存在一个垂直渐近线 $\Delta K=\Delta K_{th}$，ΔK_{th}称为疲劳裂纹扩展应力强度因子幅度门槛值。在 A 区内，当 ΔK 逐渐减小并趋向于 ΔK_{th}时，da/dN 趋向于零。C 区为裂纹快速扩展区，该区也存在一条垂直渐近线 $K_{max}=K_c$，K_c 为材料的断裂韧度，其值与材料厚度有关。通常材料的断裂韧度都用 Ⅰ 型裂纹的平面应变断裂韧度 K_{Ic} 表示。在 C 区内 da/dN 急剧增加并最终产生瞬间断裂。曲线的中间段为裂纹稳定扩展阶段，称为 B 区。该区内 $da/dN-\Delta K$ 在双对数坐标中近似呈直线，是裂纹扩展的主要阶段。以下涉及的主要是 B 区的裂纹扩展问题。

对于存在初始裂纹或类裂纹缺陷的工程实际构件，虽然开始使用时在工作载荷下并不会产生断裂，但由于绝大多数工程构件在服役期间都要承受循环载荷的作用，初始裂纹会随载荷的不断作用而不断增长，应力强度因子当然也随之增加。当应力强度因子达到材料断裂韧度时，就导致构件中的裂纹失稳扩展。从初始裂纹扩展到临界裂纹所经历的载荷循环次数即为构件的疲劳裂纹扩展寿命。下面介绍预测裂纹疲劳扩展寿命的方法。

3.1　疲劳裂纹扩展速率的基本公式

首先介绍恒幅疲劳载荷下的裂纹扩展规律。疲劳载荷作用下的裂纹扩展速率 da/dN 定义为每一个载荷循环作用下的裂纹扩展量，可表示为应力强度因子的函数，受应力比、载荷顺序、环境条件和载荷频率等因素的影响。

循环载荷一般用最大应力σ_{max}、最小应力σ_{min}、应力比 R（$R=\sigma_{min}/\sigma_{max}$）、应力幅度 $\Delta\sigma$（$\Delta\sigma=\sigma_{max}-\sigma_{min}$）、加载频率 f 等参量来描述。对于疲劳裂纹扩展速率而言，应力幅度 $\Delta\sigma$ 是最重要的控制参量。在应力幅度相同的条件下，应力比 R 越大，裂纹扩展速率 da/dN 也越大。恒幅载荷不存在载荷顺序问题。在变幅载荷情况下，如果整个载荷历程中不同载荷峰值相差较大，载荷顺序对裂纹扩展速率的影响则不能忽略。载荷顺序的影响包括拉伸超载引起的裂纹扩展延迟效应、压缩过载引起的裂纹扩展加速效应和迟滞减缓效应。计算裂纹扩展时如不考虑拉伸过载引起的延迟效应，所得裂纹扩展寿命是偏于保守的；如不考虑压缩过载引起的加速效应和迟滞减缓效应，则所得裂纹扩展寿命是偏于危险的。载荷顺序对裂纹扩展寿命的影响有时会达到一倍寿命以上。但是考虑载荷顺序影响会使裂纹扩展分析过于复杂，在超载比不大于 1.3 时，可以忽略载荷顺序对裂纹扩展速率的影响。加载频率影响的一般趋势是，加载频率减小，裂纹扩展速率增大。但是随着应力强度因子幅度的减小，加载频率的影响也逐渐减小。对于中、低加载频率的情况，可以不计加载频率的影响。但是在高温和腐蚀环境下，加载频率对裂纹扩展速率影响较大，一般不能忽略。环境影响主要表现在腐蚀介质对裂纹扩展所起的作用，通常分为应力腐蚀和腐蚀疲劳两个方面。在腐蚀介质中，构件受静应力作用而产生的持续损伤现象称为应力腐蚀；构件受交变应力作用，介质腐蚀与疲劳两种损伤机理联合产生的破坏称为腐蚀疲劳。

应力强度因子幅度 ΔK 是描述裂纹扩展的基本参量，应力比 R 是一个重要影响因素。裂纹扩展速率 da/dN 通常表示为 ΔK 的函数，即 $da/dN=f(\Delta K)$ 的形式。目前常用的裂纹扩展速率公式有Paris公式、Forman公式和Walker公式等。

（1）Paris 公式

最简明易用的裂纹扩展速率公式是 Paris 公式：

$$\frac{da}{dN}=C(\Delta K)^n \qquad (6.3\text{-}32)$$

式中，C 和 n 是由实验确定的材料常数。在工程设计中，这些参数也可从有关手册上查找。

Paris 公式可用于描述给定应力比时裂纹扩展第二阶段的裂纹扩展速率特性。如超出公式的使用限制范围，将会引起较大的寿命计算误差。

（2）Forman 公式

Forman 公式是一个适应性较强，应用范围较广的裂纹扩展速率公式：

$$\frac{da}{dN}=\frac{C(\Delta K)^n}{(1-R)K_c-\Delta K} \qquad (6.3\text{-}33)$$

式中，C 和 n 是由实验确定的材料常数；R 是应力比；K_c 是材料的平面应变断裂韧度。式（6.3-33）也可以表示为

$$\frac{da}{dN}=\frac{C(1-R)^{n-1}K_{max}{}^n}{K_c-K_{max}} \qquad (6.3\text{-}34)$$

式中，K_{max}为最大应力强度因子。

Forman 公式具有描述不同应力比下的裂纹扩展速率的能力。该公式可用于描述裂纹扩展中间段和快速扩展阶段的疲劳裂纹扩展。当应力强度因子最大值 K_{max}接近其临界值时，Forman 公式还能反映裂纹的快速扩展特性。

（3）Walker 公式

另一个适用范围较广的裂纹扩展速率公式是 Walker 公式：

$$\frac{da}{dN}=C\left[(1-R)^{M-1}\Delta K\right]^n \text{或} \frac{da}{dN}=C\left[(1-R)^M K_{max}\right]^n \qquad (6.3\text{-}35)$$

式中，C、M 和 n 为由实验确定的材料常数。Walker 公式适用于描述裂纹扩展速率特性的中间区域。该公式考虑了负应力比的影响，即考虑了压缩载荷对裂纹扩展的加速作用，故应用面较广，计算精度也较高。

在使用上述裂纹扩展速率公式时，应注意以下几点。

1）上述公式中的材料常数 C 和 n 并不是全部可以互换的。例如，Paris 公式中的 C 和 n 可以用于 Walker 公式，但不能用于 Forman 公式。

2）由实验数据拟合材料常数 C 和 n 时与公式的适用范围相匹配。例如，Paris 公式中的 C 和 n 与应力比 R 有关，

是一定应力比情况下的 C、n 值。并且只有当 $R=0$ 时其值才与 Walker 公式中的相应参数一致。Paris 公式和 Walker 公式只适用于裂纹扩展的第二阶段，故公式中的参数要由该阶段的实验数据拟合。

3) 上述公式中的常数 C 是有量纲的，其量纲为 $[\mathrm{m}\ (\mathrm{MPa}\sqrt{\mathrm{m}})^{-n}]$ /周或 $[\mathrm{mm}\ (\mathrm{N \cdot mm^{3/2}})^{-n}]$ /周。这一点在单位换算时要特别注意。

4) 应用上述公式时还应考虑环境的影响。

3.2 超载延迟效应对疲劳裂纹扩展的影响

在变幅疲劳载荷作用下工作的构件，载荷谱中若出现了能导致裂纹尖端较大屈服的高幅值载荷，会使随后一个阶段内的裂纹扩展速率降低，这种现象称为裂纹扩展的超载延迟效应。超载延迟效应能够增加构件的裂纹扩展寿命，在估算构件的裂纹扩展寿命时若不计及超载效应，会使估算结果偏于保守。下面介绍一种较为简单的超载延迟效应修正方法—Wheeler 模型。

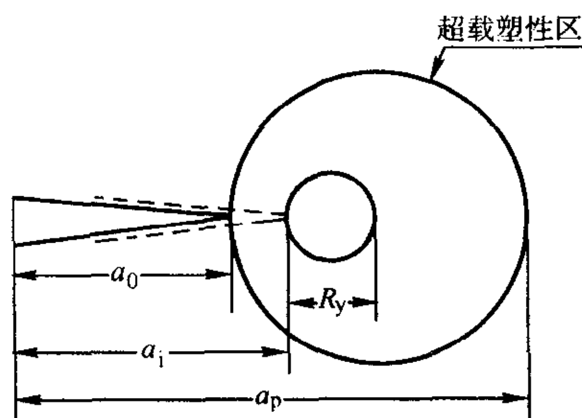


图 6.3-14 裂纹尖端超载塑性区

Wheeler 模型认为超载后裂纹扩展速率的降低与裂纹尖端在超载塑性区中的位置有关（见图 6.3-14），当与当前裂纹长度及应力水平对应的塑性区越出超载塑性区时，裂纹扩展延迟效应消失，其数学模型为

$$\left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_{\mathrm{r}} = C_{\mathrm{ri}}\left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_{\mathrm{o}} \tag{6.3-36}$$

式中，$(\mathrm{d}a/\mathrm{d}N)_{\mathrm{r}}$ 为受超载延迟效应减缓了的裂纹扩展速率；$(\mathrm{d}a/\mathrm{d}N)_{\mathrm{o}}$ 为不计超载延迟效应时的裂纹扩展速率，C_{ri} 为超载效应修正系数：

$$C_{\mathrm{ri}} = \left(\frac{R_y}{a_p - a_i}\right)^m \tag{6.3-37}$$

式中，R_y 为与当前载荷对应的裂纹尖端塑性区尺寸；$(a_p - a_0)$ 为超载峰值产生的大塑性区尺寸，a_0 为出现超载时的裂纹长度；m 是与材料及载荷谱有关的系数，需要由实验测定。

应该说明的是，Wheeler 模型比较简单，但也很粗糙，不能反映超载后裂纹扩展速率变化的真实情况，且模型系数依赖载荷谱型，应用时也不很方便。更精确的模型有修正的 Willenberg 模型和裂纹闭合模型等，但都很复杂，需要时可参照有关文献。

3.3 恒幅疲劳载荷下的裂纹扩展寿命计算

恒幅疲劳载荷下的裂纹扩展寿命 N 可以通过对裂纹扩展速率方程进行积分运算获得。由 Paris 公式可得

$$N = \int_{a_0}^{a_c} \frac{\mathrm{d}a}{C\ (\Delta K)^n} \tag{6.3-38}$$

式中，a_0 为初始裂纹尺寸；a_c 为临界裂纹尺寸；ΔK 为应力强度因子幅度，一般可表达为 $\Delta K = F\Delta\sigma\sqrt{\pi a}$，其中 F 为裂纹几何形状因子。在很多情况下，F 与裂纹尺寸 a 有关。在 F 与裂纹尺寸 a 无关或可近似地看作常数的情况下，有如下的疲劳裂纹扩展寿命计算公式

$$N = \frac{2}{(n-2)C(F\Delta\sigma\sqrt{\pi})^n}\left(a_0^{1-n/2} - a_c^{1-n/2}\right) \quad (n \neq 2) \tag{6.3-39}$$

$$N = \frac{1}{C\ (F\Delta\sigma\sqrt{\pi})^2}\ln\frac{a_c}{a_0} \quad (n=2) \tag{6.3-40}$$

由 Forman 公式，则有

$$N = \frac{2(1-R)K_C}{(n-2)C(F\Delta\sigma\sqrt{\pi})^n}\left(a_0^{1-n/2} - a_c^{1-n/2}\right) \quad (n \neq 2) \tag{6.3-41}$$

$$N = \frac{(1-R)\ K_C}{C\ (F\Delta\sigma\sqrt{\pi})^2}\ln\frac{a_c}{a_0} - \frac{2}{C\Delta\sigma\sqrt{\pi}}\left(\sqrt{a_c} - \sqrt{a_0}\right) \quad (n=2) \tag{6.3-42}$$

3.4 变幅疲劳载荷下的裂纹扩展寿命计算

在变幅疲劳载荷下计算裂纹扩展寿命的最简单的方法是忽略不同幅度的载荷之间的相互作用，应用线性累积损伤原理，分别计算各载荷幅度下的裂纹扩展量，然后相加得到裂纹在谱载荷下的总的扩展量。以 Paris 公式为例，构件在载荷 σ_i 下经历的载荷循环数为 N_i，在经受载荷 σ_i 开始时的裂纹尺寸为 a_{i-l}。在经历了 N_i 次的应力水平为 σ_i 循环载荷后裂纹尺寸 a_i 可由下式计算：

$$a_i = \left[a_{i-l}^{1-n/2} - \frac{N_i}{2}\ (n-2)\ C\ (F\Delta\sigma_i\sqrt{\pi})^n\right]^{\frac{2}{2-n}} \quad (n \neq 2) \tag{6.3-43}$$

$$a_i = a_{i-l}\exp\left[N_i C\ (F\Delta\sigma_i\sqrt{\pi})^2\right] \quad (n=2) \tag{6.3-44}$$

由上式即可算出构件经历了 $N = \sum_i N_i$ 次载荷循环之后的裂纹尺寸，或由临界裂纹尺寸计算寿命。

需要说明的是，在计算中不能把各种幅度的载荷循环的次序随意调换。

对于稳定的随机载荷，疲劳裂纹扩展速率可近似地用应力强度因子范围的均方根值 ΔK_{rms} 来描述：

$$\Delta K_{\mathrm{rms}} = F\Delta\sigma_{\mathrm{rms}}\sqrt{\pi a} \tag{6.3-45}$$

$$\frac{\mathrm{d}a}{\mathrm{d}N} = C\ (\Delta K_{\mathrm{rms}})^n \tag{6.3-46}$$

式中，σ_{rms} 为应力幅度的均方根值。

3.5 应力腐蚀裂纹扩展寿命计算

应力腐蚀裂纹扩展速率用 $\mathrm{d}a/\mathrm{d}t$ 表示，意指单位加载时间内的裂纹扩展量。试验表明 $\mathrm{d}a/\mathrm{d}t$ 与应力强度因子 K 的关系如图 6.3-15 所示。当 K 大于应力腐蚀断裂韧度 K_{ISCC} 后，裂纹开始扩展，且曲线分三个阶段。

第Ⅰ阶段，$\mathrm{d}a/\mathrm{d}t$ 随 K 的增加急剧增大

$$[\mathrm{d}a/\mathrm{d}t]_{\mathrm{I}} = C_1 + C_2 K \tag{6.3-47}$$

第Ⅱ阶段，$\mathrm{d}a/\mathrm{d}t$ 不随 K 变化，但受介质温度、压力等因素的影响

$$[\mathrm{d}a/\mathrm{d}t]_{\mathrm{II}} = C_3 \tag{6.3-48}$$

第Ⅲ阶段，$\mathrm{d}a/\mathrm{d}t$ 又开始随 K 的增加而增大

$$[\mathrm{d}a/\mathrm{d}t]_{\mathrm{III}} = C_4 K^m \tag{6.3-49}$$

式中，C_1、C_2、C_3、C_4 和 m 均为材料常数。

对式（6.3-47）积分可得第Ⅰ阶段裂纹扩展寿命为

$$t_{\mathrm{I}} = \int_{a_0}^{a_p} \frac{1}{C_1 + C_2 K}\mathrm{d}a \tag{6.3-50}$$

式中，a_p 为与裂纹扩展第Ⅰ阶段与第Ⅱ阶段分界点的裂纹长度。

第Ⅱ阶段的裂纹扩展寿命可由下式计算：

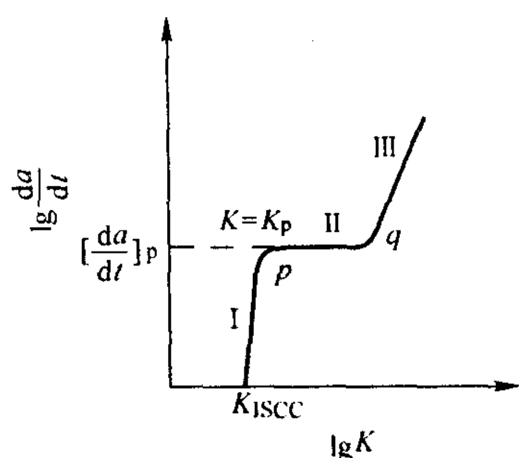


图 6.3-15 应力腐蚀裂纹扩展速率曲线

$$t_{II} = \int_{a_p}^{a_q} C_3 da = \int_{a_p}^{a_q} \frac{1}{D_3 p_l^{n_l} e^{-\Delta H/R_B T_K}} da \quad (6.3\text{-}51)$$

式中，ΔH 为材料激活能；R_B 为 Boltzman 常数；T_k 为热力学温度；p_l 为气体压力；D_3、n_l 为材料常数；a_q 为第Ⅱ阶段结束时的裂纹长度。

第Ⅲ阶段寿命较短，一般在设计时可以忽略。所以，应力腐蚀裂纹扩展的总寿命为

$$t_s = t_I + t_{II} \quad (6.3\text{-}52)$$

3.6 腐蚀疲劳裂纹扩展寿命

腐蚀疲劳裂纹的扩展速率分三种类型，如图 6.3-16 所示。

类型 a（见图 6.3-16a）。与疲劳裂纹扩展速率的规律相同，只是常数 C 变大了，$(da/dN)_{CF} > da/dN$。

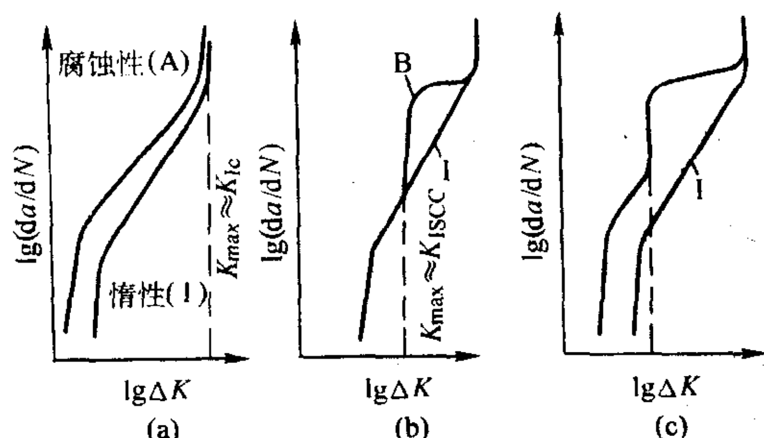


图 6.3-16 腐蚀疲劳裂纹扩展速率

类型 b（见图 6.3-16b）。当 $K > K_{ISCC}$ 时介质影响很大，裂纹扩展速率出现一个跳跃，然后 $(da/dN)_{CF}$—K 曲线呈现一个水平台阶。

类型 c（见图 6.3-16c）。大多数材料有此种类型的裂纹扩展速率曲线，它是 a 与 b 的混合型。

腐蚀疲劳的裂纹扩展速率可以分类分阶段处理，也可采用叠加模型

$$(da/dN)_{CF} = (da/dN)_{scc} + (da/dN)_F \quad (6.3\text{-}53)$$

或

$$(da/dt)_{CF} = (da/dt)_{scc} + f \times (da/dN)_F \quad (6.3\text{-}54)$$

式中，f 为频率。

将上式积分即可求出腐蚀疲劳裂纹扩展寿命。

3.7 剩余寿命计算示例

确定了初始裂纹尺寸 a_0、临界裂纹尺寸 a_c、相应的应力强度因子幅度表达式和材料的疲劳裂纹扩展速率表达式之后，就可进行剩余寿命计算。

初始裂纹尺寸 a_0 可按下述方法确定：

1）用无损检测方法测出的最大缺陷尺寸。

2）当用无损检测方法未检测出缺陷时，取初始缺陷尺寸等于所用检测方法的灵敏度。对于超声波探伤，这时一般可取 $a_0 = 2$ mm。

初始裂纹尺寸对零件的裂纹扩展寿命有显著影响，因而应尽可能准确地确定 a_0 值。

确定临界裂纹尺寸 a_c 的一般考虑是，对于低韧性材料，a_c 应由线弹性断裂判据来确定；而对于高韧性材料，a_c 则应由 COD 或 J 积分判据以及按净截面计算的静强度条件来确定。

在线弹性断裂范围内，临界裂纹尺寸 a_c 可根据以下原则确定。

1）零件的净截面应力应小于或等于强度极限 σ_b（当用静载荷计算净截面应力时，计算出的值应乘以 1.15 的动载系数）。

2）零件的应力强度因子幅度 ΔK 应小于裂纹快速扩展区起始点的对应值。没有这一数据时，也可用 K_{IC} 的下限值代替Ⅲ区起点的应力强度因子幅度值。

上述两条件中较小者即为临界裂纹尺寸 a_c。

在裂纹扩展寿命计算的过程中，对于几何形状因子 F 与裂纹尺寸 a 有关的问题，其疲劳裂纹扩展寿命的计算分为以下两种不同类型。一种是可以将 F 表示成 a 的函数，此时可直接进行积分运算以求得寿命 N。只是在积分中 F 已不再是常数，而是 a 的函数。另一种类型是 F 与 a 的关系比较复杂，不能表示成某一简单的函数。在这种情况下，须把整个疲劳裂纹扩展过程分解成多个阶段，而总寿命为各阶段寿命之和。对于每一个小阶段来说，由于裂纹长度 a 的变化很小，在计算中可以近似地把 F 做为常数。

例 如图 6.3-17 所示为含有一条径向表面裂纹的轴类零件，轴的材料为调质 40Cr，材料的断裂韧度 $K_{IC} = 1\,960$ N·mm$^{-3/2}$，裂纹长度 $2a_c = 20$ mm，裂纹深度 $l = 3$ mm，裂纹处的弯曲应力为 $\sigma = \pm 100$ MPa，试估算其疲劳寿命。

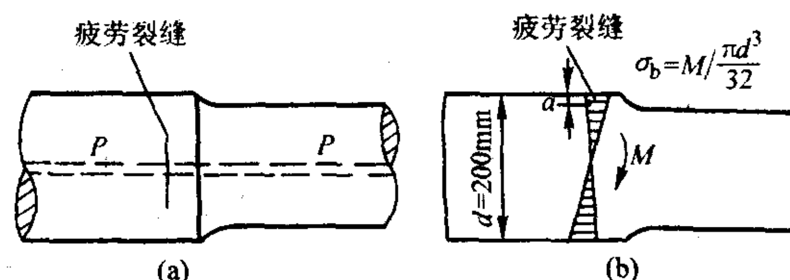


图 6.3-17 带裂纹的轴及简化形式

1）确定应力强度因子幅度 ΔK 表达式。在裂纹深度远小于裂纹长度和轴径尺寸的条件下，可以将此含裂纹体简化为处于平面应变状态下具有单边穿透裂纹的受循环弯曲载荷作用的平板（如图 6.3-17b 所示），裂纹长度 $a = 1$。由于裂纹深度比板的宽度小得多，可以用具有单边穿透裂纹的半无限大受拉平板的 K_I 的表达式作为该情况下应力强度因子的近似表达式，即

$$K_I = 1.12\sigma\sqrt{\pi a}$$

2）确定裂纹扩展速率表达式及有关参数。该轴材料为 40Cr 低合金钢，应力比为 −1 时裂纹扩展速率公式（Paris 公式）参数为 $C = 8.16 \times 10^{-14}$，$n = 3.3$。

3）已知初始裂纹尺寸 $a_0 = 3.0$ mm

4）确定临界裂纹尺寸 a_c。由已知条件 $K_{Ic} = 1\,960$ N·mm$^{-3/2}$、K_{Ic} 的表达式 $K_{Ic} = 1.12\sigma\sqrt{\pi a_c}$ 及载荷水平 $\sigma_{max} = 100$ MPa 可得：

$$a_c = \left(\frac{1\,960}{1.12 \times 100}\right)^2 \frac{1}{\pi} = 98 \text{ mm}$$

5）计算裂纹扩展寿命。由于前面裂纹扩展速率参数取的是对称循环载荷下的对应值，在下面的裂纹扩展寿命公式中 $\Delta\sigma$ 应取为 σ_{max}，所以有

$$N=\frac{2}{(n-2)\ C\ (F\Delta\sigma\sqrt{\pi})^n}\ (a_0^{\ 1-n/2}-a_c^{\ 1-n/2})$$
$$=2.16\times10^5$$

4　断裂韧度设计结构类型及设计要求

4.1　断裂韧度设计结构类型的选择原则

按照断裂韧度要求设计的结构可分为缓慢裂纹扩展结构和破损安全结构两种类型。破损安全结构又可以进一步区分为破损安全多途径传力结构和破损安全止裂结构。无止裂特性的单传力途径结构应规定为缓慢裂纹扩展结构；多途径传力和有止裂特性的结构或者规定为缓慢裂纹扩展结构，或者在指定的可检查度下规定为破损安全结构。

缓慢裂纹扩展概念是指结构中的缺陷或裂纹以稳定、缓慢的速度扩展，在预定的使用期内不发生不稳定快速扩展。对于缓慢裂纹扩展结构，通过对设计应力水平的控制和材料的选择，使裂纹保持低速扩展，在使用中裂纹不允许达到不稳定扩展的临界尺寸，在有效使用期内由裂纹缓慢扩展保证安全。

破损安全概念是指采用多途径传力或止裂后，使不稳定裂纹扩展限制在局部范围内。破损安全多途径传力结构具有多条传力途径，它是把一些本来可以设计成一体的结构人为地分成若干部分，使损伤被控制在局部范围内，以防止结构完全破坏。这类结构在主传力结构损坏后，其剩余结构在后续检查之前，由剩余结构的裂纹缓慢扩展来保证安全，在有效使用期内不允许结构强度和安全性下降到规定水平以下。破损安全止裂结构通常由多个元件组成。在结构完全破坏之前，不稳定的快速裂纹扩展能被抑制在事先设计的止裂区内。

结构类型由设计概念和可检查度决定。可检查度由结构所处的部位、可达性和可检性、损伤的性质和程度以及检查经验来确定。设计概念的选择原则主要取决于结构布局和几何构型。对于单传力途径、整体结构和静定结构，必须设计成缓慢裂纹扩展结构；而对于具有多条传力途径的结构和具有止裂能力的结构，既可设计成破损安全结构，亦可设计成缓慢裂纹扩展结构。设计成缓慢裂纹扩展结构的理由，一是难以满足破损安全结构的全部要求；二是将结构设计成破损安全是一个需要判断、分析和计算的复杂过程，而设计成缓慢裂纹扩展结构则减少了分析的复杂性，工作量少得多，而且一般偏于安全。在设计的初期阶段，往往将这类结构都设计成缓慢裂纹扩展结构。

使用中不可检的结构，例如某些内部结构只能设计成缓慢裂纹扩展结构。虽然破损安全结构的主传力途径可按使用中不可检结构设计，但是主传力途径破坏后的损伤（或剩余结构）必须是可检的。

对于具有多条传力途径和止裂能力的结构，欲设计成破损安全结构，必须是当主传力途径破坏或快速失稳扩展裂纹止裂后结构便于修复或更换。否则，结构应设计成缓慢裂纹扩展结构。

4.1.1　缓慢裂纹扩展结构

缓慢裂纹扩展结构所采用的材料和结构形式都应使结构具有较低的裂纹扩展速率和较高的剩余强度。图6.3-18所示为结构在疲劳载荷作用下裂纹长度及剩余强度随时间而变化的曲线。从图中可以看出，结构经一定使用期限后，在时间 t_0 出现裂纹。但起初裂纹尺寸很小，用现有的几何检查技术尚不能发现。再经过一段时间（如 A 点），裂纹发展到可检尺寸，此时的裂纹长度称为最小可检长度。由于裂纹的扩展，结构的剩余强度也在随着时间而降低。如果剩余强度仍大于所需要的最小强度——破损安全强度（根据设计规范确定），结构仍是安全的。当裂纹扩展到 B 点时，结构的剩余强度恰等于破损安全强度，此时裂纹长度达到临界裂纹长度。从 t_A 到 t_B 这段时间内，虽然存在裂纹，但结构是安全的。这段时间可供检测以发现裂纹。为了安全运行，在这一期间内应作多次检查，因为在一次检查中，结构的可检裂纹可能恰好漏检。因此，一般在此期间安排2～3次检查。由图可知，在裂纹较小时，有相当长的一段时间裂纹扩展速率较低。只有当裂纹较长时，裂纹扩展速率才很高。因此稍微增加一点容许的裂纹长度是没有意义的，因为裂纹扩展的最后部分只占很短的时间。将可检裂纹尺寸减小，可大大延长检查周期和提高安全可靠性。故提高检测水平是关键问题。采用缓慢裂纹扩展结构，必须使结构可检查性好，如对于裂纹发生在不易检查部位的结构，或者临界裂纹长度很小的结构，采用这种方法将会遇到困难。

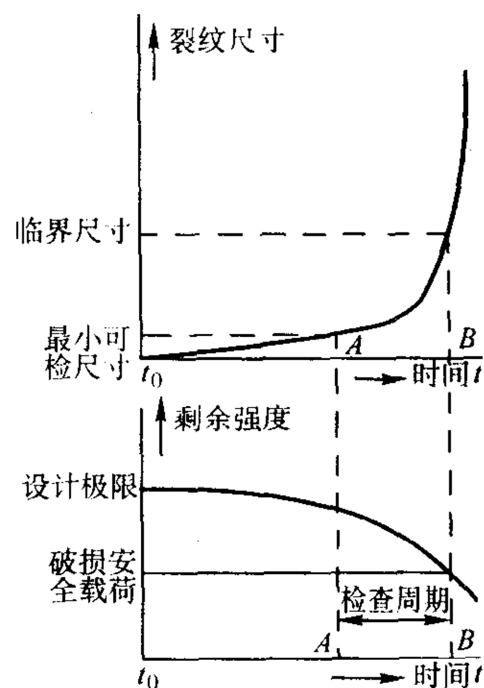


图6.3-18　由断裂韧度概念保证结构破损安全图示

4.1.2　多传力途径－破损安全结构

多传力途径－破损安全结构是由多个元件或分段组成多条传力途径的结构。结构分段可控制局部损伤，而防止结构完全破坏。结构在规定的未修理使用期内，强度不会降低到规定水平以下。该结构通过残存结构在后续检查前缓慢裂纹扩展保证安全。当结构采用两个或两个以上元件来传递载荷时，其中一个元件由于某种原因提前破坏，剩余的结构应能承受所要求的载荷。已破坏元件上的载荷可通过其他元件传递，而结构不失效。如图6.3-19所示，如果经过 t_C 小时后一个元件破坏（C 点），剩余强度仍大于规定的破损安全载荷。但由于第一个元件的破坏，其他元件上的载荷增加，第二个元件可能在短时间后（D 点）破坏，强度降低到安全值以下。从 t_C 到 t_D 之间的时间可作为发现元件破坏的时间。若第一个元件是由于初始损伤或意外损伤而引起的过早破坏，CD 区间可能相当长。若破坏不是特殊情况引起的，则 CD 区间的长度与各元件寿命的分散性有关。

4.1.3　破损安全止裂结构

破损安全止裂结构是在完全破坏前使裂纹不稳定快速扩展停止在结构的某一连续区域内的结构。残存的未损伤结构的强度在未修使用周期内不下降到规定限度以下。该结构通过残存结构中的缓慢裂纹扩展和在后续检查中发现裂纹来保证安全。此类结构一般采用多个元件组成，适当布置止裂带，如图6.3-20所示。

4.2　危险部位的选择

结构零件的破坏往往都是由危险部位的破坏引起的。所以，危险部位的选择是断裂韧度设计和评定工作中的最重要环节之一。

选择断裂韧度危险部位时须着重考虑的因素包括：

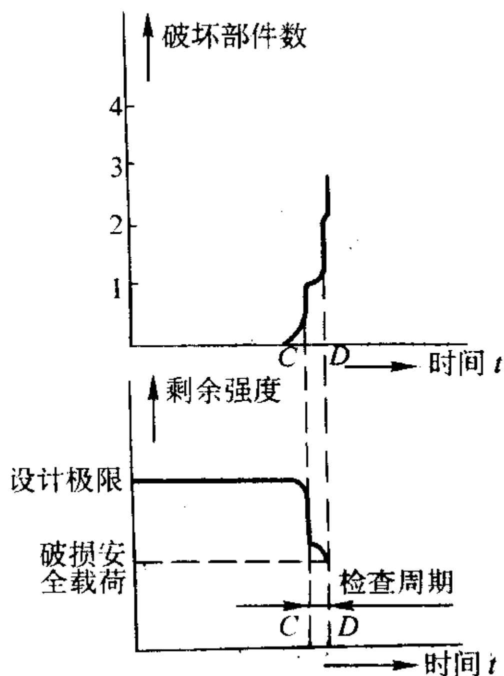


图 6.3-19　多载荷通道保证破损安全图示

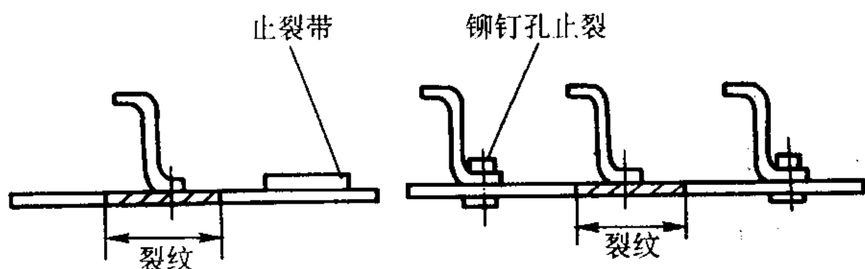


图 6.3-20　破损安全止裂结构

1) 整体结构和单传力结构，特别是不可检结构部位；
2) 应力水平高、应力集中严重的部位；
3) 采用缺口敏感材料的结构零件；
4) 承受高温或腐蚀作用的部位；
5) 加工质量不易保证的结构部位；
6) 刚度不足振动严重的部位。

危险部位的确定一般可借助应力分析确定。在分析中应特别注意诸如刃边、圆角以及圆孔等高应力集中部位。危险部位确定后，假定初始裂纹，计算裂纹扩展和剩余强度，判断是否满足设计要求。

4.3　实际缺陷与当量裂纹

在断裂韧度设计中，把机械结构在制造过程中产生的各种缺陷，例如各种形式的焊接缺陷、锻造缺陷、铸造缺陷等，都理想化为等效裂纹来处理。无损探伤技术所发现的缺陷往往都是很复杂的。而在计算应力强度因子时，需要用到裂纹的形状、尺寸及位置等参数。从应力强度因子的表达式可以看出，合理的确定裂纹参数对应力强度因子的计算以及结构寿命的预测非常重要。

从形状来看，探伤技术所能确定的缺陷有平面缺陷和非平面缺陷。平面缺陷的代表为热处理裂纹，非平面缺陷的代表为铸件内的气孔、非金属夹杂及疏松等。一般来说，当不能确定缺陷是平面或非平面时，或者已确定是非平面缺陷，但其某一方向的尺寸远小于另两个方向的尺寸时，应偏于安全地将其当做平面缺陷来处理。

根据缺陷在构件上的位置，可以分成表面裂纹、深埋裂纹和贯穿裂纹三种（如图 6.3-21 所示）。通常把深埋缺陷理想化为圆片状或椭圆片状裂纹。圆片的半径或椭圆的短半径作为裂纹的代表性尺寸代入应力强度因子的计算式中。表面缺陷一般理想化为半椭圆片状裂纹，也是以椭圆的短半径作为裂纹的特征尺寸。贯穿裂纹的特征尺寸则为图中的裂纹长度 $2a$。

对于图 6.3-22 所示的密集型缺陷，如果两个表面缺陷 A 和 B 相距很近，则应并做一个半椭圆片状裂纹来对待。例如，当两裂纹之间的距离 s 满足下面关系：

$$s < c_1 + c_2$$

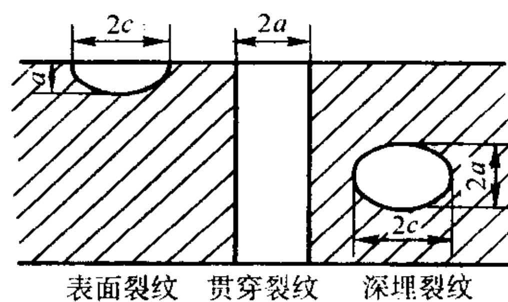


图 6.3-21　不同位置的裂纹

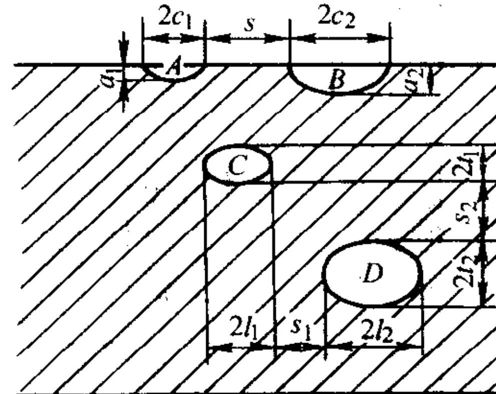


图 6.3-22　密集型缺陷

则可当作一个深为 a_2、长为$2(c_1+c_2)+s$ 的表面裂纹。

对于两个深埋裂纹 C 和 D，如果

$$s_1 < l_1 + l_2$$
$$s_2 < t_1 + t_2$$

则可当作一个深为 $2t=2(t_1+t_2)+s_2$，长为$2l=2(l_1+l_2)+s_1$ 的深埋椭圆裂纹。如果有多个难于确定其精确尺寸及间距的密集裂纹，也可以用包围此缺陷群的圆或椭圆作为当量裂纹的尺寸。

用超声波方法确定裂纹尺寸时，所得到的并不是裂纹的尺寸，而是一个当量尺寸。一般情况下，实际裂纹尺寸约为当量尺寸的 2～3 倍。

很浅而又较长的表面裂纹比同样尺寸的深埋裂纹危险得多。因为很浅，探伤检测不易发现。这时，可假设存在有一个半椭圆形表面裂纹，其形状为 $a/2c=0.3$，裂纹深度 a 按探伤技术可能达到的最小界限来确定。

对一些缺陷具体的裂纹当量化方法如下。

4.3.1　单个缺陷的当量裂纹

设计时，为了简化计算和分析，常保守地将缺陷当量处理成与最大主应力方向垂直的平面上的裂纹，即 I 型裂纹。这时，内部缺陷当量处理成圆型或椭圆型裂纹，表面缺陷则当量处理成半圆型或半椭圆型裂纹，如图 6.3-23 和图 6.3-24 所示。当量裂纹的特征尺寸为裂纹长度 l 和裂纹深度 a。

4.3.2　多个缺陷的当量裂纹

多个缺陷相邻时，相邻缺陷的应力场存在相互作用，称之为相邻缺陷的干涉。当相邻缺陷的间距小于等于缺陷深度时，干涉系数将大于 15%。设计时一般把相互干涉作用较强（例如干涉系数 > 15%）的两个或两个以上的缺陷当量处理成一个包含此缺陷群的一个较大的单个裂纹来分析（见图 6.3-25）。另外，由于共面缺陷聚集成一个裂纹的倾向比非共面缺陷更危险，在实际应用中往往把非共面密集缺陷也按共面缺陷来处理。

4.3.3　层状缺陷的处理

如果缺陷显示的长度和厚度所在的平面取向与构件表面的交角在 10°以内，则初见此种缺陷为层状缺陷，其当量裂纹尺寸如图 6.3-26 所示。

4.4　断裂判据

4.4.1　应力强度因子判据（一）：线弹性断裂－脆性断裂

根据线弹性断裂力学，当构件的裂纹应力强度因子达到某一临界值时，裂纹就开始不稳定快速扩展，导致构件突然断裂。平面应变条件下 I 型裂纹的断裂判据为

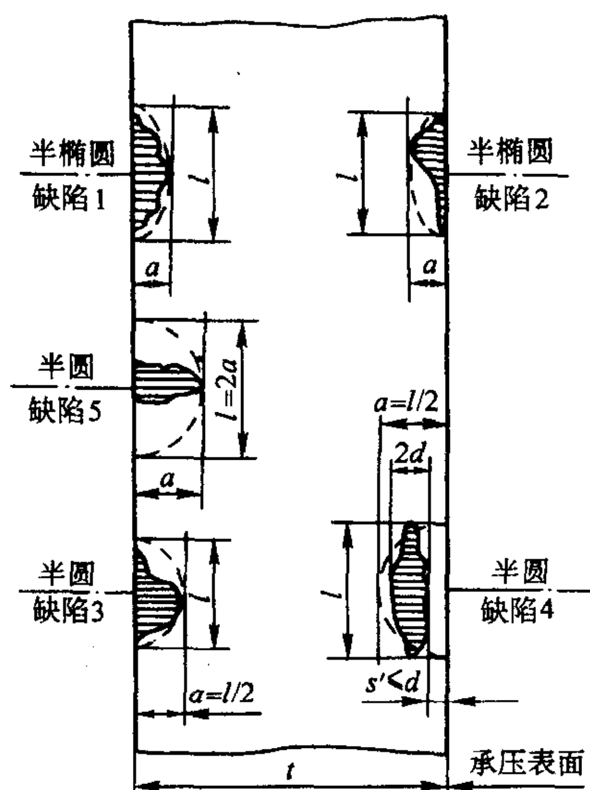


图 6.3-23 当量裂纹尺寸

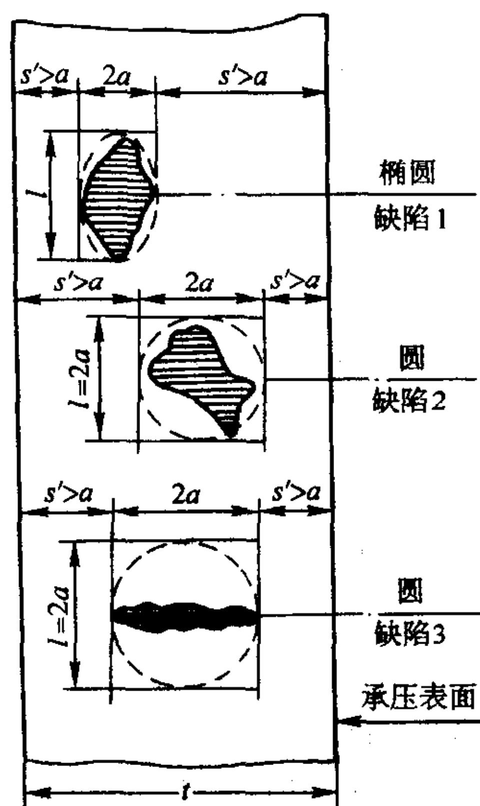


图 6.3-24 当量裂纹尺寸

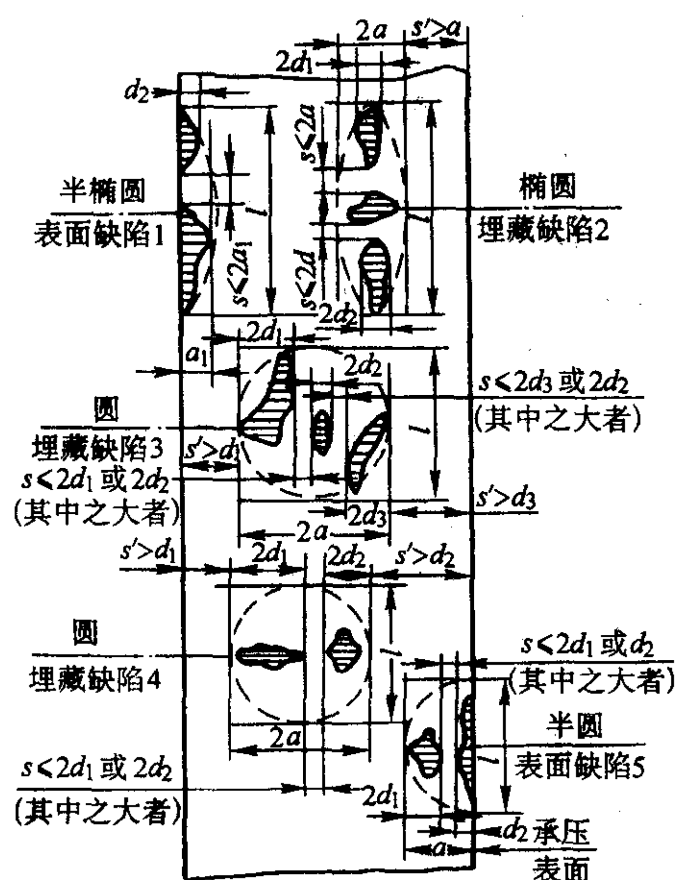


图 6.3-25 当量裂纹尺寸

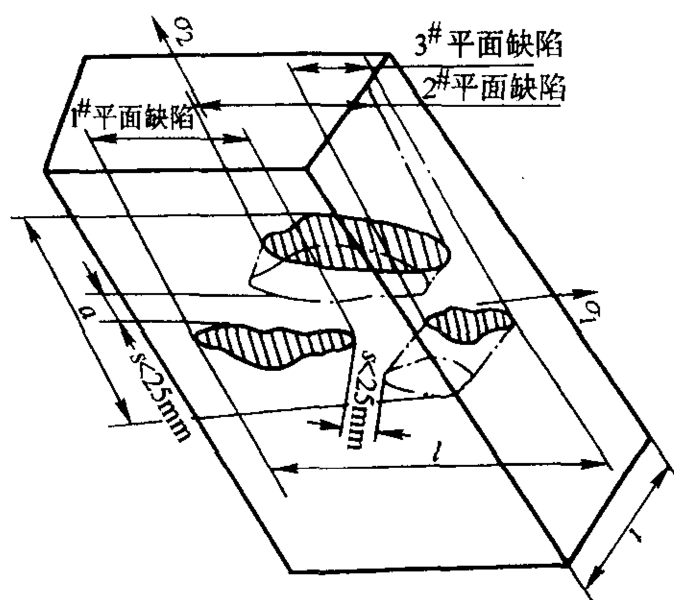


图 6.3-26 当量裂纹尺寸

$$K \geqslant K_{\mathrm{Ic}} \tag{6.3-55}$$

式中，K 为裂纹尖端应力强度因子；K_{Ic} 为材料的平面应变断裂韧度。

根据应力强度因子表达式 $K = F\sigma\sqrt{\pi a}$，可将上述断裂判据写成另外两种形式：

1）给定 σ 时，$a \geqslant a_c = \dfrac{1}{\pi}\left(\dfrac{K_{\mathrm{Ic}}}{F\sigma}\right)^2$

2）给定 a 时，$\sigma \geqslant \sigma_c = \dfrac{K_{\mathrm{Ic}}}{F\sqrt{\pi a}}$

4.4.2 应力强度因子判据(二)：线弹性断裂－平面应力状态断裂

实验表明，薄板或韧性好的材料表现出一些不同于脆性断裂的剩余强度特性，如图 6.3-27 所示。对于长度为 $2a_i$ 的初始裂纹，应力增大到 σ_i 后，随着应力的进一步增加，裂纹长度也有所增长，但不会立刻失稳。直到应力增大到 σ_c，相应的裂纹长度也增长到 $2a_c$ 时，裂纹才发生不稳定扩展。

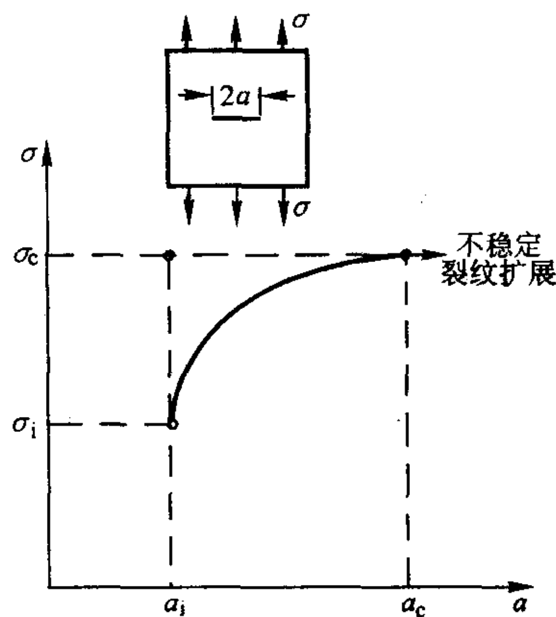


图 6.3-27 平面应力状态下裂纹扩展特性

如果对不同的初始裂纹长度重复同样的试验，则可得到如图 6.3-28 的剩余强度图。图中有 3 条不同的剩余强度曲线：最下面一条代表缓慢裂纹扩展起始的剩余强度曲线，记为 K_i，表示裂纹稳定扩展的起始点；最上面一条表示裂纹失稳扩展时的剩余强度曲线，记为 K_c，表明发生失稳扩展时的临界裂纹长度 $2a_c$ 与对应的裂纹失稳扩展应力 σ_c 的关系。由于 K_c 曲线不能预测与初始裂纹 $2a_i$ 对应的临界应力水平，a_c 又很难实测，所以工程上用容易实测的初始裂纹长度 a_i 和临界应力 σ_c 的组合来得出一条表观剩余强度曲线，记为 K_{app}，即图 6.3-28 中间的那条曲线。

相应地，破坏判据为

$$K \geqslant K_{\mathrm{app}} \tag{6.3-56}$$

式中，表观断裂韧度定义为

$$K_{\mathrm{app}} = F\sigma_c\sqrt{\pi a_i} \tag{6.3-57}$$

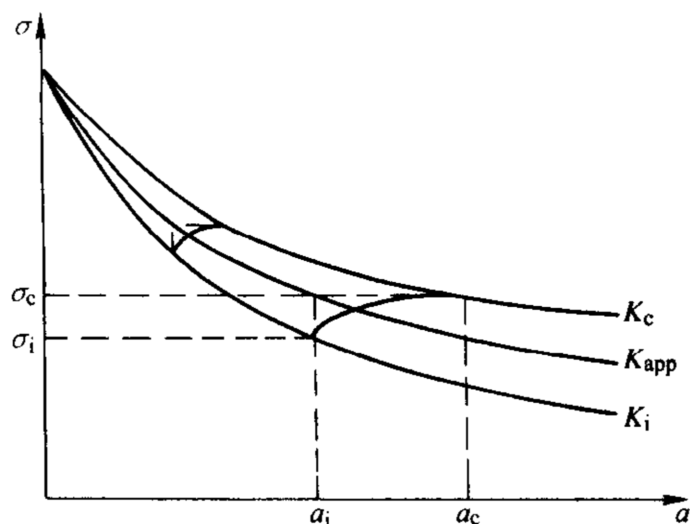


图 6.3-28 平面应力状态断裂剩余强度特性

4.4.3 净截面屈服判据

对于高韧性材料，构件上的应力会高到使整个净截面在断裂发生前先产生屈服，最后导致构件破坏。对于这种净截面屈服破坏，在工程上可以直接用截面上的净应力 σ_j 与材料屈服强度 σ_s 的关系建立破坏判据，即

1）给定裂纹长度时

$$\sigma_j \geqslant \sigma_s \tag{6.3-58}$$

2）给定外载 σ 时

$$a \geqslant a_{jq} \tag{6.3-59}$$

即当裂纹长度 a 达到或超过净截面屈服时的临界裂纹长度 a_{jq} 时，构件发生破坏。

图 6.3-29 给出了几种典型构件裂纹形式及加载方式下引起净截面屈服的远方应力 σ_{jq} 的计算公式，图 6.3-30 给出了净截面屈服时的临界裂纹长度 a_{jq} 的计算公式。图中的 σ_{xy} 为要求的最小剩余强度，a_{jq} 为 σ_{xy} 下净截面屈服允许的最大裂纹尺寸。

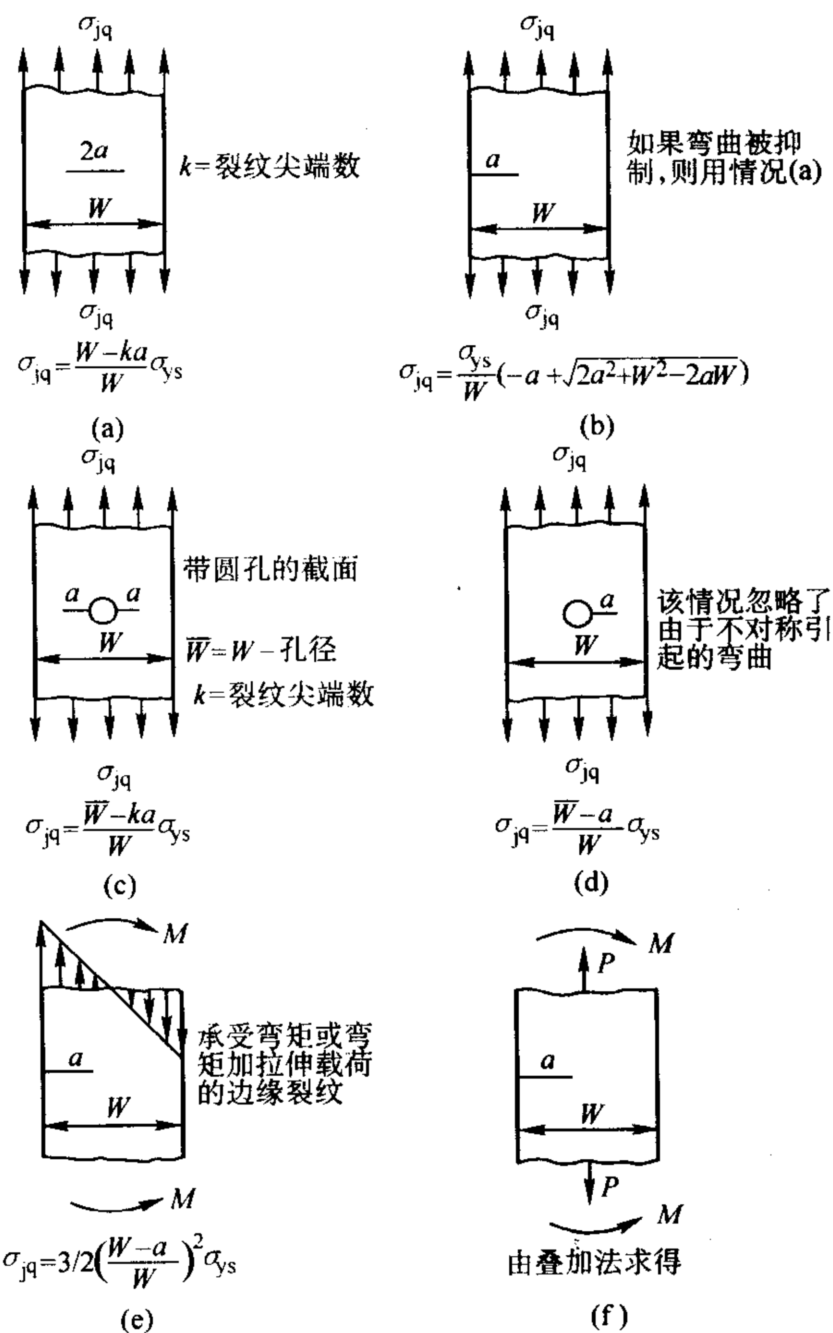


图 6.3-29 引起净截面屈服的远方应力 σ_{jq} 计算公式

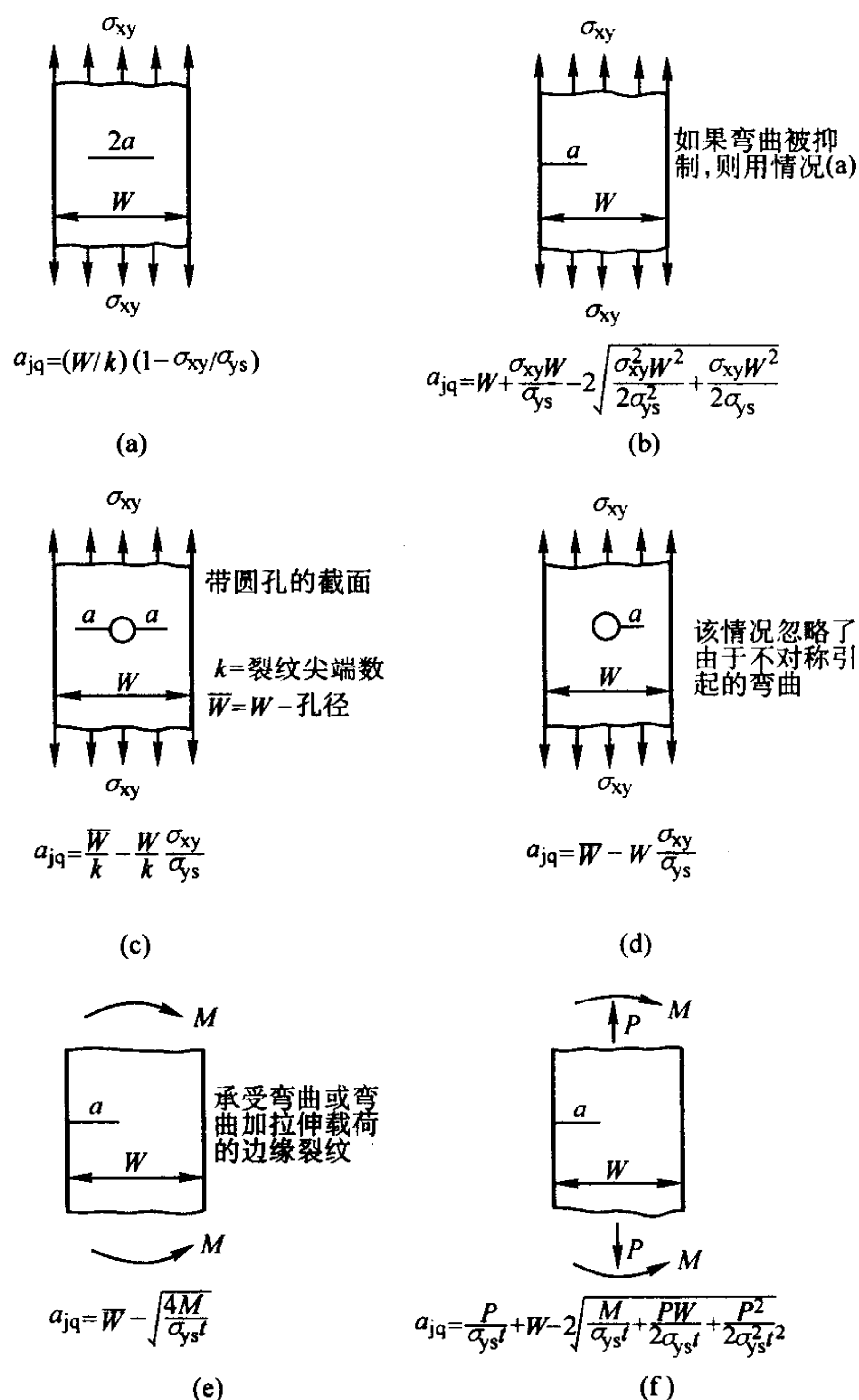


图 6.3-30 净截面屈服临界裂纹长度计算公式

4.4.4 弹塑性断裂

在线弹性断裂和净截面屈服之间，还存在另一种断裂形式。即净截面屈服尚未发生，但裂纹尖端的塑性区已大到不能忽略。这种情况属于弹塑性断裂问题。为了简化分析，工程上可以采用近似方法，下面介绍一种常用的工程近似方法——切线法。

切线法是以实验观察为基础的。对含中心穿透裂纹的无限大板，根据平面应力状态下的线弹性断裂准则可得到临界应力 σ_{cr} 与裂纹长度 $2a$ 之间的关系，如图 6.3-31 所示。σ_{cr} 和 $2a$ 呈类似双曲线关系。

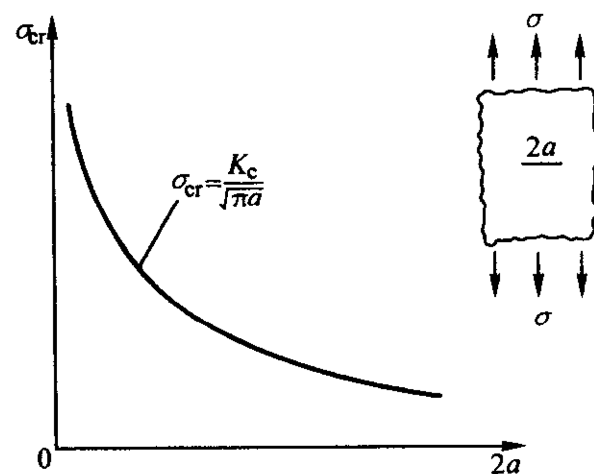


图 6.3-31 按线弹性断裂得出的剩余强度特性

然而，实际材料的强度是有限的，可能的裂纹尺寸也是有限的。有限大板的弹性不稳定断裂实验数据分布表明，中间区域与线弹性断裂情况基本一致。而在左边高应力区，由于有材料屈服极限的限制，实验数据与线弹性断裂情况有所

偏离；在右边长裂纹区，由于有限宽度的边界影响也引起偏离。有限大板的弹性不稳定断裂实验数据如图 6.3-32 所示，由图可见实验数据在高应力区和长裂纹区呈线性关系。

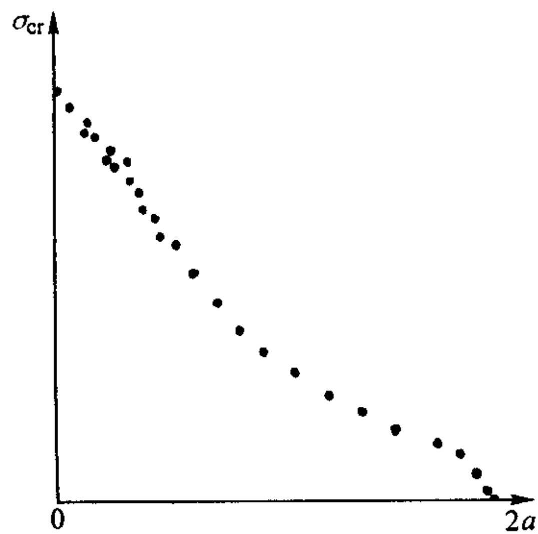


图 6.3-32 有限大板的弹性不稳定断裂实验数据示意图

基于以上考虑，可以认为实际的剩余强度曲线如图 6.3-33 所示。图中，从纵坐标轴上的 A（0，σ_s）点向理想弹性断裂曲线所作的切线表示弹塑性断裂线。在应力 σ 作用下按切线法计算出的临界裂纹长度为

$$2a_{ts}=a_{tx}\left[\frac{27}{2}\left(1-\frac{\sigma}{\sigma_s}\right)\left(\frac{\sigma}{\sigma_s}\right)^2\right] \tag{6.3-60}$$

式中，σ_s 为材料的拉伸屈服强度；a_{tx} 为在 σ 作用下按线弹性脆性断裂剩余强度条件计算所得到的临界裂纹长度；a_{ts} 为在 σ 作用下按切线法计算所得到的临界裂纹长度。

当对一个具体问题进行剩余强度分析时，可以将线弹性脆性断裂的剩余强度曲线、净截面屈服的塑性断裂线和按切线近似的弹塑性断裂线绘制在同一坐标图上（如图 6.3-34 所示），以便判断应该选用哪条曲线作为剩余强度依据。

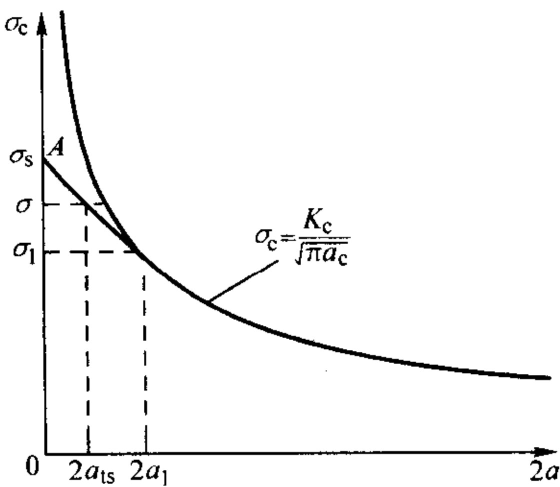


图 6.3-33 弹塑性断裂剩余强度曲线

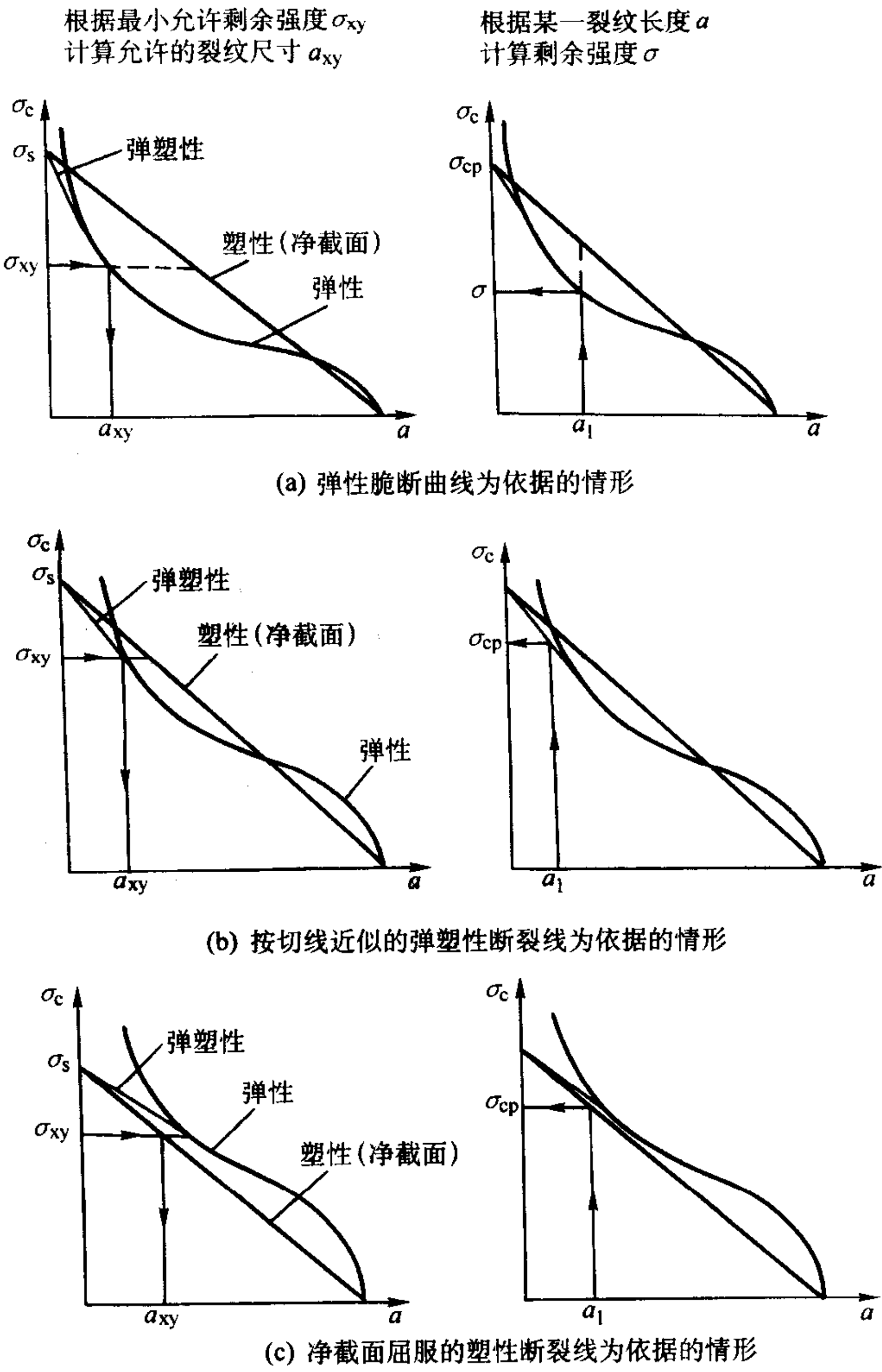


图 6.3-34 含裂纹板的不同破坏形式的剩余强度曲线

如果剩余强度要求值（即许用的最小剩余强度）σ_{xy} 等于或大于 $2\sigma_s/3$ 时，应选用弹塑性破坏线与净截面屈服破坏线中较小者作为剩余强度曲线。

如果剩余强度要求值 σ_{xy} 小于 $2\sigma_s/3$ 时，则应选用弹性破

坏曲线和净截面屈服线中较小者作为剩余强度曲线。

4.4.5　三维裂纹断裂

表面裂纹、角裂纹、孔边角裂纹、深埋裂纹等非穿透裂纹均属三维裂纹问题。三维裂纹在裂纹前缘的不同点上的应力强度因子是不同的。因此在计算剩余强度时，如果仍采用应力强度因子这个局部力学量作为断裂判据，就有一个选取哪一点的应力强度因子作判据的问题。

工程上常采用以下应力强度因子值作为构件的断裂判据。

1）采用最大应力强度因子 K_{max}　当 $K_{max} \geqslant K_{Ic}$ 时构件断裂，这是一种保守的做法。

2）采用沿裂纹前缘 45°方向处的应力强度因子 $K_{45°}$　当 $K_{45°} \geqslant K_{Ic}$ 时构件断裂。

3）采用平均应力强度因子　当 $\overline{K} \geqslant K_{Ic}$ 时构件断裂。

$$\overline{K} = \frac{K_{max} + K_{min}}{2}$$

以上判据中均采用材料的平面应变断裂韧度 K_{Ic}，而不是用平面应力断裂韧度 K_c。这是因为对于非穿透裂纹，即使板的厚度比较薄，裂纹尖端仍受到周围材料的很大约束。通常，假设所有的三维裂纹都是平面应变问题。

4.5　剩余强度分析

根据传统的材料强度理论，材料的拉伸强度就是不含裂纹构件的极限强度。但是，如果构件中存在裂纹，它的强度会明显低于不含裂纹结构的强度，且裂纹尺寸越大，其强度越低。断裂韧度设计要解决的其中一个问题就是确定含裂纹构件的承载能力，即含裂纹结构的剩余强度问题。

断裂韧度设计的基本目的是要确保结构零件在预期的整个使用寿命期内的安全。为此目的，必须按如下原则进行设计：在出现裂纹或者部分破坏的情况下，构件仍能承受可能的最大载荷。要达到这一设计要求，首要的原则是保持所要求的最小剩余强度以防止结构零件发生灾难性的破坏。由于构件材料特性、裂纹几何形状、应力状态等因素的不同，在外载荷作用下可能发生的失效形式有弹性断裂、弹塑性断裂或净截面屈服等。同时，不同的结构设计类型具有不同的剩余强度特性，而对不同的设计类型的剩余强度要求也不同。下面依次介绍不同设计类型的剩余强度要求。

1）对于裂纹扩展不可检结构，含裂纹构件在 2 倍设计寿命期内不允许由于裂纹扩展而导致结构剩余强度降低到可能遭遇到的最大载荷以下。对于可检的缓慢裂纹扩展结构，含裂纹构件在规定的最小未修使用期内（2 倍于检查间隔）不允许由于裂纹扩展而使结构剩余强度降低到可能遭遇到的最大载荷以下。

2）破损安全多途径传力结构的有效工作期分为两个阶段。第一阶段是主传力途径尚未失效，即完整结构阶段；第二阶段是主传力途径失效后的剩余结构阶段。完整结构在设计寿命内裂纹以缓慢的稳定形式扩展。在裂纹达到临界值以前，结构应能承受可能遭遇到的最大载荷。剩余结构除了能承受主传力途径破坏后的载荷外，还必须有能力承担主传力途径失效的瞬间转移过来的载荷，包括在传力途径破坏时相邻元件的载荷重新分配并考虑其动载荷效应。

3）破损安全止裂结构的基本思想是在结构中提供一种减小裂纹尖端应力强度因子的方法，以便能够限制或阻止裂纹不稳定扩展。设计上常常采用附加加强元件，诸如肋条等结构元件。对破损安全止裂结构剩余强度特性的要求是，这些加强元件应具有足够大的刚度，以便能够有效地卸掉基本结构上的部分载荷，从而使基本结构快速扩展的裂纹暂停在结构的某个连续区域内。

一般来说，在进行剩余强度分析以前，应当完成以下工作：

1）确定断裂危险部位；

2）明确剩余强度要求；

3）确定裂纹开裂顺序，并按开裂顺序绘制开裂结构简图；

4）计算对应开裂形式的应力强度因子；

5）选取有关的材料数据；

6）确定载荷谱和应力谱。

关于剩余强度分析的具体方法，较简单的问题可直接用线弹性断裂力学计算最大允许裂纹长度 a_{xy}，复杂问题则需绘制剩余强度曲线，做出完整的剩余强度图，这种方法适用于下述诸情况：

1）用于加肋板或者其他具有止裂作用的或多途径传力的任何结构的剩余强度分析。

2）在进行破损安全止裂结构或多途径传力结构设计时，绘制出各种设计方案的剩余强度图，可以更直观地进行方案比较，优化结构参数，选择最佳设计。

3）对于不很熟悉断裂力学和剩余强度分析的工程师来说，绘制全部剩余强度图有助于了解带裂纹构件剩余强度特性的概貌和总体特征。

4.6　工程实际问题的简化处理方法

虽然用剩余强度图一步一步地计算剩余强度这一方法并不复杂，但问题的复杂性可能会出现在构件本身的复杂性上。例如，构件具有复杂的几何形状、载荷形式、复杂的载荷传递等。对于这些情况，可以用有限元法解决。通常，用偏于保守的工程近似方法更为简洁、有效。下面针对几种常见的工程问题推荐其近似处理方法。

第一种情况是已断裂加强件附近的裂纹，如图 6.3-35 所示。

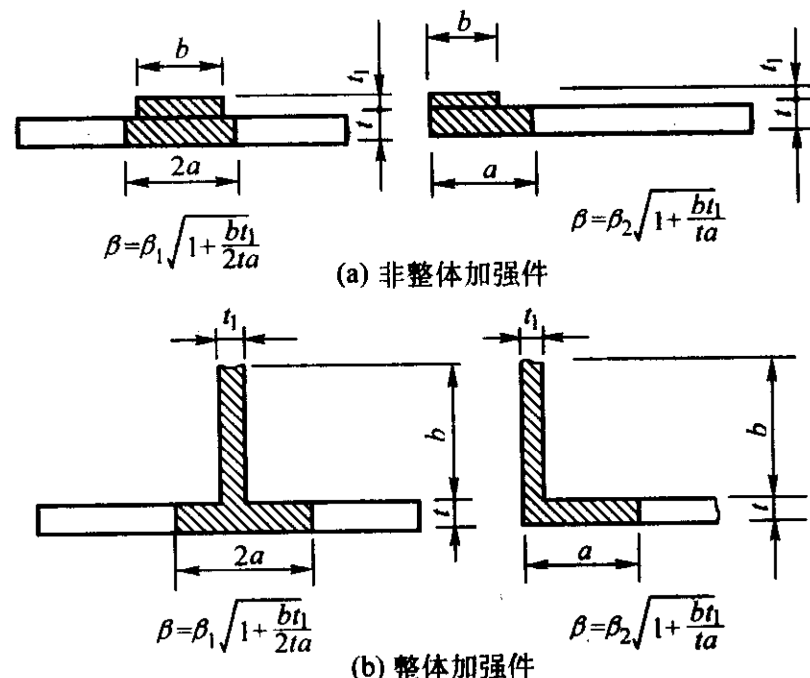


图 6.3-35　已断裂的加强件附近裂纹扩展及构件宽度

构件的当量宽度与当量裂纹长度按下式计算：

当量宽度　$W' = W + \dfrac{bt_1}{t}$

当量裂纹长度　$2a' = 2a + \dfrac{bt_1}{t}$

净截面屈服应力　$\sigma_{jq} = \dfrac{W' - 2a'}{W'}\sigma_s$

第二种情况是腹板开裂情况，如图 6.3-36 所示。

图中所示为工字梁腹板开裂情况。计算净截面积时，总面积必须包括缘条面积：

$$\sigma_{jq} = \frac{\text{总面积} - \text{裂纹面积}}{\text{总面积}}\sigma_s$$

如果总面积还包括孔，则须从总面积中将孔的面积减去。

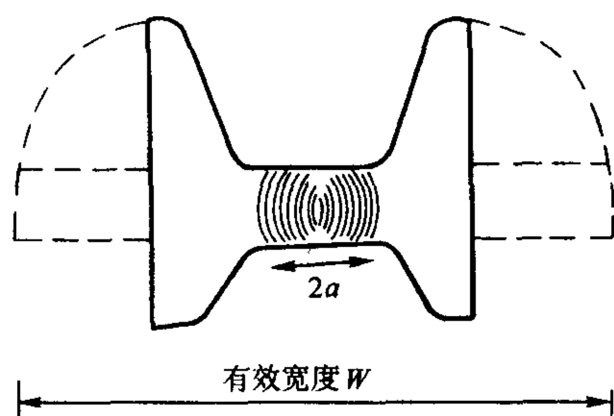


图 6.3-36 腹板开裂情况的剩余强度简化计算图

对于以上两种复杂几何形状的情况，按以下方法建立其剩余强度图。

首先根据实际裂纹尺寸及应力强度因子确定弹性断裂曲线；然后用前面介绍的公式计算净截面屈服应力，作出净截面屈服线；最后从净截面屈服线与纵坐标的交点向弹性断裂线的 $2\sigma_s/3$ 点作直线，构成弹性断裂的切线近似。以上 3 条曲线的最低包络线就是每一裂纹长度所对应的剩余强度。

第三种情况是变化的几何构形，如图 6.3-37 所示。

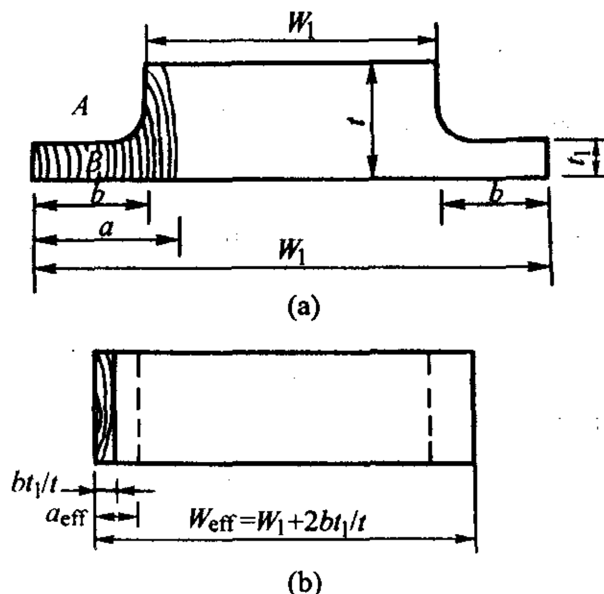


图 6.3-37 变化的几何构形

图 6.3-37a 所示的变化的几何构形情况也是很常见的问题之一，其剩余强度图按如下步骤绘制：

缘条开裂时，在裂纹扩展接近厚度变化处，首先须计算应力强度因子修正系数 F。当裂纹扩展到较厚的材料部分时，F 由规则的边缘裂纹解给出（参见图 6.3-37b。工程上常常用有效裂纹长度作为较厚缘条部分的计算裂纹长度。有效裂纹长度定义为

$$a_{eff}=\frac{bt_1}{t}+a-b$$

式中，a 为总裂纹长度；b 为薄缘条宽度。

用 a_{eff} 计算出较厚元件的 F 值后，就可以按下式计算较厚材料部分的弹性断裂剩余强度

$$\sigma_{tx}=\frac{K_c\ (\text{或}\ K_{Ic})}{\sqrt{\pi a}F}$$

尽管如此，在剩余强度图上绘制 σ_c 和 a 的关系曲线时，必须用实际的裂纹尺寸。因为实际裂纹尺寸与检查者有关，并且为检验者所见到的。

同时，当裂纹从一个材料厚度变到另一个材料厚度时，K_c 也要变化。

对于净截面屈服曲线，可以用下面的简化公式计算：

$$\sigma_{jq}=\frac{\text{总面积}-\text{裂纹面积}}{\text{总面积}}\sigma_s$$

5 破损安全结构断裂控制

破损安全结构的特点是缓慢的裂纹扩展速度、较大的临界裂纹长度和便于检查等因素。破损安全结构一般分三种类型：缓慢裂纹扩展结构、多传力途径－破损安全结构，破损安全止裂结构。在设计中应通过有效的结构布局、合理地选择材料、降低结构细节的应力集中、提高裂纹的可检性、合理地选用延缓裂纹扩展速率的工艺措施、必要时对某些构件的制造方法做出特殊规定等各种技术途径来提高结构性能。

5.1 结构布局

为了将裂纹扩展速率有效地控制在容许范围内，避免在规定的使用期内发生意外断裂，必须采取破损安全结构，并要求结构便于检修。破损安全结构可以采用以下方法实现：

1）采用多通道载荷结构。其设计思想是，当结构中有一个构件断裂，完全丧失承载能力时，载荷可以由其他构件承受，而不致发生事故。

2）采取止裂措施。①一个构件由若干个元件组成，如其中一个元件出现裂纹，不致扩展到其他元件上。②在零件的预期裂纹扩展途径上钻一个小孔，或设止裂缝。当裂纹扩展到孔或缝时，裂纹因尖端变钝而停止扩展。③在裂纹的扩展途径上设加强件，也称为止裂件。

3）采用断裂前自动报警的安全措施，如压力容器的断裂前渗漏预警。

考虑结构布局时，传力路线不宜过于集中。这样，就有可能采用静强度较低但断裂韧度特性好的材料，而避免用高强度钢。高强度钢临界裂纹长度短，裂纹扩展速率高，按断裂韧度设计时难以满足裂纹扩展寿命要求。

分散传力有利于安全这一原理还体现在传力途径上。一般来说，多传力途径比单传力途径有利于安全。分散传力结构的一个典型例子是机翼多交点厚蒙皮多墙结构。由于蒙皮传力比较分散，在机翼根部又是通过多个接头将载荷传到机身，从而有可能使机翼主体材料全部采用静强度较低但断裂韧度特性好的铝合金。

5.2 材料选择

按静强度观点设计结构，往往追求材料的静强度指标以求减小结构质量。但从断裂力学观点看这是不充分的。静强度高的材料往往断裂韧度较差，一旦材料有某种缺陷或裂纹，很容易导致材料的低应力脆性断裂。因此，为了提高结构的断裂韧度特性，选择材料时首先要考虑的是其抗断裂能力，选择断裂韧度高、裂纹扩展速率小的材料。

裂纹扩展速率是影响裂纹扩展寿命的主要因素，理想的选择是在短裂纹阶段和长裂纹阶段的扩展速率都低的材料。前者是为了满足耐久性要求，后者是为了满足断裂韧度要求。但是，很难找到一种在短裂纹阶段和长裂纹阶段裂纹扩展速率都低的材料。因此，同样是按裂纹扩展寿命要求选择材料，也要分清楚最关心的是哪个裂纹范围内的寿命。如果设计对象耐久性要求的矛盾更突出，应该选择在短裂纹范围内裂纹扩展速率低的材料；如果断裂韧度要求的矛盾更突出，则应该选择长裂纹范围内裂纹扩展速率低的材料。

在选择材料时，还应尽可能地考虑到具体的载荷与环境情况。一种材料在给定温度和介质下表现出的断裂韧度特性可能优于其他材料，但是在另一种环境下可能会出现相反的结果。一种材料在等幅加载时可能表现出较好的特性，而在谱载荷下另外一种材料可能显示出较好的断裂特性。此外，不能仅考虑等幅加载下的 da/dN 特性，还应考虑迟滞效应、应力比及环境等影响因素。例如，当有腐蚀介质时，应首先考虑较高的 K_{ISCC} 和较低的腐蚀裂纹扩展速率 da/dt。

5.3 细节设计

裂纹往往在结构不连续的局部区域开始，并在环境载荷作用下逐渐扩展。因此，结构细节的具体形式及应力水平对

构件的使用寿命有很大影响。由裂纹扩展速率公式$\frac{\mathrm{d}a}{\mathrm{d}N}=C\ (F\Delta\sigma\sqrt{\pi a})^n$可知，裂纹扩展速率是应力强度因子幅度 ΔK 的函数，降低应力幅值 $\Delta\sigma$ 和应力强度因子修正系数 F 都能有效地降低裂纹扩展速率。

降低应力意味着增大结构尺寸，或增加另外的传力构件，这将导致结构重量的增加。如果在较大范围内增加构件尺寸，还可能会影响邻近结构的工作状态，甚至影响载荷的重新分配。所以，降低应力水平不一定是个好办法。

更多的注意力应放在提高结构细节设计水平上。通过认真而精细的构形，改善结构细节的应力分布，降低应力集中，提高结构细节自身的抗裂纹能力，可以获得显著的寿命增益而不会付出重量的代价。

5.4　裂纹检测

在进行破损安全结构细节设计时，裂纹的可检查度是必须考虑的内容。对于缓慢裂纹扩展结构，在质量和结构条件允许下，也应尽可能设计成可检结构。这样，在一个寿命期内可安排 4 次检查。通过检查，对裂纹扩展进行监测，及时排除可能导致结构破坏的裂纹，可大大提高结构的安全性。对于破损安全结构，在允许的条件下，应尽量开设较多的检查通道。特别是对断裂韧度关键部位，在使用中能对裂纹扩展情况进行监测是确保使用安全性的重要环节。

提高裂纹可检度的方法很多，对不同的结构和不同的结构构形所采用的方法也不一样。图 6.3-38 是一个壁板设计中提高结构裂纹可检度的例子。图 6.3-38a 所示的结构很难检查板的搭接边缘出现的裂纹；而图 6.3-38b 所示的结构是将搭接板端设计成波浪形的，裂纹可以被方便地检查出来。

图 6.3-39 是另一个提高裂纹可检度的例子。如将大梁设计成图 6.3-39a 所示的结构，一旦缘条上出现了裂纹，由于大梁缘条完全被蒙皮盖住，裂纹很难被检查出来。在工艺许可的条件下，将其设计成如图 6.3-39b 所示的形式，紧固件孔边出现的裂纹就容易被检查出来。

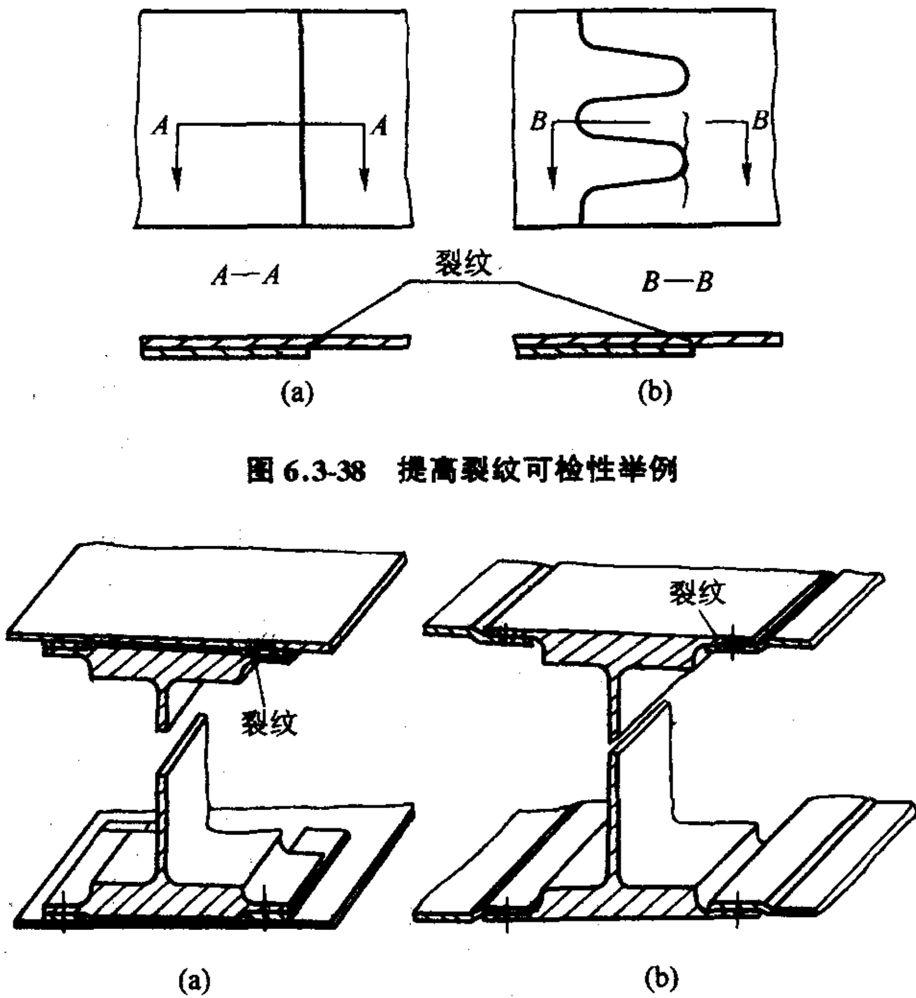


图 6.3-38　提高裂纹可检性举例

图 6.3-39　提高裂纹可检性举例

检验程序是断裂控制的主要环节之一，包括裂纹长度检测和检验周期的确定。测定裂纹尺寸必须采用适当的检验方法。按照断裂韧度设计的零件都必须易于检测。当要求的初始裂纹尺寸小于质量控制方法的检测能力时，必须改变材料，或降低应力水平，以求得到较大的初始缺陷容许值。对于单载荷通道设计，则要求有较大的临界裂纹尺寸，以便于检测。

按照断裂韧度设计，零件的剩余寿命必须大于其检查周期。为确保零件在检修期内的安全使用，一般取检修周期小于或等于其剩余寿命之半，这样便可以保证检验人员在零件发生破坏之前至少有两次发现裂纹的机会。

编写：谢里阳（东北大学）
王永岩（青岛科技大学）

第4章　材料冲击强度设计

1　冲击载荷作用下材料的力学响应

高速作用于物体上的载荷称为冲击载荷。冲击载荷与静载荷的主要差异，在于加载速率不同。加载速率是指载荷施加于试样或机件的速度，用形变速率可以间接地反映加载速率的变化。

以采矿和筑路中广泛应用的凿岩机为例，其活塞以6～8 m/s的速率冲击钎杆，钎杆再将冲击能量传递到钎头，从而使岩石破碎。又例如作为反坦克武器的长杆穿甲弹，以1.5～2.0 km/s的速度着靶后实现侵彻穿孔。人造卫星和导弹在宇宙飞行过程中与陨石碰撞的相对速度可高达20 km/s。冲击速率愈高，则材料的应变率愈高。一般说来，准静态加载的应变率为$10^{-5}\sim10^{-1}s^{-1}$量级，冲击加载的应变率通常为$10^{2}\sim10^{4}s^{-1}$，甚至可达$10^{6}s^{-1}$以上。随着应变率的提高，本构关系中的动态效应便显示出来，表现为材料强度的提高，而塑性、韧性的变化则较为复杂；材料的应力水平与其应变程度、应变率和温度有关。

材料对冲击载荷的响应可分为三种情况。

1）弹性响应　当外载荷产生的应力低于材料的屈服点，应力波的传播不造成材料不可逆的变化，材料表现为弹性行为，线性胡克定律适用。

2）弹塑性响应　当应力超过屈服点而低于1×10^{4} MPa时，材料的响应可用耗散过程来描述，要考虑大变形、黏滞性、热传导等，本构方程十分复杂，呈非线性。

3）流体动力学-热力学响应　当应力超过材料强度几个数量级、达到几兆帕或更高时，材料可作为非黏性可压缩流体处理，其真实结构可不予考虑；材料的响应可用热力学参数来描述，其本构关系可用状态方程表示，也为非线性。

图6.4-1表示低碳钢在不同温度和应变率区域中塑性流动的耗散机制。在区域Ⅰ中，屈服强度对应变率和温度不敏感，塑性流动主要为非热机制所控制，位错、析出相质点和晶界等所引起的长程内应力场起着主要的作用。在区域Ⅱ中，屈服强度对温度和应变率较为敏感，一般认为，对于hcp，fcc和bcc金属，塑性变形速率主要为位错运动的热激活所控制，包括Peierls-Nabarro、林位错、螺型位错的交割和交滑移，以及刃型位错的攀移等。在区域Ⅲ中，屈服应力受应变率和温度的影响也较小，这和孪晶的形成有关。区域Ⅳ包括很高的应变率，即$10^{3}\sim10^{5}s^{-1}$。此时屈服强度对应变率极为敏感，这与位错运动的黏滞性阻力有关，可以用声子黏滞性理论来解释。

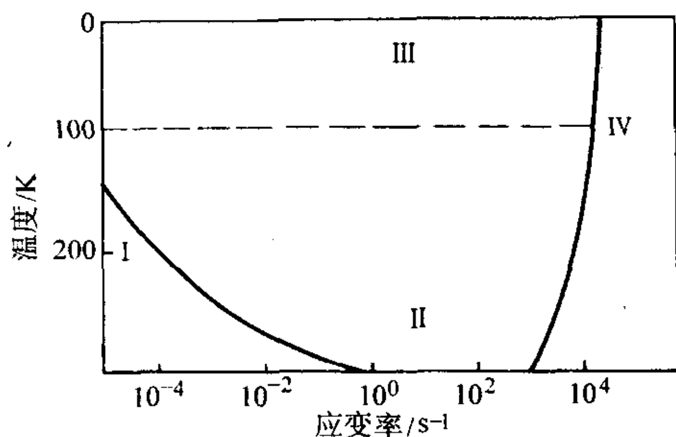


图6.4-1　塑性流动时的耗散机制

Ⅰ—非热机制；Ⅱ—热激活机制；Ⅲ—孪晶机制；Ⅳ—阻尼机制

高应变率下本构关系的实验测量较为困难，不仅需要高精密的仪器设备，还需要应力波和动态响应的理论知识。

2　冲击载荷下材料的损伤和破坏

冲击载荷下材料的失效和破坏比较特殊，常常难以归类和分析。在高应变率加载下，不仅材料的强度和塑性会发生变化，而且其变形和断裂机制也可完全不同于准静态加载。

（1）应力波反射引起的破坏

由于应力波的反射和叠加，会引起冲击载荷下材料的层裂、纵裂和角裂。

如果由爆炸和高速撞击所产生的应力波施加于平板上，假定这是一个锯齿形的压缩波，它将沿着板厚方向传播而不改变其形状和烈度。从图6.4-2可见，这个应力波在到达板的背面以后就会发生反射，而变成拉伸波。在$t=\frac{1}{2}\frac{p}{c}$的瞬间，在离开背面$\frac{1}{2}$波长的位置，会出现一个最大拉伸应力σ_m。如果此应力低于材料的断裂强度，即$\sigma_m<\sigma_F$，则不会发生破坏，这个拉伸应力波将继续向左行进。相反，当$\sigma_m>\sigma_F$，就会在该处开裂；其裂纹平行于板面，称为层裂。由于金属板中可能存在着非金属夹杂物，这些夹杂物和金属基体的结合较弱，即这个局部区域中断裂强度σ_F较低，就往往在该处产生层裂。按照拉伸应力波和夹杂物薄弱面的组合，也可能产生多层的层裂。

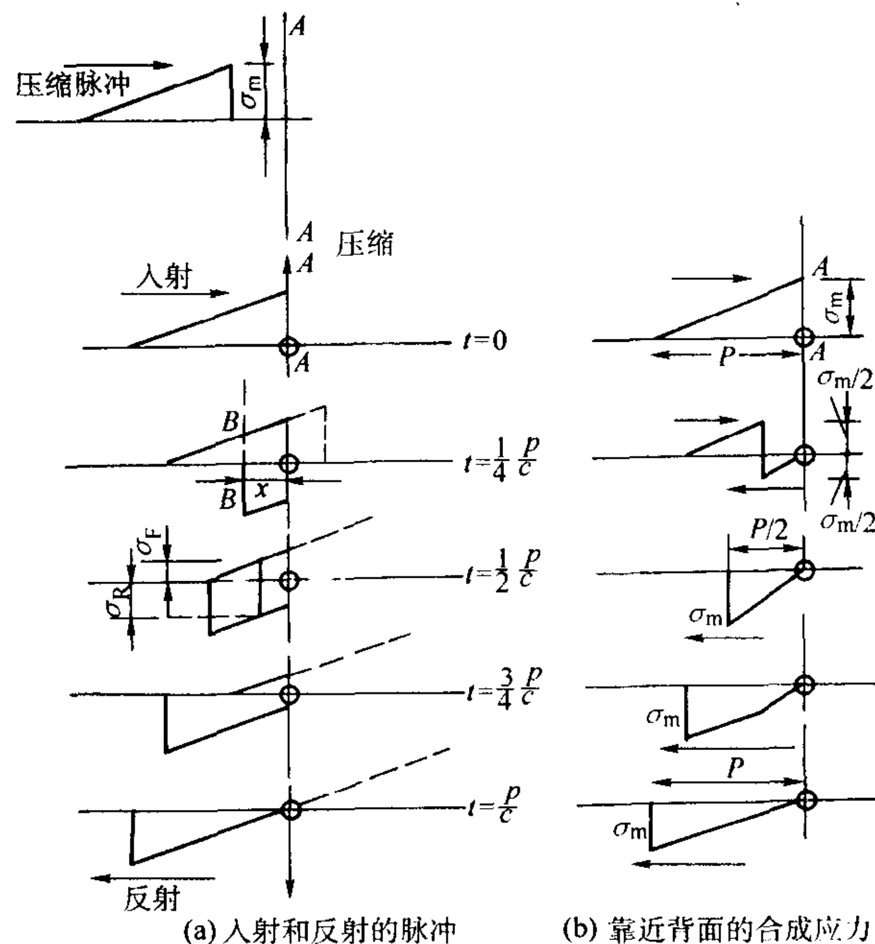


图6.4-2　入射的压缩波经板的背面反射而成拉伸波

图6.4-3表示炸药在一方柱体的端面产生应力波，此应力波在柱体的侧面发生反射；两个反射的应力波在中心线上相遇，因而可能产生中心纵裂。

图6.4-4表示另一种形状的柱体；此时两个反射的应力波在角上相遇，会产生角裂。

图6.4-5为一系列角裂的例子。值得注意的是，沿着对角线方向材料的厚度最大，在静载作用下不可能在角上造成断裂；但在动载作用下，角裂是很常见的现象。

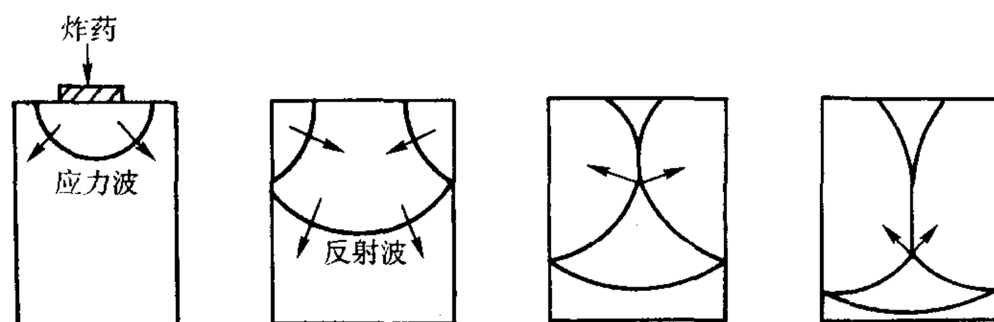


图 6.4-3 由于反射波造成的中心纵裂

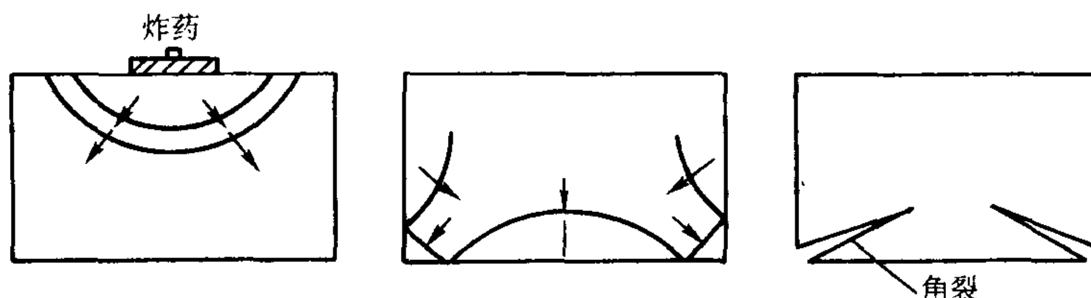


图 6.4-4 由于反射波造成的角裂

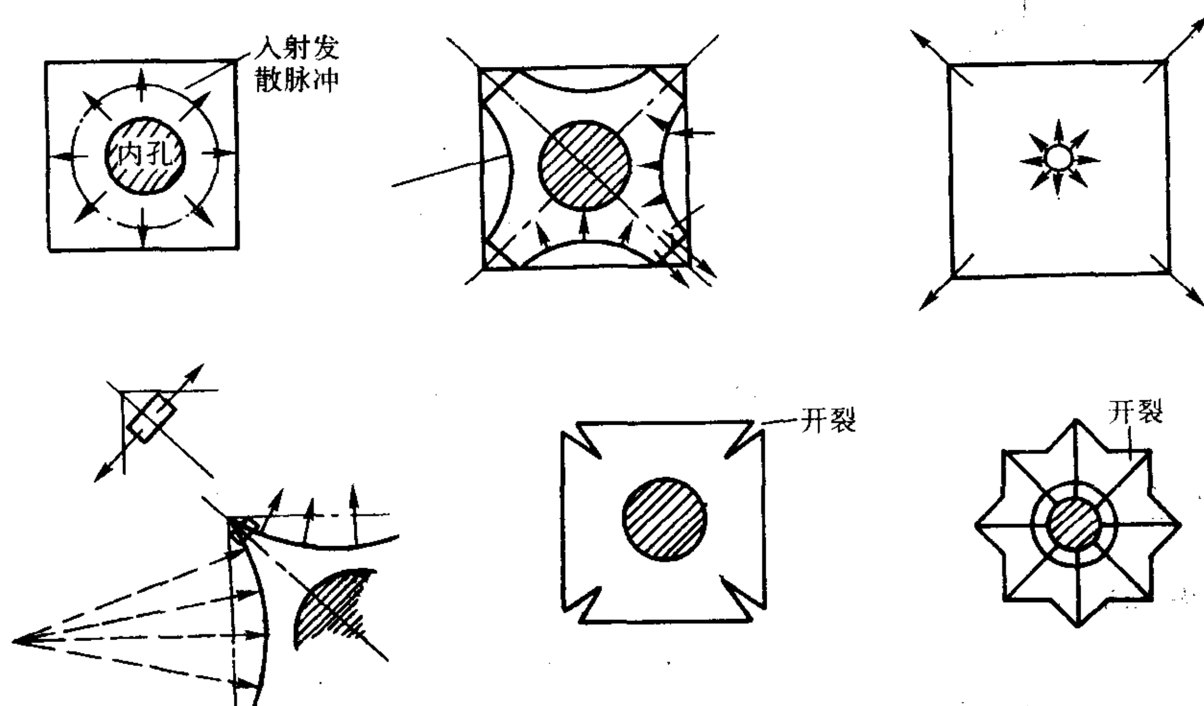


图 6.4-5 由于反射波造成的各种类型的角裂

图 6.4-6 表示锥形的物体，在底面上施加冲击力，由于应力波的反射，可在中心线上产生裂纹，也可能在腰部碎裂。

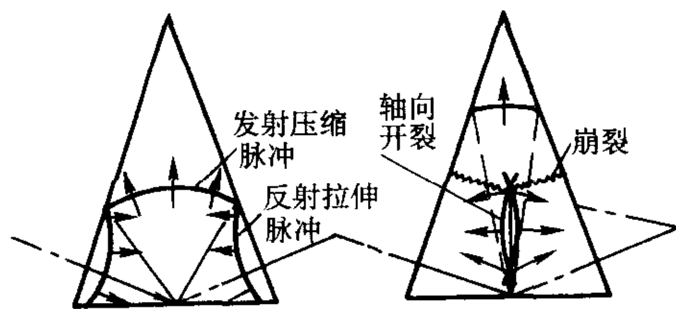


图 6.4-6 锥形物体的开裂

图 6.4-7a 为一软钢制成的厚壁圆筒，内孔承受爆炸载荷；断裂沿着内孔表面最大切应力方向进行，并因环箍应力沿半径方向进一步开裂。图 6.4-7b 为 24ST 铝合金制成的筒体，其外壁受爆炸压缩应力波；由于应力波在内表面会聚并反射，而产生剪切开裂。

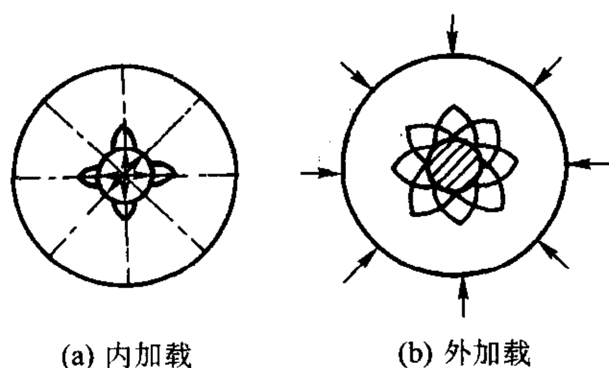


图 6.4-7 厚壁圆筒的开裂

(2) 损伤参数和动态失效模式

冲击失效的影响因素很多。在两个物体中发生冲击时，一般将动的物体称为“弹”，静的物体称为“靶”。除了两者的相对速度以外，“弹”和“靶”的力学性能（如各种强度、塑性、韧性指标和弹性模量、塑性模量等）、物理性能（如密度、比热容、导热系数、熔化热和气化热等）及其几何形状等都可能对失效过程产生影响。Johnson 最早提出可根据损伤参数来划分动态失效模式。所谓损伤参数是指“弹”的动能密度和“靶”的失效强度的比值，即 $D=\rho v^2/\overline{Y}$。式中 ρ 是弹的密度，v 是弹的速度，$\overline{Y}$ 是靶的失效强度，通常可用抗剪强度来表征。D 值的大小表明惯性过程对材料行为的影响。D 值愈高，则意味着惯性过程对材料行为起愈益重要的作用，而显微组织等愈退居次要地位。

在低速冲击情况下，$D<10^{-3}$。对于脆性材料，断裂以裂纹扩展的方式进行，其过程可用线弹性断裂力学来描述；对于塑性材料，可能表现为与时间有关的空洞的萌生、扩展与聚合，也包括因重复冲击而导致的动态疲劳断裂。此时材料的性能决定于温度和应变率，热激活过程对位错运动起着重要作用。

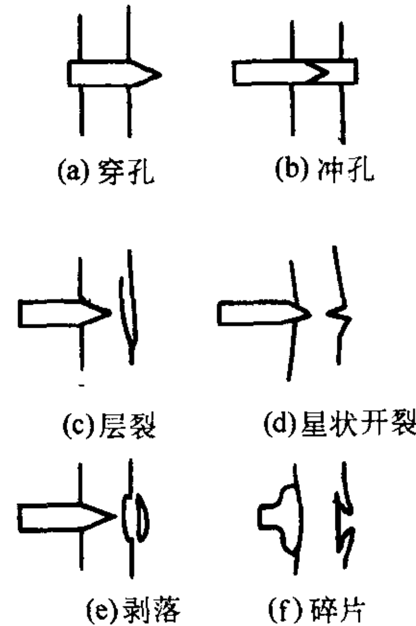
在中速冲击范围，$10^{-3}<D<1$。当“弹”的速度达到 500 m/s 量级，情况就发生根大转变，位错运动不再依赖于热激活，而是取决于声子拖曳作用，因而流变应力对应变率的敏感性大为提高。裂纹出现分叉现象。裂纹速度 c_b 相当于纵向声波传播速度 c_0 的 3/10，即 $c_b=0.3c_0$；应力强度因子 K_{Ib} 接近于材料断裂韧性 K_{Ic} 的 4 倍，即 $K_{Ib}=4K_{Ic}$，这是

开始出现裂纹分叉的两个条件。裂纹分叉的角度可以在20°~120°的宽广范围内变动。对于脆性材料，则发生多重裂纹前沿扩展，即导致碎裂现象。

在高速冲击时，$D\geqslant 1$，失效过程决定于惯性作用，显微组织的影响减小。在极端情况下，当冲击速度 > 1 000 m/s时，"弹"和"靶"两者都在冲击点发生气化；当速度稍低，即使不发生气化，也会发生绝热熔化，材料性能表现为似黏滞件的流体，冲击波的压力大大越过材料的动态屈服强度。

图6.4-8表明中速冲击时穿甲弹和装甲板的失效模式。图中进一步说明了6种可能的失效模式以及装甲板硬度（强度）和失效模式的关系。

当装甲板硬度较低时（小于350HV），将因塑性变形而被穿孔。当硬度稍高（350~450HV），则发生局部的剪切带，并冲下直径与炮弹相同的圆柱体。继续提高装甲板的硬度（450~560HV），则由于应力波的作用，在板的背面附近产生层裂，或在板的背面产生星状开裂，裂纹沿半径方向扩展。最后的两种失效模式是在板的背面产生碎片和剥落。



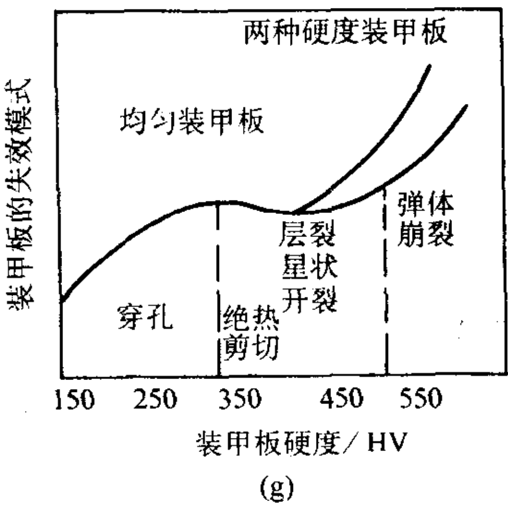


图6.4-8　装甲板的失效模式及与其硬度的关系

3　材料的应变率响应

3.1　复合材料的高应变率响应

1）单向增强复合材料的应变率效应　对于单向纤维增强材料，在平行增强的方向施加拉伸载荷时，复合材料的力学响应主要由纤维决定。因为纤维比基体具有较高的强度和韧性，因此基体的应变率相对来说不是主要的。因此对于单向增强CFRP材料，断裂强度都随应变率增加而增加。对于低纤维复合材料（如纤维体积只占整个材料体积的17%）而言，试验应变率在 $2\times10^{-4}\sim20\ \mathrm{s}^{-1}$ 范围时，环氧树脂的基体的应力-应变响应表示在图6.4-9a中，而用玻璃纤维单层增强的环氧树脂材料的应力-应变响应表示在图6.4-9b中。当应变率进一步增加到891 s^{-1}时，材料失效模式发生了变化。所以纤维起着一定的控制作用。值得注意的是，在低应变率下，环氧树脂的拉伸应变直至失效都比复合材料来得大，而在高应变率下并非如此。在高应变率下，环氧树脂的失效应变限制了复合材料的抗拉强度。在压缩载荷下，在单向增强材料中，环氧树脂的特性是重要的，对于抗弯曲而言同样基体的切变模量应该足够大才能起到支持纤维的作用。

2）编织增强复合材料的应变率效应　编织增强的复合材料在拉伸或弯曲载荷下，基体局部应变增加，纤维与基体之间存在着较大的相互干涉。对于GFRP材料，在压缩载荷下，导致了压缩强度对应变率敏感性的增加，而且抗拉强度的速率敏感性也增加，如图6.4-10所示。对于编织增强CFRP材料也导致了强度对小应变率敏感的可能性，如图6.4-11所示。

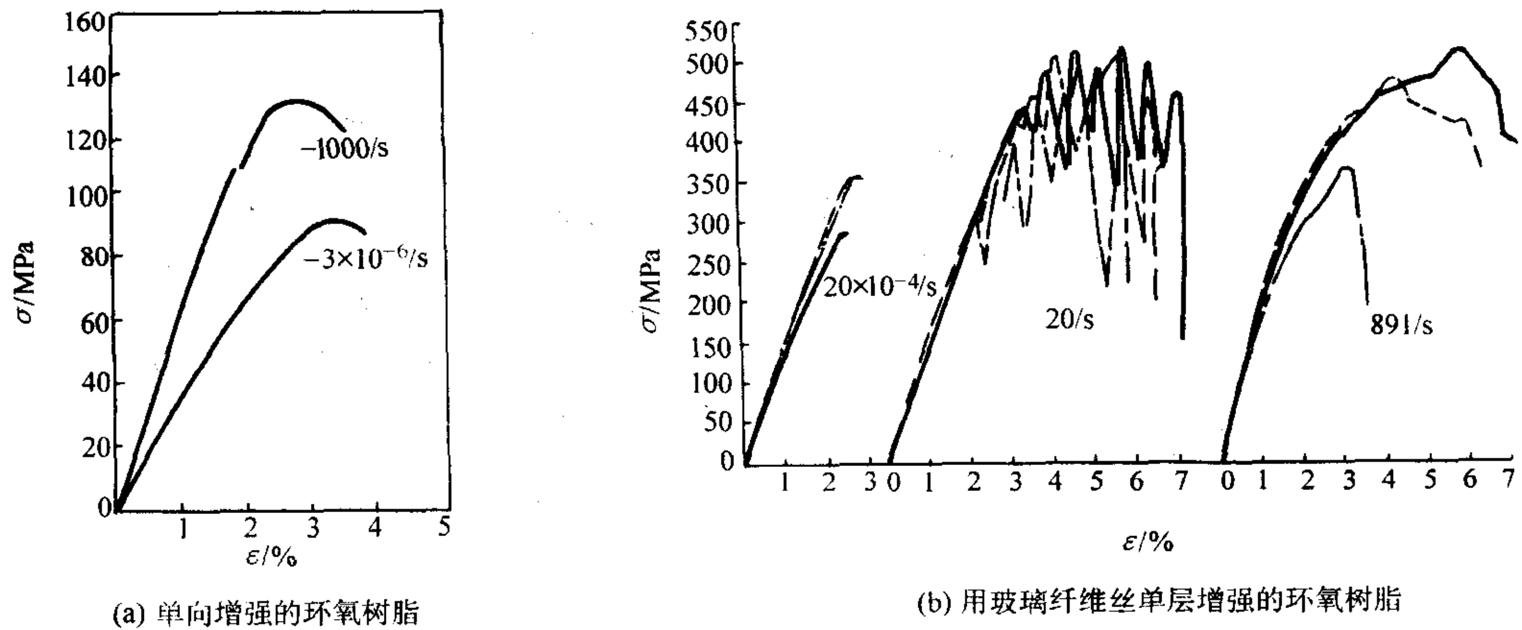


图6.4-9　单向增强GFREP试件对拉伸应力-应变响应的应变率效应

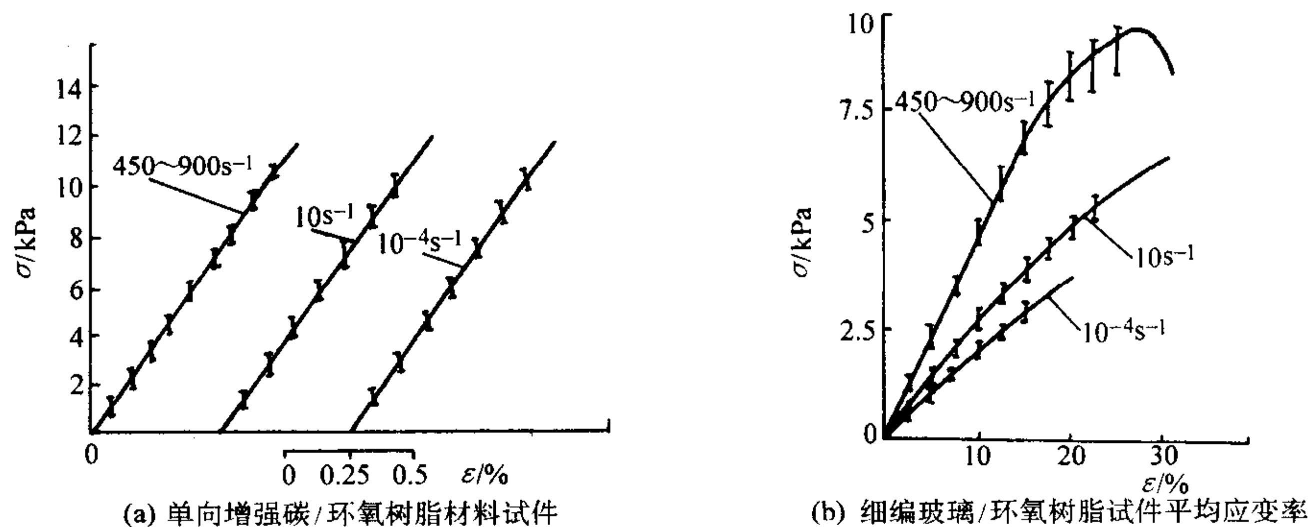


图 6.4-10　复合材料拉伸应力-应变曲线的应变率效应

在编织增强的复合材料中，纤维与基体间的干涉增强也可以解释为拉伸模数对应变率敏感性增加，因为有时在试验中可以观察到曲线斜率变化。例如在细织玻璃增强材料试验中（图 6.4-10b），当应变率由 10^{-4} ~ 970 s^{-1} 变化时，弹性模量由 20 GPa 增加到 45 GPa，随着应变率的增加弹性模量都有明显的增加。对于粗编织材料相对增加少些。具有相同编织增强的几何形状、相同聚合脂基体的一组 CFRP、KFRP、GFRP材料在相同应变率范围内弹性模量的试验结果如图 6.4-12a 所示。图中表明，这三类增强纤维材料由于纤维不同，模量的绝对值也不同，但它们的应变率敏感性却是相似的。这充分说明在编织增强复合材料中，模量的应变率敏感性由沿轴向排列的纤维与基体之间的弹性干涉决定。纤维与基体之间的干涉强度将随着增强的几何形状而变化。而且细编织多半大于粗编织的干涉强度，这就导致了细编织的弹性模量比粗编织的速率敏感性大。图 6.4-10b 与图6.4-11b的试验结果表明了这一点。

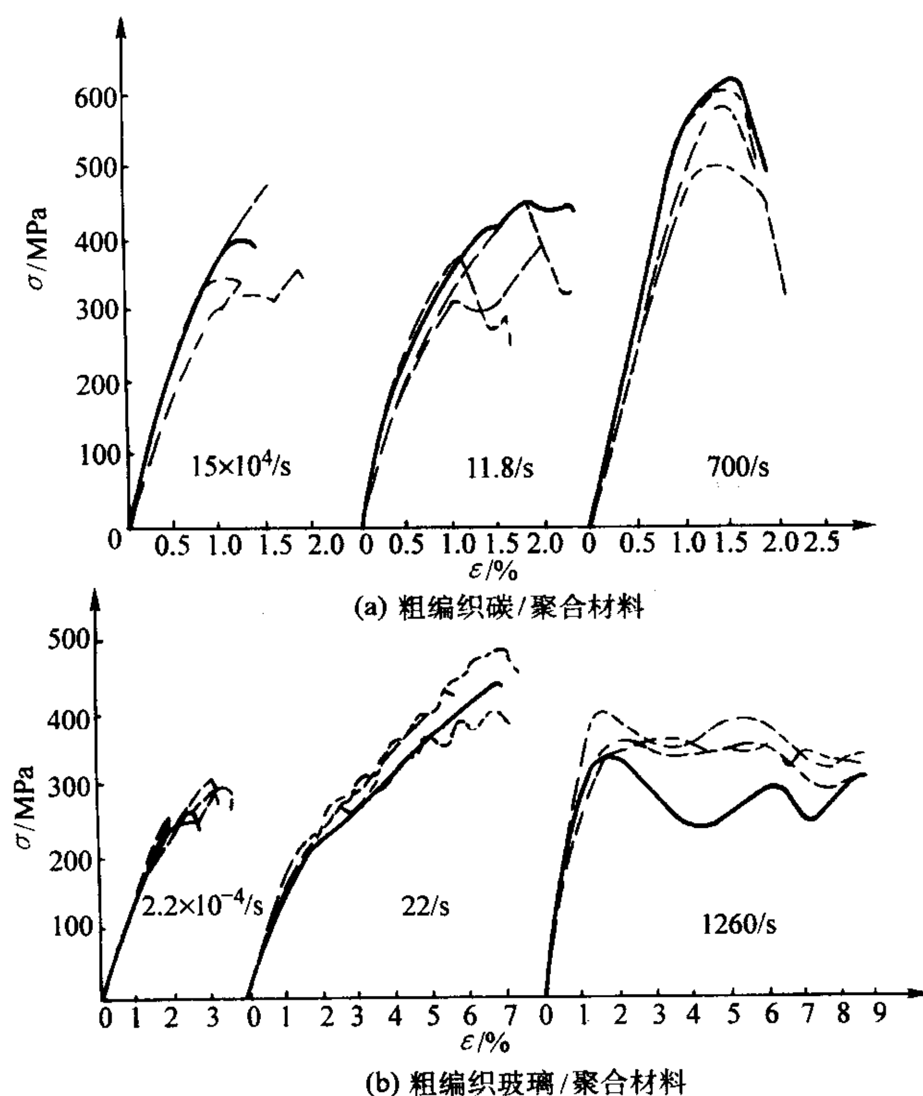


图 6.4-11　增强纤维类型材料在动力拉伸应力-应变响应方面的效应

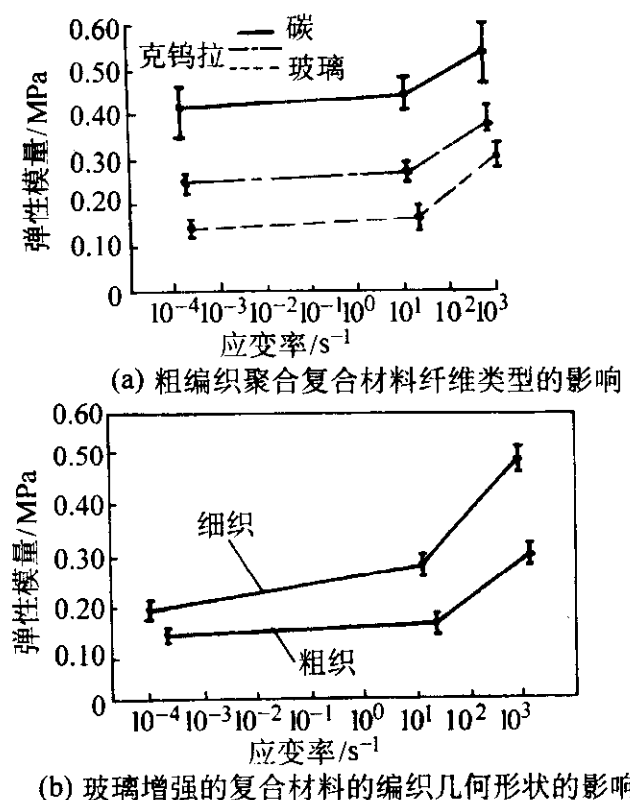


图 6.4-12　编织纤维复合材料中弹性模量对应变率的敏感性

由图 6.4-11b 看到，低应变率或中等应变率的应力 - 应变曲线与拉伸断裂相联系，而在撞击试验中断裂是由于纤维的拔拉，因此应力 - 应变曲线有明显的变化。尤其值得一提的是在冲击速度下试件所能支撑的最大载荷比中等应变率条件下要小，也许这是因为纤维与基体之间的边界强度随着应变率增加而减小。

3.2　金属材料的应变率响应

了解材料的力学特性对应变率的依赖关系是很重要的。如果对撞击载荷进行适当的分析就可以合理地设计有关的撞击结构。这不仅需要知道材料的动力学特性，还要知道材料的应力、应变、应变率以及与温度相关的力学状态方程，从而用来进行数值分析，并了解结构的动态响应。至今为止，已经在不同程度上研究了相当多的金属和合金以及一部分非金属材料的动态力学响应问题。

在考虑应变率效应时，首先要区别材料响应与结构响应。作为一个设计者不但要了解结构响应而且要了解材料响应。作为粗略的近似，应变率等于撞击速度除以结构的变形区域尺寸。在许多情况下，结构尺寸是相当大的，这将导致一个平均的应变率。在许多结构碰撞问题中，如爆炸加载圆筒、爆炸成形、高速穿甲等都属于大变形问题。然而如果进一步考虑狭窄的变形区域即靠近撞击点附近的局部变形区域，由于应变率很大，有可能引起结构突然失效。作为一个例子，考虑短柱形射弹高速横向撞击金属靶板，当撞击速度

低于某个临界值时，靶板的塑性变形集中在撞击坑附近（呈鼓形），由于撞击点附近的大塑性变形以及靶板周边固定处的摩擦效应，射弹的动能完全被靶板吸收。在这种情况下，塑性变形率很低。而在高速碰撞情况下，塑性流动集中在更小的撞击点附近的区域里，靶板被穿孔，而靶板的其他地方变形较小，靶板吸收的能量大大减少。为了描述这种破坏方式，除了要求有一个破坏准则外，还要有一个反映材料力学特性的合理的本构关系。

塑性流动可以集中在更狭窄的区域里，例如刻槽、裂纹或截面积变化等，这就导致了脆性或低应变率断裂或由于热效应而断裂，例如绝热剪切带中的裂纹。本章主要讨论材料对应变率的响应，但所有的结果对结构碰撞工程问题都是有用的。

1）高应变率下的应力－应变曲线　试验研究表明，大多数金属和合金的强度都随着应变率的增加而增加。图6.4-13a、b分别表示铝和钛在扭转加载下的应力－应变曲线，图6.4-13c表示软钢在拉伸加载下的应力－应变曲线，图6.4-13d为纯铜在冲压加载荷下的载荷－位移曲线，其应变率或加载速率在几个数量级范围内变化。这里，各种金属材料的晶格类型不同，铝和铜为fcc金属，铁为bcc金属，钛为hcp金属，加载方式则包括扭转、拉伸和冲击，但一个共同的规律是，随着应变率的增加，材料的强度也随之提高，其中bcc金属显示出最大的应变率敏感性，软钢的屈服强度可以成倍地提高。

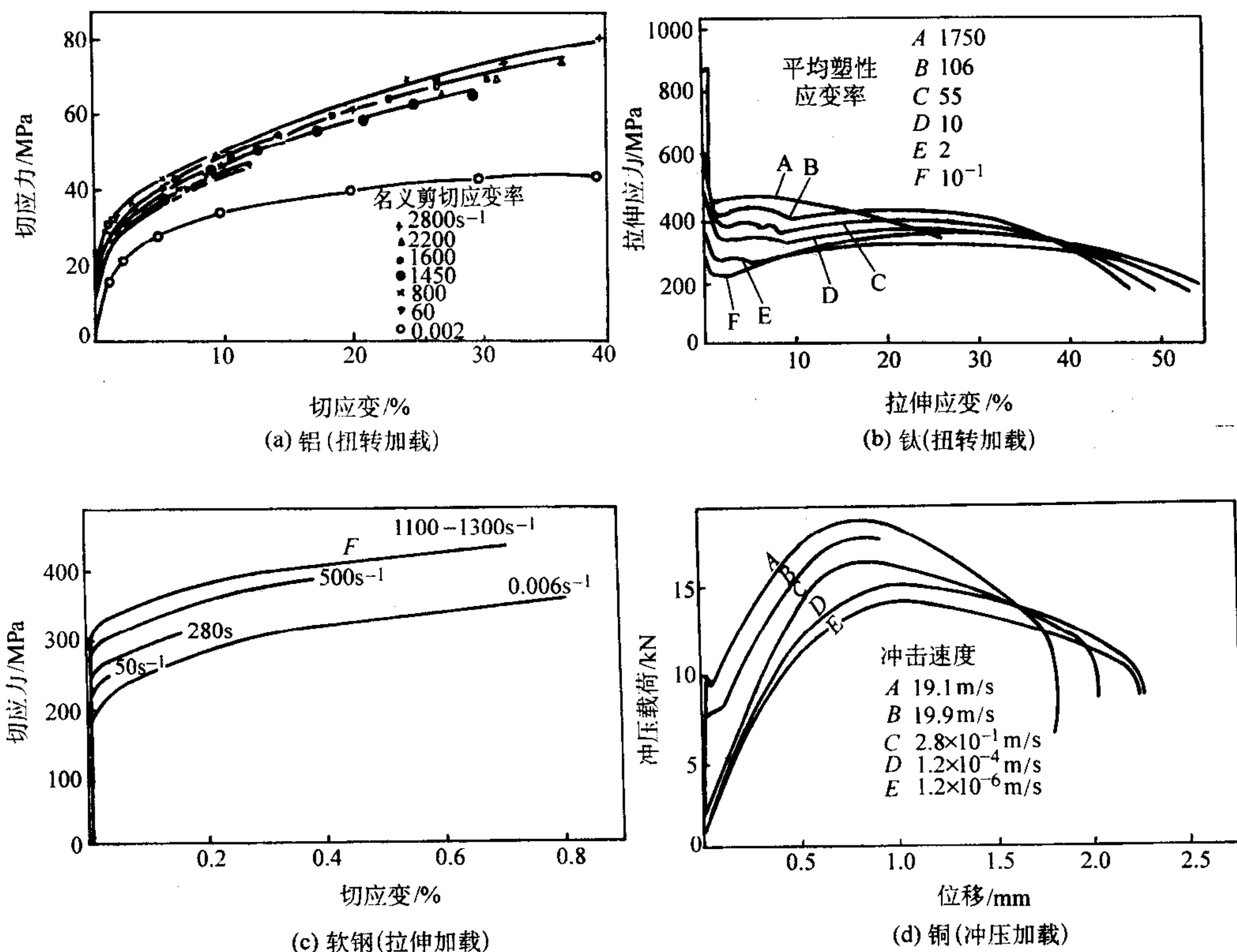


图6.4-13　应变率对金属应力应变曲线的影响

对于合金来说，这种变化的总趋势仍是保持不变，只是应变率对流变应力的影响有所减弱。例如，纯铝是应变率敏感材料，其高应变率（$10^3 s^{-1}$）的流变应力比准静态（$10^{-4} \sim 10^{-3} s^{-1}$）提高20%～60%；但对于经时效强化的变形铝合金，其动态应力－应变曲线与准静态却几乎重合。图6.4-14a、b分别表示淬火回火合金钢、钛－铝（6%）－钒（4%）合金在不同应变率下的应力－应变曲线。可以看出，随着应变率的提高，其应力－应变曲线的位置也上移，但提高的幅度则不如纯金属。

2）应变率对强度和塑性的影响　大量试验证明提高应变率可使屈服强度、极限强度和流变应力提高。曾有研究报道，当应变率从 $\dot{\varepsilon} = 10^{-4} s^{-1}$ 提高到 $\dot{\varepsilon} = 10^3 s^{-1}$，fcc金属的延伸率随之提高，而bcc金属的延伸率却有所降低。但这个结论未能得到普遍的证实。

图6.4-15为经淬火回火的35NiCrMoV钢的试验结果。从图中可以看出，随应变率的提高，钢的强度也随之提高，但并非必然使塑性降低；恰恰相反，在 $10^2 \sim 10^4 s^{-1}$ 应变率范围内，断面收缩率 Ψ 和延伸率 δ 都有所提高。在CrMoV钢中也得到类似结果。

图6.4-16表示18Ni马氏体时效钢的试验结果。18Ni是一种新型的超高强度钢，含18% Ni、8% Co、5% Mo和0.6Ti，依靠析出金属间化合物（Ni_3Ti等）而时效强化，具有很好的强度与韧性的配合。这种钢的屈强比很高，均匀延伸率很小，但仍有相当高的塑性。随着应变率提高，其 σ_b 和 σ_s 单调地提高，但其 ψ 几乎保持不变，或者说略有下降，但绝对值仍在40%以上。

图6.4-17表示奥氏体钢的试验结果。这种钢屈服强度较低，易于变形，均匀延伸率可达较高数值。随应变率的提高，σ_s 急剧提高，σ_b 的提高较为缓慢，而 ψ 和 δ 则略有降低，与马氏体时效钢相类似。值得注意的是，在低应变率下，奥氏体钢的强度远低于中碳铬钼钢；但在高应变率下，奥氏体钢的强度提高很多，甚至可超过铬钼钢。

因此，不同合金成分和显微组织的钢，对应变率的敏感性有很大不同，必须通过试验确定强度和塑性的变化趋势。

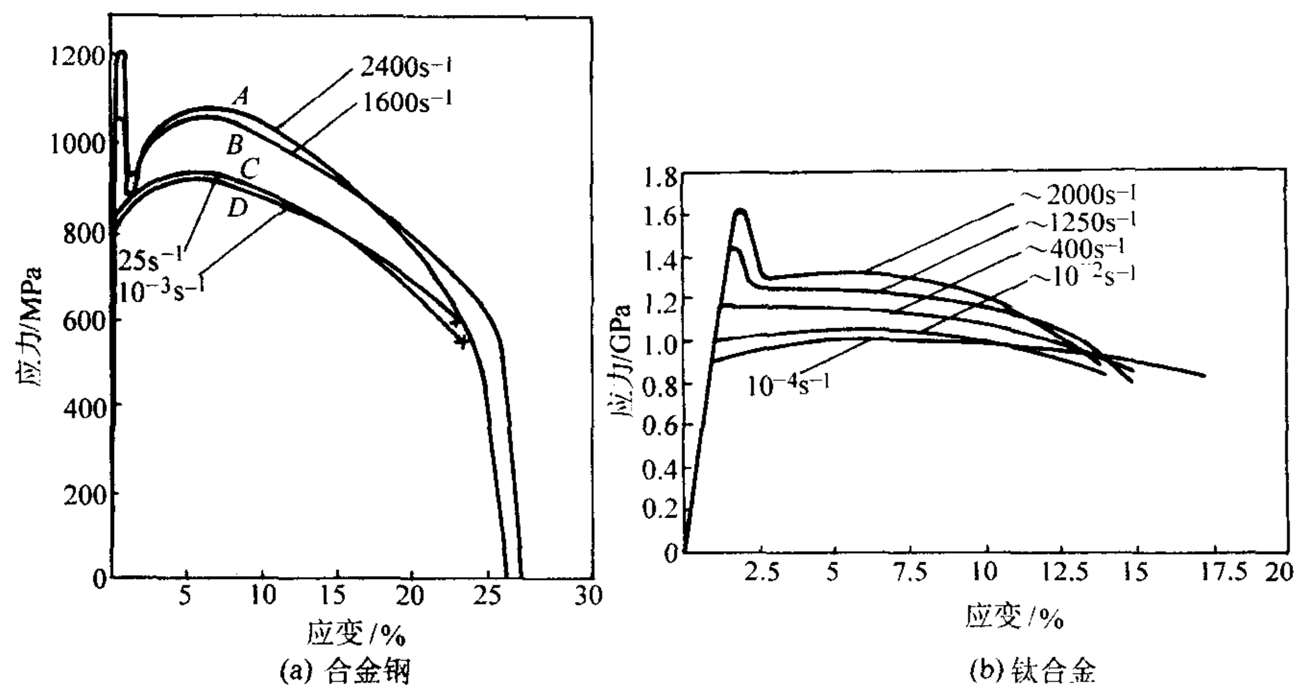


图 6.4-14 应变率对合金应力-应变曲线的影响

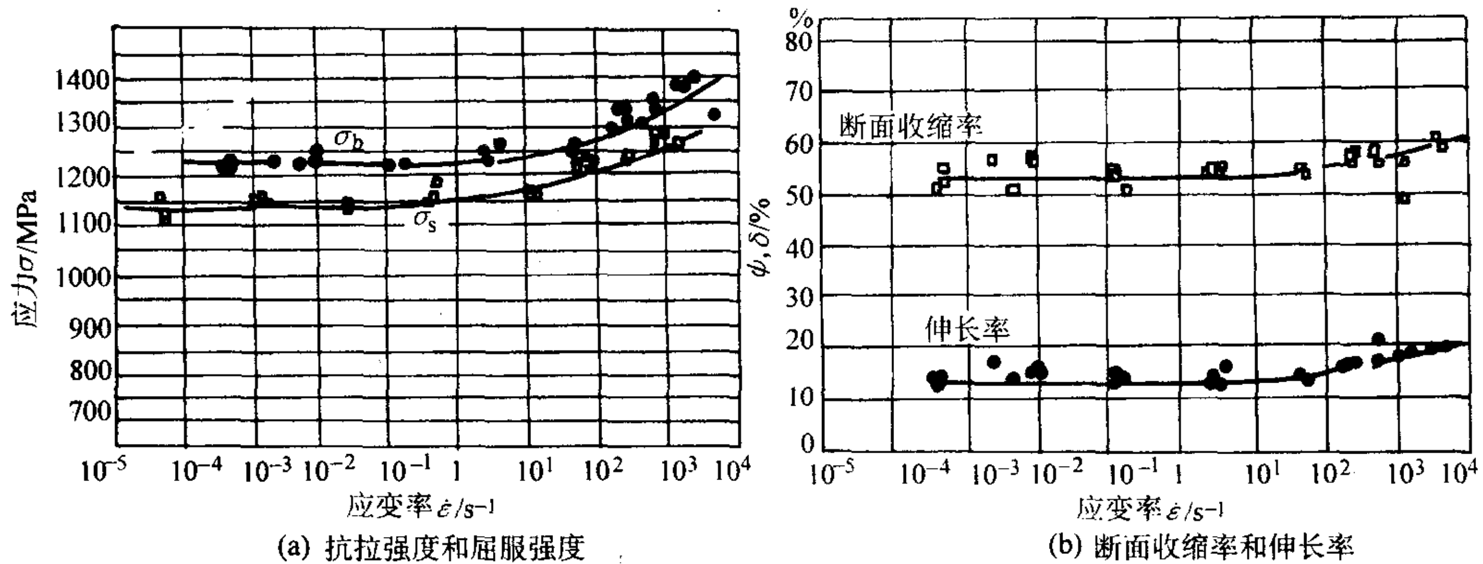


图 6.4-15 应变率对淬火回火钢强度和塑性的影响

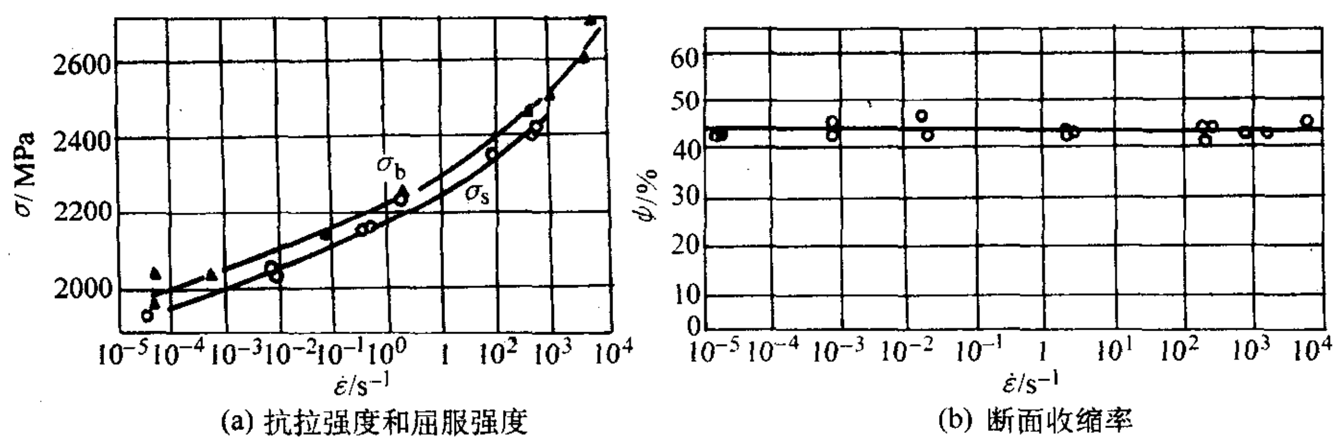


图 6.4-16 应变率对马氏体时效钢强度和塑性的影响

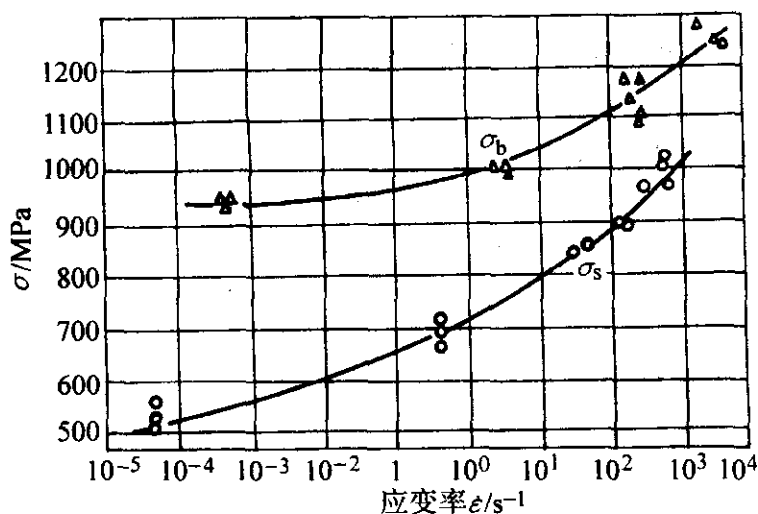


图 6.4-17 应变率对奥氏体钢强度的影响

4 材料在冲击载荷作用下的性能指标

4.1 动态断裂韧性

根据加载速度的不同，断裂韧性的测试方法大体上可分为三类：① $\dot{k} \leqslant 10^3$ MPa·m$^{1/2}$·s^{-1}，即通常在静载下测试的断裂韧性；② 10^3 MPa·m$^{1/2}$·s$^{-1} \leqslant \dot{k} \leqslant 10^5$ MPa·m$^{1/2}$·s^{-1}，指摆锤冲击弯曲试验方法测定断裂韧性；③ $\dot{k} \geqslant 10^5$ MPa·m$^{1/2}$·s^{-1}，指用应力波加载方法测定断裂韧性，常用的是 HSBT 技术。后两者即称作材料的动态断裂韧性。

4.2 冲击试验和冲击性能指标

在冲击载荷下，由于载荷的能量性质使整个承载系统（包括机件）承受冲击能，因此，机件及与机件相连物体的刚度都直接关系到冲击过程的持续时间，从而影响加速度和惯性力的大小。由于冲击过程持续时间很短而测不准确，难以按惯性力计算机件内的应力。所以，机件在冲击载荷下所受的应力，通常是假定冲击能全部转换成机件内的弹性能，再按能量守恒法计算。

在冲击载荷下，由于加载速率大，塑性变形得不到充分发展，故将缺口试样在冲击载荷下进行试验能更灵敏地反映材料的变脆倾向。常用的缺口试样冲击试验一般是冲击弯

曲，有时也用冲击拉伸，但应用较少。

缺口试样静拉伸和偏斜拉伸试验结果的比值，用以评定材料缺口敏感性；缺口试样冲击试验则是按冲击性能指标的绝对值评定材料的缺口敏感性。

冲击弯曲试验方法和原理见图6.4-18和图6.4-19。

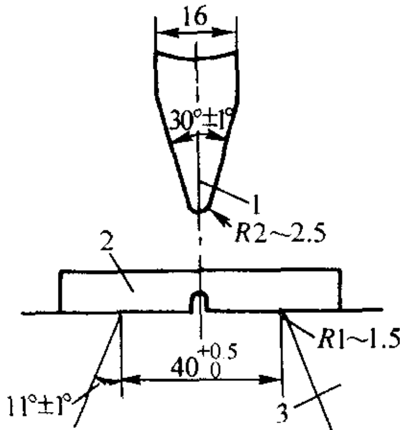


图6.4-18 弯曲冲击试样的安放

1—摆锤；2—试样；3—支座

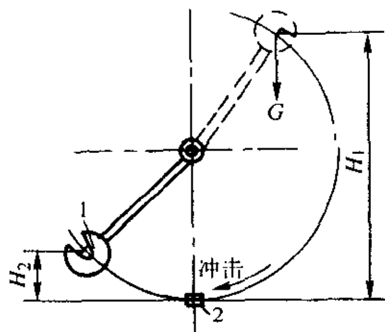


图6.4-19 冲击试验原理

1—摆锤；2—试样

试验是在摆锤式冲击试验机上进行的。将试样水平放在试验机支座上，缺口位于冲击相背方向，并用样板使缺口位于支座中间。然后将具有一定重量摆锤举至一定高度 H_1，使其获得一定位能 GH_1。释放摆锤冲断试样，摆锤的剩余能量为 GH_2，则摆锤冲断试样失去的位能为 GH_1-GH_2，此即为试样变形和断裂所消耗的功，称为冲击吸收功。根据试样缺口形状不同，冲击吸收功分别为 A_{kv} 和 A_{ku}。A_{kv}（A_{ku}）$=G$（H_1-H_2），单位为J。A_{kv} 亦有用CVN或 C_v 表示的。

用试样缺口处截面积 F_N（cm^2）去除 A_{kv}（A_{ku}），即得到冲击韧度 a_{kv}（a_{ku}）

$$a_{kv}\ (a_{ku}) = \frac{A_{kv}\ (A_{ku})}{F_N} \tag{6.4-1}$$

通常单位为J/cm²。

国家标准（GB/T 229—1994）规定冲击试验标准试样是U形缺口或V形缺口。分别称为夏比（Charpy）U形缺口试样和夏比V形缺口试样。测量球铁或工具钢等脆性材料的冲击韧性，常采用10×10×55的无缺口冲击试样。

冲击试验所得到的冲击吸收功 A_{kv} 或 A_{ku} 包括试样在冲击断裂过程中吸收的弹性变形功、塑性变形功和裂纹形成及扩展功等。简单的冲击试验不能将这些不同阶段的功耗区分开来，因此虽然冲击功属于韧性指标，但只是一种混合的性能指标，其物理含义是不明确的，在设计中不能定量使用。在夏比冲击试验机上装备冲击过程的监测系统（示波冲击系统），可以记录试样冲击变形和断裂的全过程，从而得以对断裂过程进行分析。示波冲击系统得到的载荷-挠度（$P-f$）曲线如图6.4-20，曲线所围成的面积即冲吸收击功。曲线上 P_{Gy} 之前为弹性变形阶段，从 P_{Gy} 开始，试样进入塑性变形和形变强化阶段，由于缺口的存在，塑性变形只发生于缺口附近的局部范围，而且缺口越尖锐，参与塑性变形的材料体积越小，得到的冲击吸收功越低。缺口形式对某低合金钢冲击吸收功的影响列于表6.4-1。由此可见缺口形式对冲击试验结果的影响之巨。一般评定材料时，希望揭示不同材料在冲击吸收功方面的差异，因此，应根据材料的韧性情况，选择合适的缺口形式。如对于一组韧性很高的材料，应选用尖锐缺口试样，而对于韧性差的材料，则应选用钝缺口试样甚至不开缺口。

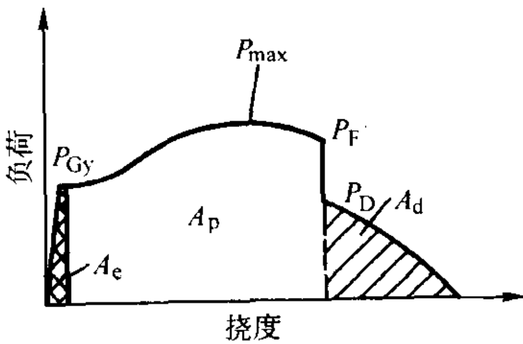


图6.4-20 缺口冲击试样的载荷-挠度曲线

表6.4-1 缺口锐度对冲击某低合金钢冲击功的影响

缺口形式	U形	V形	预制裂纹
冲击韧度/J·cm⁻²	68	28	10

当载荷达到 P_{max} 时，塑性变形已贯穿整个缺口截面，缺口根部开始横向收缩（相当于颈缩变形），承载面积减小，试样承载能力降低，载荷下降。在 P_{max} 附近试样内部萌生裂纹，视材料韧性情况，裂纹可能萌生于 P_{max} 之前，也可能在 P_{max} 之后。缺口根部为三向应力状态，应力最大值不在缺口根部表面，而是在试样内部距缺口根部一定距离处，因而裂纹萌生于距缺口一定距离的试样内部。裂纹形成以后，向两侧宽度方向和前方深度方向扩展，其机制遵循微孔聚集型断裂规律。在裂纹扩展过程中，载荷继续下降，载荷达 P_F 时，裂纹已扩展到缺口根部的整个宽度。因试样中部约束较强，裂纹扩展较快，形成缺口前方的脚跟形纤维区。随着裂纹尺寸增大，裂纹在 P_F 点开始失稳扩展，形成试样中心的结晶状断口区，呈放射状特征，与此对应的载荷陡降到 P_D。此时裂纹前沿已进入试样的压应力区，尚未断裂的截面积已比较小，与两侧一样已处在平面应力状态下，变形比较自由，形成二次纤维区和剪切唇，相应的载荷由 P_F 降到零。研究表明，试样背面横向扩展量，缺口根部横向收缩量以及剪切唇的厚度都是衡量材料韧性的参数。

根据对断裂过程的分析，可将冲击吸收功分为 A_e，A_p 和 A_d，如图6.4-20所示。可以近似认为，A_e 为弹性变形功，A_p 为塑性变形、形变强化以及裂纹形成等过程吸收的功，A_d 为裂纹扩展功。不同材料，或相同材料但试样不同，各阶段吸收的功的相对比例不同。因此，有时尽管冲击吸收功相同，但断裂的物理过程不同，并由此而引起对材料评定的差异。这也是冲击吸收功不能用作定量设计指标的原因。

在冲击吸收功中，只有裂纹形成和扩展功表示材料的韧性。但在计算冲击韧度时，将冲击吸收功除以缺口截面有效面积，这是缺乏科学依据的。因此认为冲击韧度 a_k 是没有物理意义的，也正是由于此，a_k 值不可能作为定量性能指标用于设计。目前，在材料评定中，较多采用冲击吸收功 A_{kv} 或 A_{ku}，冲击吸收功表示在一定条件下冲断试样所消耗的功，可以相对比较材料的缺口敏感性。还可利用冲击断口上的结晶区面积的比例表示材料的脆性倾向或评定材料冶金缺陷的严重程度。

编写：何雪浤（东北大学）

第 5 章　材料疲劳强度设计

1　概述

1.1　疲劳的分类

根据研究对象、载荷条件、环境和介质情况，疲劳有多种分类方法。常用的疲劳分类见表 6.5-1。

1.2　疲劳强度设计方法

疲劳强度设计方法有名义应力法、局部应力应变法、损伤容限设计法和概率疲劳设计法等。

表 6.5-1　疲劳分类

分类	名称	特征说明	举例
按研究对象分	材料疲劳	通过标准试样研究材料的失效机理、化学成分和微观组织对疲劳强度的影响，疲劳试验方法和数据处理方法，材料的基本疲劳特性，环境和工况的影响，疲劳断口的宏观和微观形貌等	用国家规定的标准试样做的各种疲劳试验
	结构疲劳	以零部件、接头以至整机为研究对象，研究其疲劳性能、抗疲劳设计方法、寿命估算方法、疲劳试验方法，以及形状、尺寸、表面状态和工艺因素的影响，提高其疲劳强度方法等	各种工程结构，机械零部件等
按失效周次分	高周疲劳	材料或结构在低于其屈服强度的循环应力作用下，经过 $10^4 \sim 10^5$ 次以上的循环产生的失效。高周疲劳一般应力较低，材料处于弹性范围内，其应力应变是成比例的，也称应力疲劳，它是机械中最常见的疲劳	弹簧、轴、螺栓等
	低周疲劳	材料或构件在接近或超过其屈服强度的循环应力作用下，在低于 $10^4 \sim 10^5$ 次塑性应变循环产生的失效。由于其应力超过弹性极限，产生较大塑性变形，应力应变不成比例，其主要参数是应变，也常称为应变疲劳	高压容器，汽轮机转子，飞机起落架等
按载荷条件分	随机疲劳	应力幅和频率都随时间变化的疲劳	如汽车底盘、半轴悬挂系统等零件
	冲击疲劳	小能量多次冲击引起的疲劳	内燃机阀杆等
	接触疲劳	零件接触表面在接触压力循环作用下出现麻点、剥落或表层压碎剥落，从而造成零件失效的疲劳	如齿轮传动、滚动轴承、车轮等
	微动磨损疲劳	当两零件表面相接触，并作小幅度的往复相对运动时，在接触表面上产生的疲劳，经过附着、氧化、疲劳三个阶段，是机械过程和化学过程综合的结果	如铆钉联接件、螺栓联接件、紧配合件、销钉、花键、键联接等
	声疲劳	由气体动力噪声、结构噪声或电磁噪声等噪声使结构件产生的疲劳。只有当作为激振力的噪声使结构件产生的应力—应变响应足够大，足以对结构材料造成疲劳损伤时才可能产生声疲劳	如火箭和飞机的涡轮发动机作为噪声源，使飞行器和机翼表面产生高声压水平的噪声场，足以对其结构的局部危险区造成声疲劳
按温度环境分	高温疲劳	在高温环境下零件承受循环载荷发生的疲劳。高温指约在 0.5 T_m 或再结晶温度以上，T_m 为以热力学温度表示的金属熔点。高温疲劳是疲劳与蠕变共同作用结果	如燃气轮机的叶片由机械振动发生的高温高周疲劳；燃气轮机转子由装置的起动和停车而发生的高温低周疲劳等
	低温疲劳	在低于室温环境下零件承受循环应力作用发生的疲劳	寒冷地区露天机械结构产生的疲劳
	热疲劳	由温度循环变化而引起应变循环变化产生的疲劳	锅炉水冷壁管子因冷水分层现象使管子发生的疲劳
	腐蚀疲劳	在腐蚀介质（如酸、碱、海水、淡水、活性气体等）和循环载荷联合作用下产生的疲劳	如化工机械、石油机械某些零件，在酸、碱液体和气体中工作等

1.2.1 名义应力法

名义应力法是以材料的 $S-N$ 曲线为依据，以零构件的名义应力为设计参数，计人了有效应力集中系数 K_σ、零件尺寸系数 ε、表面系数 β 和平均应力影响系数 ψ_σ 等因素，得到零件的 $S-N$ 曲线，依此进行疲劳强度设计。

当 $S-N$ 曲线的纵轴 σ 和横轴 N 都取对数时，则成为如图6.5-1所示的以 P 为交点的两条直线段组成的折线。对于钢材，交点的循环次数（或横坐标）$N_0 \approx 10^7$。N_0 称为循环基数。

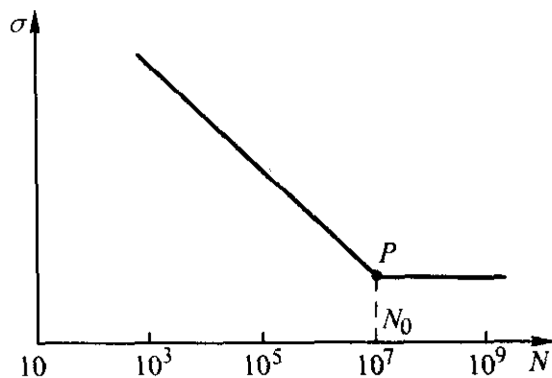


图 6.5-1 取对数坐标的 $\sigma-N$ 曲线

根据平行于横轴的直线进行设计称为无限寿命设计。根据左边斜线进行设计称为有限寿命设计。名义应力法通常也称为常规疲劳设计法或影响系数法。详见本章第4节。

1.2.2 局部应力应变法

以零构件应变集中处的局部应力、应变为基本设计参数的一种疲劳强度设计方法。其基本思路：零构件的破坏都是从应变集中部位的最大应变集中处开始。应变集中处的塑性变形是疲劳裂纹形成和扩展的先决条件，因此应变集中处的局部最大应变决定了零构件的疲劳强度和寿命。对于同一种材料，只要其局部最大应力、应变相同，疲劳寿命就相同。

根据相同应变条件下损伤相等原则，可以用光滑试样的应变－寿命曲线估算零构件危险部位的损伤，从而得到零构件疲劳裂纹形成的寿命。详见本章第5节。

1.2.3 损伤容限设计法

这种设计方法允许零构件内存在有初始裂纹，应用断裂力学方法来估算其剩余寿命，并通过试验来检验，确保其在使用期内裂纹不致于扩展到引起破坏的程度，保证有裂纹零构件在服役期内的安全。

1.2.4 概率疲劳设计法

根据零构件的工作应力与疲劳强度相联系的统计方法而进行的疲劳强度设计方法，是概率统计方法与疲劳强度设计相结合的一种设计方法，也称疲劳可靠性设计。

2 疲劳强度设计中的参量及处理

2.1 疲劳载荷

载荷可分为两大类，即静载荷和动载荷。动载荷又分为周期载荷、非周期载荷和冲击载荷。周期载荷和非周期载荷统称为疲劳载荷。

一般机器和零件承受的载荷，大都是一个连续的随机载荷。承受随机载荷的零件，在进行疲劳强度计算、寿命估算和疲劳试验之前，必须先确定其载荷谱。在机器工作时直接测得的载荷－时间历程称为工作谱或使用谱，由于随机载荷的不确定性，这种谱无法使用，必须对它进行处理。经过处理后的载荷－时间历程称为载荷谱，该载荷谱具有统计特性，它能本质地反映零件的载荷变化情况。将实测的载荷－时间历程处理成具有代表性的典型载荷谱的过程称为编谱，编谱的重要环节是应用统计理论来处理所获得的实测子样。

统计处理分析随机载荷的方法主要有：循环计数法和功率谱法。循环计数法是从载荷－时间历程中确定出不同载荷参量值及其出现的次数。功率谱法是借助傅氏变换，将连续变化的随机载荷分解为无限多个具有各种频率的简单变化，得出其功率谱密度函数。

对于疲劳强度来说，最主要的是载荷幅值的变化情况，故广泛使用循环计数法。

2.1.1 循环应力和循环应变

1）循环应力　最简单的循环应力为恒幅循环应力。图6.5-2是四种不同的应力变化规律。

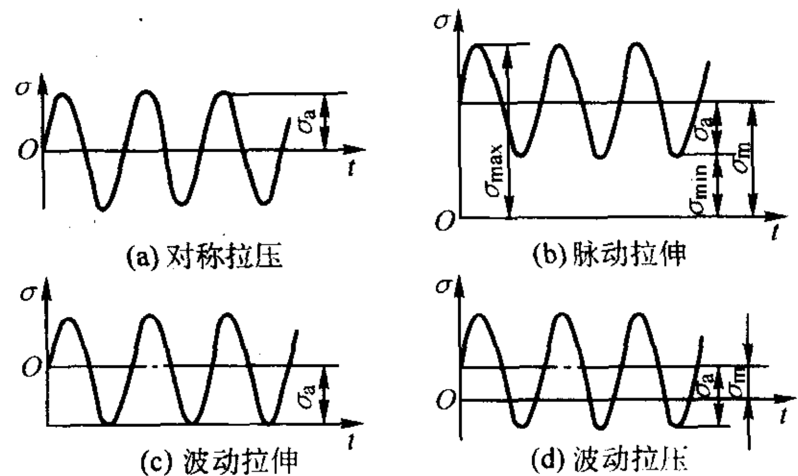


图 6.5-2 恒幅循环应力的种类

图中 σ 为正应力，t 为时间。各应力分量：σ_{max} 为应力循环中具有最大代数值的应力；σ_{min} 为应力循环中具有最小代数值的应力；σ_m 为应力循环中最大应力和最小应力的代数平均值；σ_a 为应力循环中最大应力和最小应力代数差的一半；规定拉应力为正，压应力为负。平均应力 σ_m、应力幅 σ_a、与最大应力 σ_{max}、最小应力 σ_{min} 之间有如下关系

$$\sigma_m = \frac{\sigma_{max} + \sigma_{min}}{2} \tag{6.5-1}$$

$$\sigma_a = \frac{\sigma_{max} - \sigma_{min}}{2} \tag{6.5-2}$$

$$\sigma_{max} = \sigma_m + \sigma_a \tag{6.5-3}$$

$$\sigma_{min} = \sigma_m - \sigma_a \tag{6.5-4}$$

应力每一周期性变化称为一个应力循环。定义应力比 r 为

$$r = \frac{\sigma_{min}}{\sigma_{max}} \tag{6.5-5}$$

对于对称循环，$r=-1$；对于脉动循环，$r=0$；静应力可以看作应力幅为零的循环应力，此时 $r=+1$。任何一个应力循环的应力比都可以在 $-1 \leqslant r \leqslant +1$ 范围内取值。

一种循环应力状态，一般可用 σ_{max}、σ_{min}、σ_m、σ_a 和 r 五个参数中的任意两个来确定。如果作用的应力是切应力时，各应力分量之间的关系有

$$\tau_m = \frac{\tau_{max} + \tau_{min}}{2} \tag{6.5-6}$$

$$\tau_a = \frac{\tau_{max} - \tau_{min}}{2} \tag{6.5-7}$$

$$\tau_{max} = \tau_m + \tau_a \tag{6.5-8}$$

$$\tau_{min} = \tau_m - \tau_a \tag{6.5-9}$$

2）循环应变　由疲劳试验得到的应力－寿命曲线中，当循环加载的应力水平较低时，疲劳的全过程中弹性应变起主导作用，这时的应力－寿命曲线（$\sigma-N$）能正确反映出应力与寿命之间的关系。当应力水平较高时，塑性应变起主导作用，高应力水平部分达到屈服应力，应力－寿命曲线出

现平坦部分，见图 6.5-3a。此时的应力不能描述实际寿命的变化，需要用应变 ε 来代替应力 σ，即成为应变－寿命（$\varepsilon-N$）曲线，见图 6.5-3b。把应变随时间的变化称为循环应变。与上述循环应力相同，循环应变也有最大应变 ε_{max}，最小应变 ε_{min}，平均应变 ε_m，应变幅 ε_a 和循环特性参数 r。

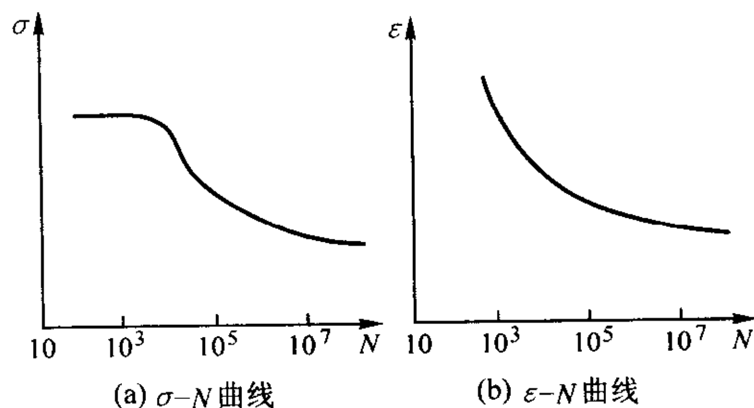


图 6.5-3　用 ε 代替 σ 的 $S-N$ 曲线

2.1.2　循环计数法

把一个随机的载荷－时间历程处理成一系列的全循环或半循环的过程称为循环计数法。将其分成两大类：单参数计数法和双参数计数法。单参数计数法只记录载荷谱中的一个参量，如峰值或范围，不能给出载荷循环的全部信息。属于这种计数方法有：峰值计数法，穿级计数法和范围计数法等。双参数计数法可以记录载荷循环中的两个参量。由于载荷循环中只有两个独立参量，因此双参数计数法可以记录载荷循环的全部信息，是一种较好的计数方法。属于这种计数方法有：范围对计数法，跑道计数法和雨流计数法等。使用最广泛的是雨流计数法。该法在计数原理上有一定的力学依据，并具较高的正确性，也易于实现自动化程序化。

雨流法的计数原理如下。

如图 6.5-4 所示。对一个实际的载荷时间历程，取一垂直向下的纵坐标轴表示时间，横坐标轴表示载荷。这样载荷时间历程形同一座宝塔，雨点以峰值、谷值为起点向下流动，根据雨点向下流动的迹线，确定载荷循环，这就是雨流法（或称塔顶法）名称的由来。其计数规则如下。

1）雨流的起点依次在每个峰（谷）值的内侧开始。

2）雨流在下一个峰（谷）值处落下，直到对面有一个比开始时的峰（谷）值更大（更小）值为止。

3）当雨流遇到来自上面屋顶流下的雨时就停止。

4）取出所有的全循环，并记下各自的振程。

5）按正、负斜率取出所有的半循环，并记下各自的振程。

6）把取出的半循环按雨流法第二阶段计数法则处理并计数。

根据上述规则，图 6.5-4 中的第 1 个雨流应从 O 点开始，流到 a 点落下，经 b 与 c 之间的 a' 点继续流到 c 点落下，最后停止在比谷值 O 更小的谷值 d 的对应处。取出一个半循环 $O-a-a'-c$。第二个雨流从峰值 a 的内侧开始，由 b 点落下，由于峰值 c 比 a 大，故雨流停止于 c 的对应处，取出半循环 $a-b$。第三个雨流从 b 点开始流下，由于遇到来自上面的雨流 $O-a-a'$，故止于 a'点，取出半循环 $b-a'$。因 $b-a'$与 $a-b$ 构成闭合的应力－应变回线，则形成一个全循环 $a'-b-a$。依次处理，最后可以得到在图 6.5-4a 所示的载荷－时间历程中三个全循环：$a'-b-a$，$d'-c-d$，$g'-h-g$ 和三个半循环，$O-a-a'-c$，$c-d-d'-f$，$f-g-g'-i$。

图 6.5-4b 是该载荷历程作用下的材料应力－应变回线，可见与雨流法计数所得结果是一致的。

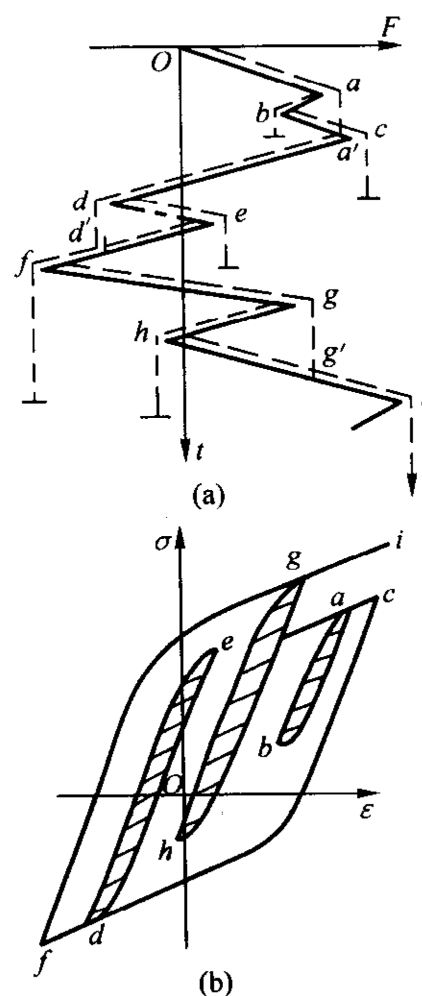


图 6.5-4　雨流法计数原理图

一个实际的载荷时间历程，经过雨流法计数并取出全循环之后，剩下的半循环构成了一个发散－收敛的载荷谱，按上述雨流法规则无法继续计数。如把它改造一下使之变成一个收敛－发散谱后，就可继续用雨流法计数，这就是雨流法计数第二阶段。

图 6.5-5a 为一发散－收敛谱，从最高峰值 a_1 或最低谷值 b_1 处截成两段，使左段起点 b_n 和右段末点 a_n 相连接，构成如图 6.5-5b 那样的发散－收敛谱，则可继续用雨流法计数直到完毕。

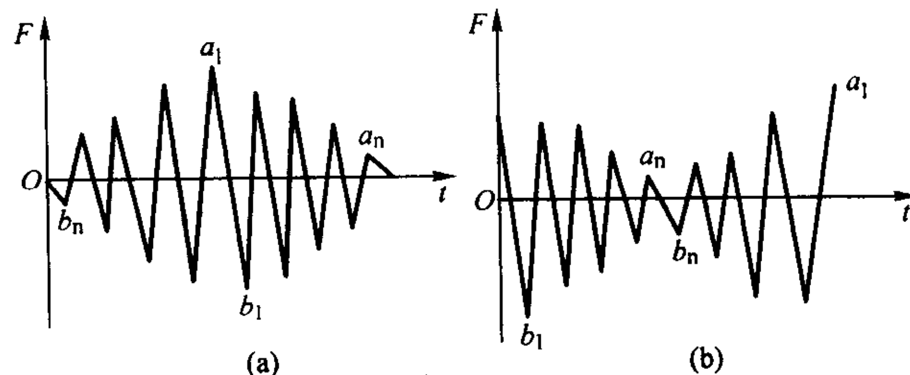


图 6.5-5　雨流法第二阶段计数原理图

2.2　材料疲劳性能数据

2.2.1　材料疲劳极限

对于结构钢和钛合金等材料，试验得到的 $S-N$ 曲线上有一平行于横坐标的水平区段，与此水平线段相应的最大应力 σ_{max} 称为该材料的疲劳极限，简称疲劳极限。

对于有色金属和在腐蚀疲劳条件下的试验，在 $S-N$ 曲线上没有水平区段，因此不存在疲劳极限。在一般情况下，以 10^7 或 10^8 次循环失效时的最大应力 σ_{max} 作为条件疲劳极限。这时的失效循环数称为循环基数。

表 6.5-2～表 6.5-4 是国产材料的疲劳极限。表 6.5-5 是铝合金材料疲劳极限。

表6.5-2 常用国产机械材料的旋转弯曲疲劳极限

序号	材 料	热处理	抗拉强度 σ_b/MPa	疲劳极限（$N=10^7$）			疲劳比 f（$N=10^7$）
				平均值 σ_{-1} /MPa	标准差 S_{-1} /MPa	变异系数 ν_{-1}	
1	Q235A	热轧	439	210	7.8	0.037	0.048
2	Q235A（F）	热轧	428	198	9.4	0.047	0.46
3	Q235B	热轧	441	250	3.9	0.016	0.57
4	20	正火	463	250	4.7	0.019	0.54
5	20g	热轧	432	209	2.6	0.012	0.48
6	20R	—	386	209	—	—	0.54
7	35	正火	593	261	4.1	0.016	0.44
8	45	正火	624	285.1	7.0	0.026	0.43
9	45	调质	710	388	10.1	0.026	0.53
10	45	电渣重熔	934	433	19.5	0.048	0.43
11	50	正火	661	278	10.3	0.037	0.42
12	55	调质	834	386	13.3	0.034	0.46
13	70	淬火后中温回火	1 138	489	17.9	0.037	0.43
14	Q345	热轧	586	298.1	8.6	0.031	0.48
15	Q345g	热轧	507	271	—	—	0.53
16	20MnVB	碳氮共渗	1 210	809	0.6	0.001	0.67
17	25MnTiBRE	碳氮共渗	1 193	834	23.2	0.028	0.70
18	35Mn2	调质	937	520	—	—	0.55
19	40MnB	调质	970	436	19.5	0.045	0.45
20	40MnVB	调质	1 111	531	9.0	0.017	0.48
21	45Mn2	调质	952	485	8.0	0.016	0.51
22	12Cr2Ni4	调质	793	441	22.6	0.051	0.56
23	18CrNiW	调质	1 039	491	23.1	0.047	0.47
24	20Cr	渗碳	577	273	3.9	0.014	0.47
25	20CrMnTi	淬火后低温回火	1 416	566	37.4	0.066	0.40
26	20CrMnSi	调质	788	299	13.7	0.046	0.38
27	20Cr2Ni4A	淬火后低温回火	1 483	602	14.1	0.023	0.41
28	30CrMnTi	碳氮共渗	1 771	730	35.3	0.048	0.41
29	30CrMnSiA	调质	1 110	641	25.2	0.039	0.58
30	35CrMo	调质	924	431	13.3	0.031	0.47
31	40Cr	调质	940	422	10.1	0.024	0.45
32	40CrMnMo	调质	977	470	17.2	0.037	0.48
33	40CrMnSiMoVA	淬火后低温回火	1 843	677	—	—	0.37
34	40CrNiMo	调质	972	498	7.8	0.016	0.51
35	40CrNiMoA	调质	1 040	524	19.7	0.038	0.50
36	42CrMo	调质	1 134	504	12.5	0.025	0.44
37	16MnCr5	淬火后低温回火	1 373	592	10.9	0.018	0.43
38	20MnCr5	淬火后低温回火	1 482	634	8.0	0.013	0.43
39	25MnCr5	淬火后低温回火	1 587	509	37.4	0.074	0.32
40	28MnCr5	淬火后低温回火	1 307	479	21.1	0.044	0.37
41	50CrV	淬火后中温回火	1 586	747	32.0	0.043	0.47
42	55Si2Mn	淬火后中温回火	1 866	658	10.5	0.016	0.35
43	60Si2Mn	淬火后中温回火	1 625	660	24.2	0.043	0.41
44	65Mn	淬火后中温回火	1 687	708	31.1	0.044	0.42
45	0Cr17Ni4Cu4Nb	固溶时效	740	400	—	—	0.54

续表 6.5-2

序号	材　料	热处理	抗拉强度 σ_b/MPa	疲劳极限（$N=10^7$）			疲劳比 f（$N=10^7$）
				平均值 σ_{-1} /MPa	标准差 S_{-1} /MPa	变异系数 ν_{-1}	
46	1Cr12Mo	调质	768	382	—	—	0.50
47	1Cr13	调质	721	374	12.5	0.033	0.52
48	2Cr13	调质	687.5	374	14.0	0.038	0.54
49	3Cr13	调质	842	370	12.5	0.034	0.44
50	4Cr5MoVSi	调质	1 496	730	—	—	0.49
51	7Cr7Mo3V2Si	调质	2 353	512	24.2	0.047	0.22
52	Cr12	淬火后低温回火	2 272	709	20.4	0.029	0.31
53	Cr12MoV	淬火后低温回火	2 059	633	—	—	0.31
54	ZG20SiMn	正火	515	226	7.5	0.033	0.44
55	ZG230－450	正火	543	207	9.4	0.045	0.38
56	ZG270－500	调质	823	272	5.5	0.020	0.33
57	ZG40Cr	调质	977	294	10.9	0.037	0.30
58	ZG340－640	调质	1 044	322	12.6	0.039	0.31
59	ZG0Cr13Ni6Mo	正火后两次回火	779	289	16.8	0.058	0.37
60	ZG1Cr13	退火后正火	789	328	14.8	0.045	0.42
61	QT400－15	退火	484	243	10.9	0.045	0.50
62	QT400－18	退火	453	219	7.4	0.034	0.48
63	QT500－7	退火	625	206	10.9	0.053	0.33
64	QT600－3	正火	809	271	7.4	0.027	0.33
65	QT700－2	正火	754	219	9.9	0.045	0.29
66	QT800－2	正火	842	352	10.1	0.029	0.42

注：S_{-1}和 ν_{-1}为对称循环下疲劳的标准差和变异系数。

表 6.5-3　某些国产机械材料的拉－压疲劳极限

序号	材　料	热处理	抗拉强度 σ_b/MPa	疲劳极限（$N=10^7$）			疲劳比 f（$N=10^7$）
				平均值 σ_{-1l} /MPa	标准差 S_{-1l} /MPa	变异系数 ν_{-1l}	
1	20	正火	464	241	7.8	0.032	0.52
2	45	调质	735	329①	18.7	0.057	0.45
3	Q345	热轧	586	327	14.0	0.043	0.56
4	09SiVL	热轧	529	284	13.3	0.047	0.54
5	12CrNi3	调质	833	363	14	0.039	0.44
6	25Cr2MoV	调质	1 090	335	—	—	0.31
7	35CrMo	调质	924	317	—	—	0.34
8	35VB	热轧	741	331	13.3	0.040	0.45
9	40CrMnSiMoVA	等温淬火	1 765	718	—	—	0.41
10	40CrNiMo	调质	927	389	15.6	0.040	0.40
11	45CrNiMoV	淬火后中温回火	1 553	486	17.2	0.035	0.31
12	55SiMnVB	淬火后中温回火	1 536	536	21.1	0.039	0.35
13	HT200	去应力退火	250	96.5	5.4	0.056	0.39
14	HT300	去应力退火	353	133.3	5.0	0.38	0.38
15	ZG310－570	调质	1 012	303	17.2	0.057	0.30

注：1. 应力比 $r=0.1$。
2. S_{-1l}和 ν_{-1l}为拉－压时对称循环疲劳的标准差。

表 6.5-4 调质结构钢的疲劳极限

材 料	静强度指标	试验条件		寿命	疲劳极限均值	标准差	变异系数
		r	α_σ	N	$\bar{\sigma}_r$/MPa	S_r/MPa	$\nu = S_r\sqrt{\sigma_r}$
45（调质）	σ = 833.6 MPa σ_s = 686.5 MPa δ = 16.7% 硬度 250 ~ 270 HBS	-1	1.9	5×10^4	411.9	13.07	0.031 73
				10^5	343.2	9.807	0.028 58
				5×10^5	309.9	7.845	0.025 31
				10^6	294.2	7.845	0.026 67
				5×10^6	286.4	7.845	0.027 39
				10^7	279.5	8.169	0.029 23
18Cr2Ni4WA (950℃正火， 860℃淬火， 540℃回火)	σ_b = 1 145.5 MPa δ = 18.6%	-1	2	10^5	463.9	22.23	0.047 92
				5×10^5	411.9	17.00	0.041 27
				10^6	384.4	15.69	0.040 82
				5×10^6	368.7	13.73	0.037 24
				10^7	360.9	11.77	0.032 61
30CrMnSiA (890 ~ 989℃ 油淬火， 510 ~ 520℃ 回火)	σ_b = 1 108.2 ~ 1 186.6 MPa σ_s = 1 088.6 MPa δ = 15.3% ~ 18.6%	-1	1	10^5	784.6	35.96	0.045 83
				5×10^5	676.7	19.61	0.028 98
				10^6	655.1	17.65	0.026 94
				5×10^6	639.4	17.00	0.026 59
				10^7	637.5	18.63	0.029 22
30CrMnSiA (890 ~ 989℃ 油淬火， 510 ~ 520℃ 回火)	σ_b = 1 108.2 ~ 1 186.6 MPa σ_s = 1 088.6 MPa δ = 15.3% ~ 18.6%	-1	2	10^5	411.3	19.61	0.047 68
				5×10^5	379.5	14.71	0.038 76
				10^6	359.9	10.13	0.028 15
				5×10^6	356.0	10.13	0.028 46
				10^7	353.1	9.807	0.027 77
			3	10^5	308.9	14.71	0.047 62
				5×10^5	270.7	10.13	0.037 42
				10^6	250.1	9.807	0.039 21
				5×10^6	243.2	9.150	0.037 62
				10^7	241.3	9.150	0.037 92
			4	10^5	285.4	11.11	0.038 93
				5×10^5	245.2	9.807	0.035 00
				10^6	221.6	9.150	0.041 29
				5×10^6	210.9	8.169	0.038 73
				10^7	204.0	6.865	0.033 65
30CrMnSiA (890 ~ 989℃ 油淬火， 510 ~ 520℃ 回火)	σ_b = 1 108.2 ~ 1 186.6 MPa σ_s = 1 088.6 MPa δ = 15.3% ~ 18.6%	0.1	1	10^5	1 176.8	52.30	0.044 44
				5×10^5	1 108.2	42.49	0.038 34
				10^6	1 090.5	39.23	0.035 97
				5×10^6	1 088.6	39.55	0.036 33
				10^7	1 088.6	39.89	0.036 64
			3	10^5	455.0	29.42	0.064 66
				5×10^5	377.6	17.00	0.045 02
				10^6	347.2	14.39	0.041 45
				5×10^6	335.4	15.69	0.046 78
				10^7	328.5	16.35	0.049 77
		0.5	3	10^5	676.7	35.96	0.053 14
				5×10^5	642.4	31.06	0.048 35
				10^6	612.0	27.46	0.044 87
				5×10^6	609.0	24.84	0.040 79
				10^7	608.0	24.84	0.040 86

续表 6.5-4

材　料	静强度指标	试验条件		寿命	疲劳极限均值	标准差	变异系数
		r	α_σ	N	$\overline{\sigma}_r$/MPa	S_r/MPa	$\nu = S_r\sqrt{\sigma_r}$
30CrMnSiNi2A（900℃淬火，260℃回火）	σ_b = 1 422 ~ 1 618 MPa σ_s = 1 109 MPa δ = 12.5% ~ 18.5%	−0.5	5	5×10^4	415.8	20.92	0.050 31
				10^5	343.2	13.73	0.040 01
				5×10^5	272.6	10.46	0.038 37
				10^6	251.1	9.150	0.036 44
				5×10^6	248.1	9.150	0.036 88
				10^7	245.2	9.807	0.040 00
		0.1	3	10^4	662.0	33.02	0.049 88
				5×10^4	539.4	26.80	0.049 68
				10^5	441.3	17.98	0.040 74
				5×10^5	415.8	16.67	0.040 09
				10^6	402.1	16.35	0.040 66
				5×10^6	392.3	15.69	0.039 99
				10^7	382.5	14.71	0.038 46
			4	10^4	686.5	49.04	0.071 43
				5×10^4	510.0	29.42	0.057 69
				10^5	328.5	17.98	0.054 73
				5×10^5	241.3	9.150	0.037 92
				10^6	187.3	6.865	0.036 65
		0.445	3	10^4	1 059.2	58.84	0.055 55
				5×10^4	858.1	34.32	0.040 00
				10^5	686.5	27.78	0.040 47
				5×10^5	583.5	20.59	0.035 29
				10^6	578.6	20.27	0.035 03
				5×10^6	572.7	19.29	0.033 68
				10^7	571.7	18.96	0.033 16
		0.5	5	5×10^4	731.6	29.74	0.040 65
				10^5	624.7	26.16	0.041 88
				5×10^5	525.7	18.31	0.034 83
				10^6	517.8	17.33	0.033 47
				5×10^6	513.9	16.67	0.032 44
				10^7	510.0	16.35	0.032 06
40CrNiMoA（850℃油淬火，580℃回火）	σ_b = 1 040 ~ 1 167 MPa σ_s = 917 ~ 1 126 MPa δ = 15.6% ~ 17%	−1	1	5×10^4	760.0	44.13	0.058 07
				10^5	666.9	37.59	0.056 37
				5×10^5	590.4	26.16	0.044 31
				10^6	559.0	20.92	0.037 42
				5×10^6	539.4	20.92	0.038 78
				10^7	523.7	19.61	0.037 45
			2	10^5	392.3	25.17	0.064 16
				5×10^5	333.4	14.05	0.042 14
				10^6	318.7	11.44	0.035 90
				5×10^6	310.9	10.46	0.033 64
				10^7	307.9	9.807	0.031 85
			3	10^5	294.2	15.03	0.051 09
				5×10^5	245.2	9.807	0.040 00
				10^6	217.7	8.169	0.037 52
				5×10^6	210.9	6.865	0.032 55
				10^7	208.9	6.865	0.032 86

续表 6.5-4

材 料	静强度指标	试验条件		寿命	疲劳极限均值	标准差	变异系数
		r	α_σ	N	$\overline{\sigma}_r$/MPa	S_r/MPa	$\nu = S_r \sqrt{\sigma_r}$
40CrNiMoA (850℃油淬火，580℃回火)	σ_b = 1 040 ~ 1 167 MPa σ_s = 917 ~ 1 126 MPa δ = 15.6% ~ 17%	0.1	1	5×10^4	1 259.2	60.15	0.047 77
				10^5	1 211.2	45.77	0.037 79
				5×10^5	1 157.2	42.49	0.036 72
				10^6	1 110.2	39.89	0.035 93
				5×10^6	1 066.0	38.25	0.035 88
				10^7	1 029.7	32.69	0.031 75
			3	5×10^4	490.4	22.88	0.046 66
				10^5	384.4	17.65	0.045 92
				5×10^5	326.6	11.44	0.035 03
				10^6	305.0	10.79	0.035 38
				5×10^6	292.2	10.79	0.036 93
				10^7	284.4	9.807	0.034 48
42CrMnSiMoA (GC-4电渣钢) (920℃加热，300℃等温，空冷)	σ_b = 1 894 MPa σ_s = 1 388 MPa δ = 13%	-1	1	5×10^4	965.0	65.38	0.067 75
				10^5	874.8	49.69	0.056 80
				5×10^5	799.3	38.25	0.047 85
				10^6	761.0	29.42	0.038 66
				5×10^6	735.5	26.80	0.036 44
		-1	1	10^7	717.9	24.84	0.034 60
			3	10^4	513.9	45.44	0.088 42
				5×10^4	421.7	32.04	0.075 98
				10^5	373.6	18.31	0.049 01
				5×10^5	323.6	13.07	0.040 39
				10^6	284.4	11.44	0.040 23
				5×10^6	251.1	9.807	0.039 06
				10^7	239.3	9.150	0.038 24
		0.1	1	5×10^4	1 216.1	65.38	0.053 76
				10^5	1 118.0	52.30	0.046 78
				5×10^5	1 074.8	41.19	0.038 32
				10^6	1 069.0	39.23	0.036 70
				5×10^6	1 067.0	39.23	0.036 77
				10^7	1 065.0	38.57	0.036 22
			3	10^4	672.8	33.02	0.049 08
				5×10^4	555.1	26.48	0.047 70
				10^5	485.4	18.63	0.038 38
				5×10^5	460.9	16.35	0.035 47
				10^6	447.2	16.67	0.037 28
				5×10^6	433.5	15.69	0.036 19
				10^7	427.6	15.03	0.035 15

注：S_r 为循环特性为 r 条件下疲劳的标准差。

表 6.5-5　铝合金的疲劳极限

材　料	静强度指标	试验条件 r	试验条件 α_σ	寿命 N	疲劳极限均值 $\bar{\sigma}_r$/MPa	标准差 S_r/MPa	变异系数 $\nu = S_r\sqrt{\sigma_r}$
2A12B（"B"为预拉伸加工硬化）	$\sigma_b = 455 \sim 480$ MPa $\sigma_s = 343 \sim 438$ MPa $\delta = 8\% \sim 19\%$	0.1	1	10^4	411.9	22.88	0.055 55
				5×10^4	369.7	18.63	0.050 39
				10^5	329.5	13.41	0.040 70
				5×10^5	293.2	11.77	0.040 14
				10^6	264.8	9.807	0.037 04
				5×10^6	243.2	9.150	0.037 62
				10^7	223.6	7.522	0.033 64
			3	10^4	245.2	13.07	0.053 30
				5×10^4	191.2	9.150	0.047 86
				10^5	161.8	7.522	0.046 49
				5×10^5	134.4	5.227	0.038 89
				10^6	114.7	4.246	0.037 02
				5×10^6	106.9	3.923	0.036 70
				10^7	103.0	3.599	0.034 94
			5	10^4	194.2	9.150	0.047 12
				5×10^4	148.1	6.541	0.044 17
				10^5	120.6	4.904	0.040 66
				5×10^5	99.05	3.923	0.039 61
				10^6	87.28	3.266	0.037 42
				5×10^6	84.34	3.266	0.038 72
				10^7	82.38	2.941	0.035 70
		0.5	1	5×10^4	459.0	21.58	0.047 02
				10^5	405.0	17.33	0.042 79
				5×10^5	360.9	15.03	0.041 65
				10^6	347.2	13.73	0.039 54
				5×10^6	328.5	12.09	0.036 80
				10^7	319.7	11.77	0.036 82
			3	10^4	343.2	16.35	0.047 64
				5×10^4	268.7	11.77	0.043 80
				10^5	211.8	8.826	0.041 67
				5×10^5	169.7	6.541	0.038 54
				10^6	151.0	5.227	0.034 62
				5×10^6	145.1	5.227	0.036 02
				10^7	143.2	4.904	0.034 25
			5	10^4	299.1	14.71	0.049 18
				5×10^4	222.6	10.46	0.046 99
				10^5	161.8	6.541	0.040 43
				5×10^5	129.4	5.227	0.040 39
				10^6	115.7	4.246	0.036 70
				5×10^6	109.8	3.923	0.035 73
				10^7	104.0	2.941	0.028 28
		−0.5	3	10^5	117.5	5.816	0.049 50
				5×10^5	108.5	4.776	0.044 02
				10^6	100.0	3.923	0.039 23
				5×10^6	92.19	3.599	0.039 04
				10^7	87.77	2.942	0.033 52

续表 6.5-5

材料	静强度指标	试验条件		寿命	疲劳极限均值	标准差	变异系数
		r	α_σ	N	$\bar{\sigma}_r$/MPa	S_r/MPa	$\nu = S_r \sqrt{\sigma_r}$
2A12-T4	$\sigma_b = 407$ MPa $\sigma_s = 270$ MPa $\delta = 13\%$	0.1	1.16	10^5	202.0	9.483	0.046 95
				5×10^5	146.1	6.541	0.044 77
				10^6	125.5	4.580	0.036 49
				5×10^6	115.7	4.246	0.036 70
				10^7	110.8	3.923	0.035 41
	$\sigma_b = 457$ MPa $\sigma_s = 336$ MPa $\delta = 18.7\%$	0.02	1	10^5	277.5	14.05	0.050 63
				5×10^5	195.2	8.826	0.045 22
				10^6	144.2	5.561	0.038 56
				5×10^6	132.4	4.580	0.034 59
		0.6	1	5×10^5	331.5	15.69	0.047 33
				10^6	309.9	12.43	0.040 11
				5×10^6	274.6	9.807	0.035 71
2A12-T6	$\sigma_b = 429 \sim 433$ MPa $\sigma_s = 364 \sim 370$ MPa $\delta = 6.6\% \sim 7.8\%$	0.1	1	5×10^4	353.1	21.25	0.060 18
				10^5	240.5	11.44	0.047 57
				5×10^5	176.5	7.189	0.040 73
				10^6	139.3	5.227	0.037 52
				5×10^6	133.4	4.904	0.036 76
				10^7	131.4	4.680	0.035 62
		0.5	1	5×10^4	470.7	22.88	0.048 61
				10^5	372.7	16.35	0.043 87
				5×10^5	304.0	11.77	0.038 72
				10^6	255.0	8.826	0.034 61
				5×10^6	225.6	7.846	0.034 78
				10^7	206.9	6.865	0.033 18
7A09	$\sigma_b = 647$ MPa $\sigma_s = 603$ MPa $\delta = 17.2\%$	-1	1	5×10^4	303.0	14.05	0.046 37
				10^5	261.8	12.75	0.048 70
				5×10^5	220.7	8.826	0.039 99
				10^6	188.3	7.189	0.038 18
				5×10^6	170.6	6.208	0.036 39
				10^7	161.8	5.561	0.034 37
			2.4	5×10^4	187.3	8.826	0.047 12
				10^5	154.0	6.541	0.042 47
				5×10^5	131.4	5.227	0.039 78
				10^6	113.8	4.246	0.037 31
				5×10^6	98.07	3.599	0.036 70
				10^7	93.17	3.266	0.035 05
		0.1	1	10^5	269.7	13.41	0.049 72
				5×10^5	199.1	8.169	0.041 03
				10^6	161.8	5.561	0.034 37
				5×10^6	142.2	4.904	0.034 49
				10^7	130.4	4.246	0.032 56
			3	10^5	124.5	5.884	0.047 26
				5×10^5	93.17	4.246	0.045 57
				10^6	76.49	2.942	0.038 46
				5×10^6	70.61	2.618	0.037 08
				10^7	66.69	2.285	0.034 26

续表 6.5-5

材料	静强度指标	试验条件		寿命	疲劳极限均值	标准差	变异系数
		r	α_σ	N	$\bar{\sigma}_r$/MPa	S_r/MPa	$\nu = S_r\sqrt{\sigma_r}$
7A09	$\sigma_b = 647$ MPa $\sigma_s = 603$ MPa $\delta = 17.2\%$	0.1	5	5×10^4	115.7	6.865	0.059 33
				10^5	81.40	4.246	0.052 16
				5×10^5	63.75	2.618	0.041 07
				10^6	57.86	2.285	0.039 49
				5×10^6	54.92	1.961	0.035 71
				10^7	52.96	1.795	0.033 89
		0.5	1	10^5	431.5	26.16	0.060 63
				5×10^5	262.8	13.41	0.051 03
				10^6	228.5	10.79	0.047 22
				5×10^6	204.0	7.846	0.038 46
				10^7	186.3	5.884	0.031 58
			3	10^5	178.5	9.807	0.054 94
				5×10^5	144.2	6.208	0.043 05
				10^6	127.5	4.904	0.038 46
				5×10^6	116.3	4.119	0.035 42
				10^7	109.8	3.599	0.032 78
			5	5×10^4	166.7	8.169	0.049 00
				10^5	117.7	4.904	0.041 67
				5×10^5	92.19	3.599	0.039 04
				10^6	82.38	3.267	0.039 67
				5×10^6	78.46	2.618	0.033 37
				10^7	76.49	2.618	0.034 23

当缺乏疲劳极限的数值时，可采用经验公式估算。

1）对于结构钢对称循环应力下的疲劳极限

拉压 $\sigma_{-1l} = 0.23\ (\sigma_s + \sigma_b)$

弯曲 $\sigma_{-1} = 0.27\ (\sigma_s + \sigma_b)$

扭转 $\tau_{-1} = 0.15\ (\tau_s + \tau_b)$

2）对于结构钢脉动循环应力下的疲劳极限

拉压 $\sigma_{0l} = 1.42\sigma_{-1}$

弯曲 $\sigma_0 = 1.33\sigma_{-1}$

扭转 $\tau_0 = 1.50\tau_{-1}$

3）对于铸铁的疲劳极限

拉压 $\sigma_{-1l} = 0.4\sigma_b$ $\sigma_{0l} = 1.42\sigma_{-1l}$

弯曲 $\sigma_{-1} = 0.45\sigma_b$ $\sigma_0 = 1.33\sigma_{-1}$

扭转 $\tau_{-1} = 0.36\sigma_b$ $\tau_0 = 1.35\tau_{-1}$

4）对于球墨铸铁的疲劳极限

$\tau_{-1} = 0.26\sigma_b$

5）对于铝合金的疲劳极限

$\sigma_{-1l} = \sigma_b/6 + 75$

$\sigma_{-1} = \sigma_b/6 + 75$

$\sigma_{0l} = 1.5\sigma_{-1l}$

6）对于青铜的弯曲疲劳极限

$\sigma_{-1} = 0.21\sigma_b$

2.2.2 材料的 S－N 曲线

图 6.5-6～图 6.5-52 是金属材料的 S－N 曲线。钢材的图注中 δ 表示板材厚度，ϕ 表示棒材的直径。铝合金尾部字母 B 表示预拉伸加工硬化；T4 表示固溶热处理后自然时效；T6 表示固溶热处理后人工时效。

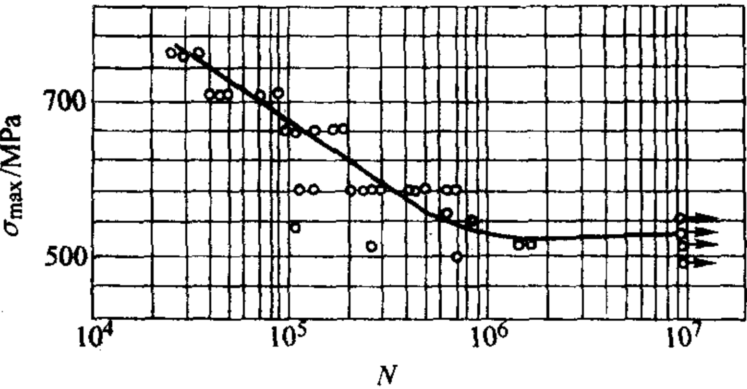


图 6.5-6 40CrNiMoA 钢棒材光滑试样的 S－N 曲线（棒材 ϕ30 mm）

热处理：850℃油淬火，580℃回火

材料 $\sigma_b = 1\ 039$ MPa

悬臂旋转弯曲，$r = -1$

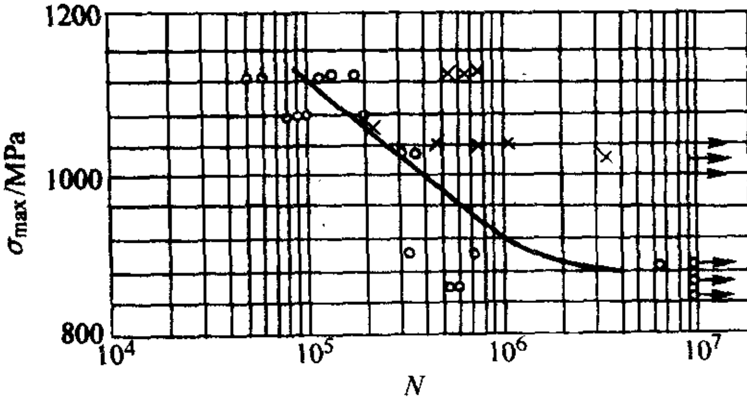


图 6.5-7 40CrNiMoA 钢棒材光滑试样的 S－N 曲线（棒材 ϕ180 mm）

热处理：850℃油淬火，570℃回火

材料 纵向 $\sigma_b = 1\ 167$ MPa，横向 $\sigma_b = 1\ 172$ MPa

轴向加载试验，$r = 0.1$

“×”—纵向，“○”—横向

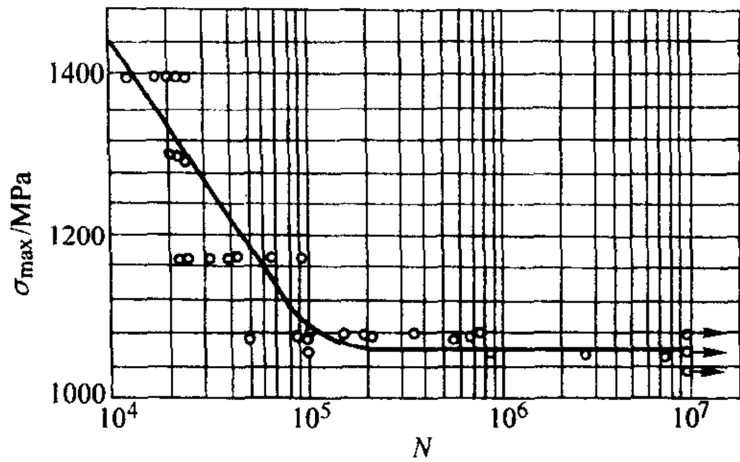


图 6.5-8　40CrMnSiMoA 钢棒材光滑试样的 $S-N$ 曲线
（棒材 $\phi42$ mm）
热处理：920℃加热，300℃等温，空冷
材料　$\sigma_b = 1\ 893$ MPa
轴向加载，$r = 0.1$

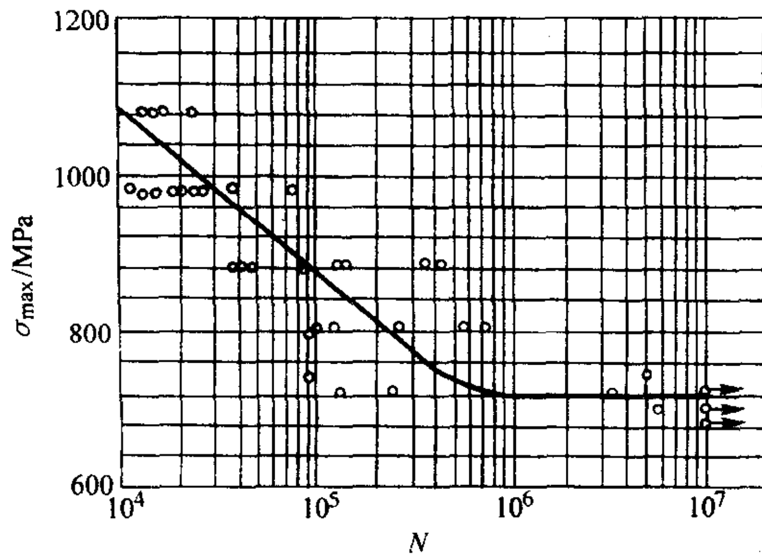


图 6.5-9　40CrMnSiMoA 钢棒材光滑试样的 $S-N$ 曲线
（棒材 $\phi42$ mm）
热处理：920℃加热，300℃等温，空冷
材料　$\sigma_b = 1\ 893$ MPa
轴向加载试验，$r = -1$

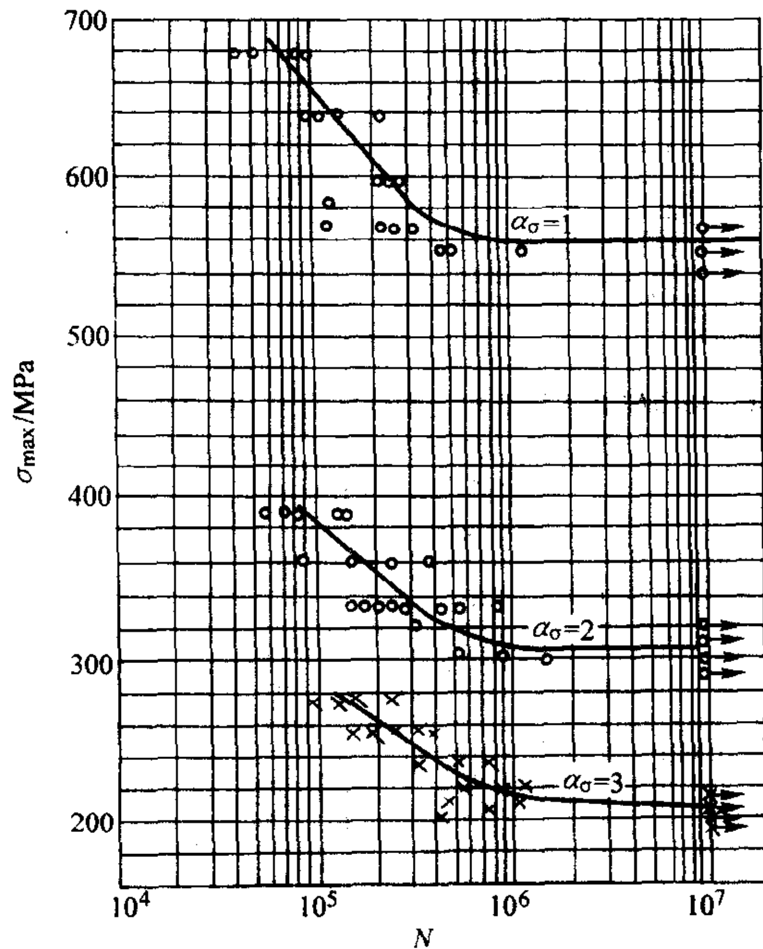


图 6.5-10　40CrNiMoA 钢棒材的 $S-N$ 曲线（棒材 $\phi22$ mm）
热处理：850℃油淬火，580℃回火
材料　$\sigma_b = 1\ 049$ MPa
试样：光滑（$\alpha_\sigma = 1$）和缺口（$\alpha_\sigma = 2$、3）
试样旋转弯曲试验，$r = -1$

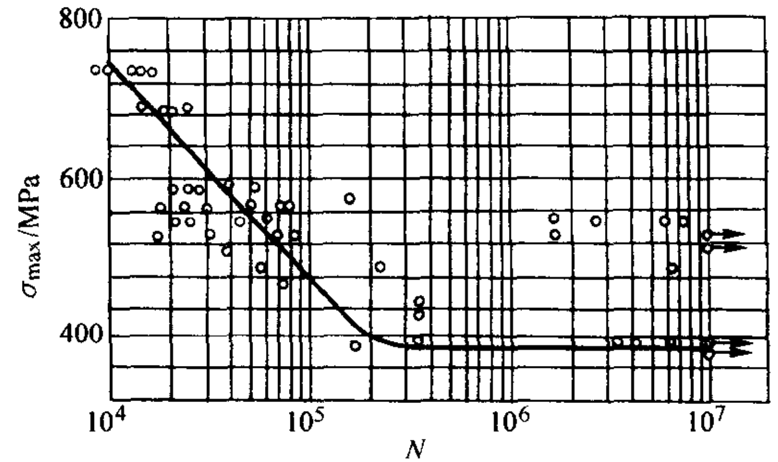


图 6.5-11　40CrMnSiMoA 钢棒材缺口试样（$\alpha_\sigma = 3$）的 $S-N$ 曲线
（棒材 $\phi42$ mm）
热处理：920℃加热，180℃等温，260℃回火
材料　$\sigma_b = 1\ 971$ MPa
轴向加载试验，$r = 0.1$

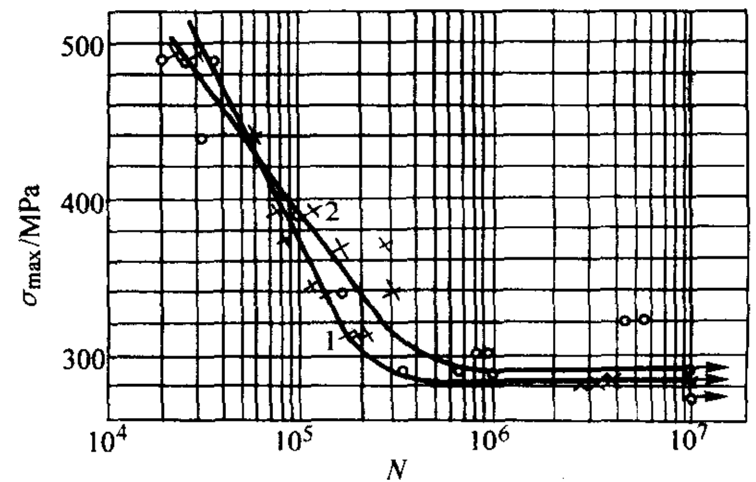


图 6.5-12　40CrNiMoA 钢棒材缺口试样（$\alpha_\sigma = 3$）的 $S-N$ 曲线
热处理：850℃油淬火，570℃回火
材料　纵向 $\sigma_b = 1\ 167$ MPa，横向 $\sigma_b = 1\ 172$ MPa
轴向加载试验，$r = 0.1$
曲线 1—纵向　曲线 2—横向

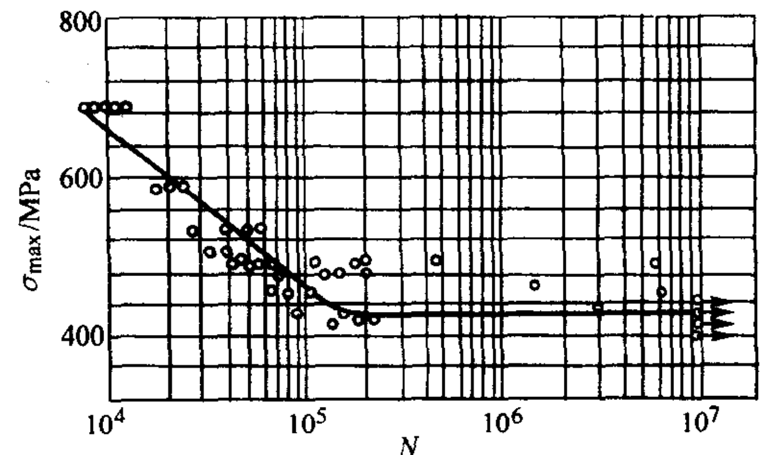


图 6.5-13　40CrMnSiMoA 钢棒材缺口试样（$\alpha_\sigma = 3$）的 $S-N$ 曲线
（棒材 $\phi42$ mm）
热处理：920℃加热，300℃等温，空冷
材料　$\sigma_b = 1\ 893$ MPa
轴向加载，$r = 0.1$

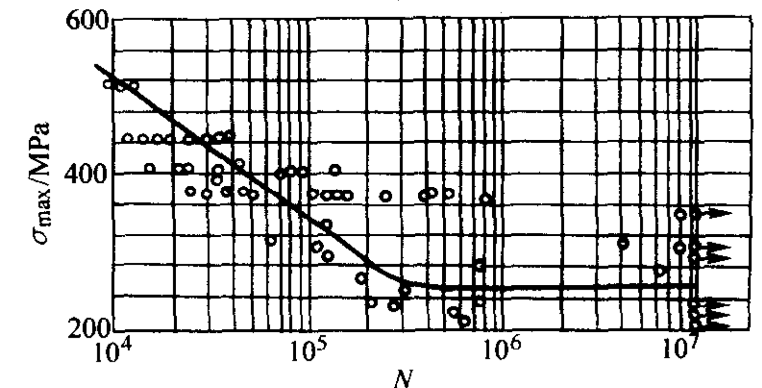


图 6.5-14　40CrMnSiMoA 钢缺口试样（$\alpha_\sigma = 3$）的 $S-N$ 曲线
（棒材 $\phi42$ mm）
热处理：920℃加热，300℃等温，空冷
材料　$\sigma_b = 1\ 893$ MPa
轴向加载，$r = -1$

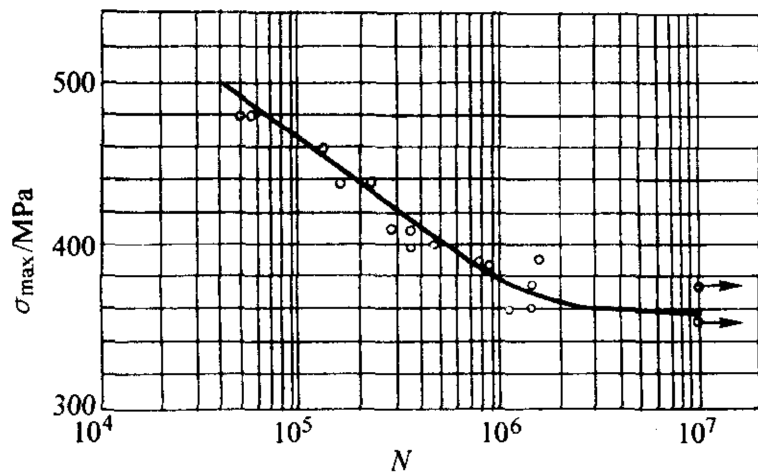


图 6.5-15　18Cr2Ni4WA 钢棒材缺口试样（$\alpha_\sigma=2$）的 $S-N$ 曲线

（棒材 ϕ18 mm）

热处理：950℃正火，860℃淬火，540℃回火

材料　$\sigma_b=1\ 145$ MPa

旋转弯曲试验，$r=-1$

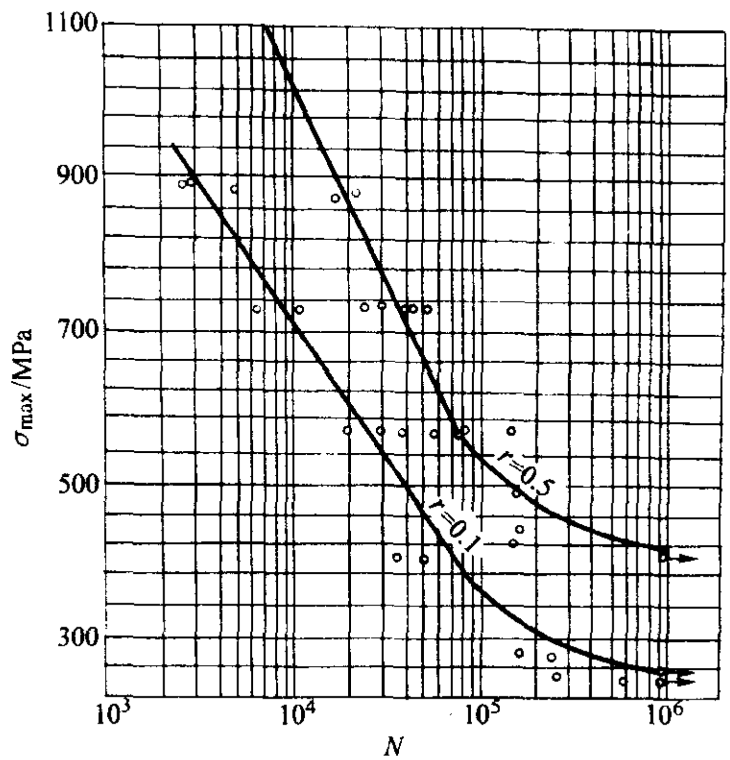


图 6.5-18　30CrMnSiNi2A 钢锻压板缺口试样($\alpha_\sigma=2.9$)的 $S-N$ 曲线

热处理：900℃淬火，250℃回火

材料　$\sigma_b=1\ 618$ MPa

轴向加载，$r=0.1$，0.5

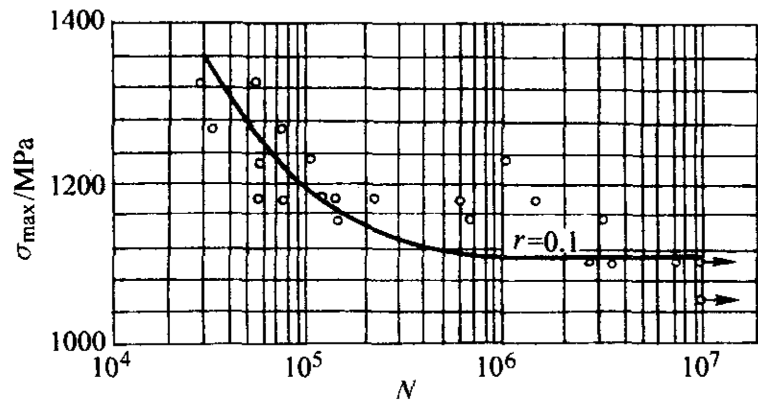


图 6.5-16　30CrMnSiNi2A 钢棒材光滑试样的 $S-N$ 曲线

（棒材 ϕ25 mm）

热处理：900℃淬火，250℃回火

材料　$\sigma_b=1\ 584$ MPa

轴向加载，$r=0.1$

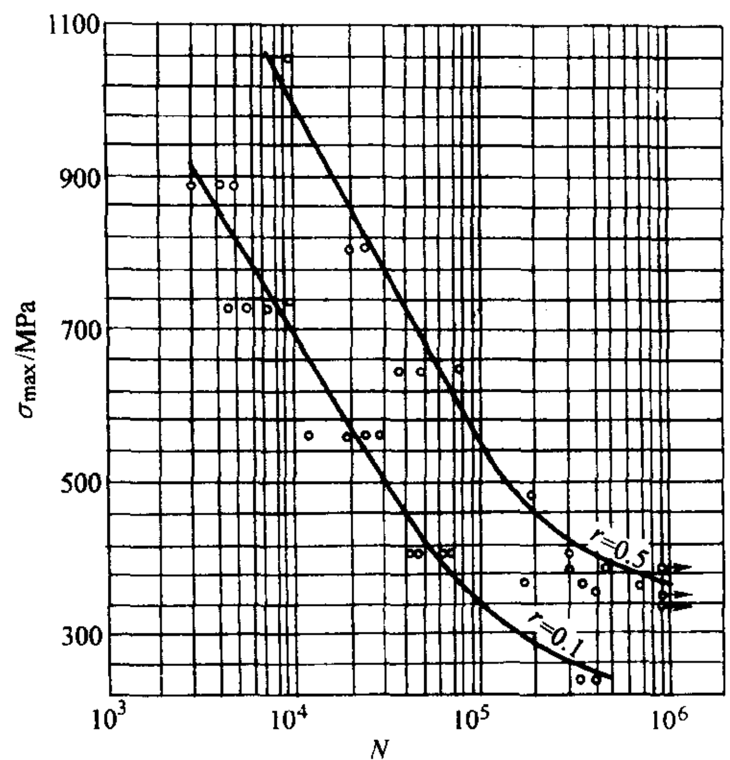


图 6.5-19　30CrMnSiNi2A 钢锻压板缺口试样($\alpha_\sigma=3.7$)的 $S-N$ 曲线

热处理：900℃淬火，250℃回火

材料　$\sigma_b=1\ 618$ MPa

轴向加载，$r=0.1$，0.5

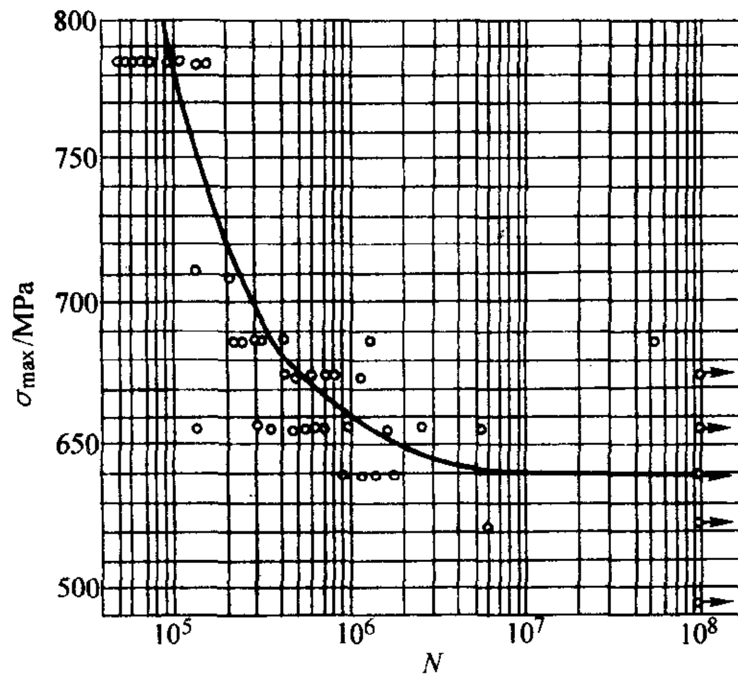


图 6.5-17　30CrMnSiA 钢锻件光滑试样的 $S-N$ 曲线

热处理：900℃油淬火，510℃回火

材料　$\sigma_b=1\ 110$ MPa

悬臂旋转弯曲试验，$r=-1$

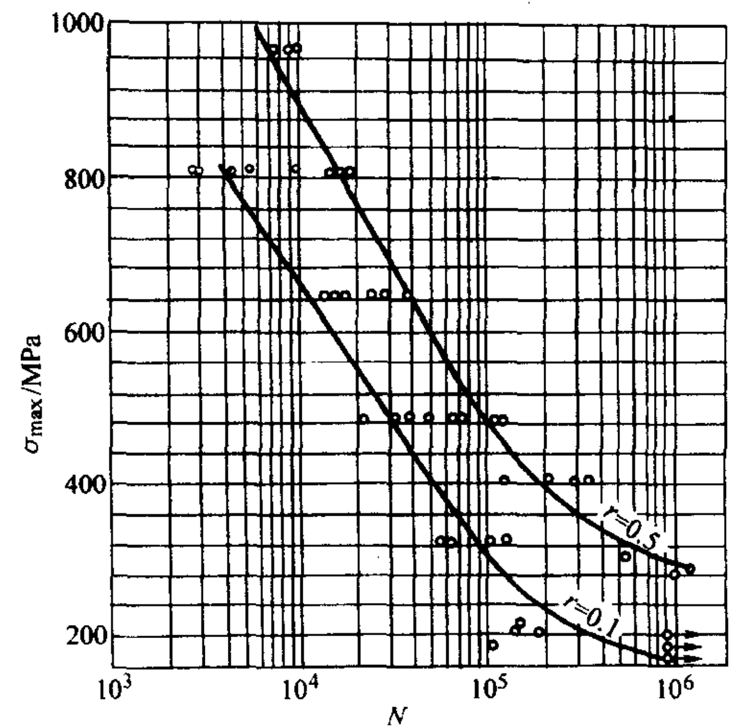


图 6.5-20　30CrMnSiNi2A 钢锻压板缺口试样($\alpha_\sigma=4.1$)的 $S-N$ 曲线

热处理：900℃淬火，250℃回火

材料　$\sigma_b=1\ 618$ MPa

轴向加载，$r=0.1$，0.5

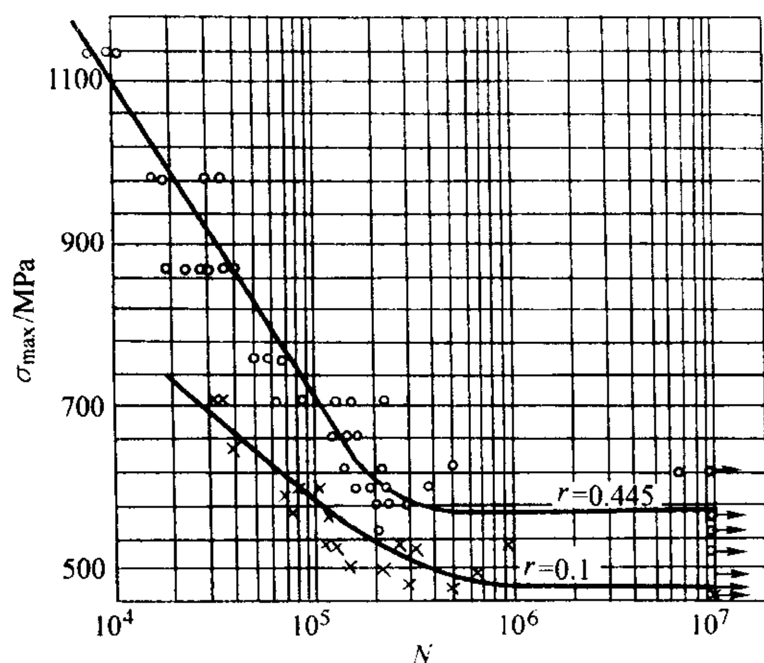


图 6.5-21 30CrMnSiNi2A 钢棒材缺口试样（$\alpha_\sigma=3$）的 $S-N$ 曲线

（棒材 ϕ25 mm）

热处理：900℃淬火，260℃回火

材料 $\sigma_b=1\ 569$ MPa（$r=0.445$）

$\sigma_b=1\ 665$ MPa（$r=0.1$）

轴向加载，$r=0.1$，0.445

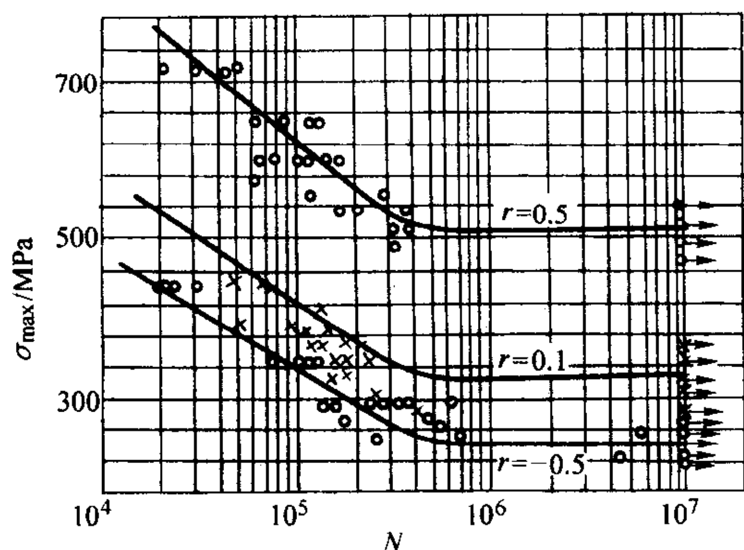


图 6.5-22 30CrMnSiNi2A 钢棒材缺口试样（$\alpha_\sigma=5$）的 $S-N$ 曲线

（棒材 ϕ25 mm）

热处理：900℃淬火，260℃回火

材料 $\sigma_b=1\ 569$ MPa（$r=0.5$，-0.5）

$\sigma_b=1\ 665$ MPa（$r=0.1$）

轴向加载，$r=0.5$，0.1，-0.5

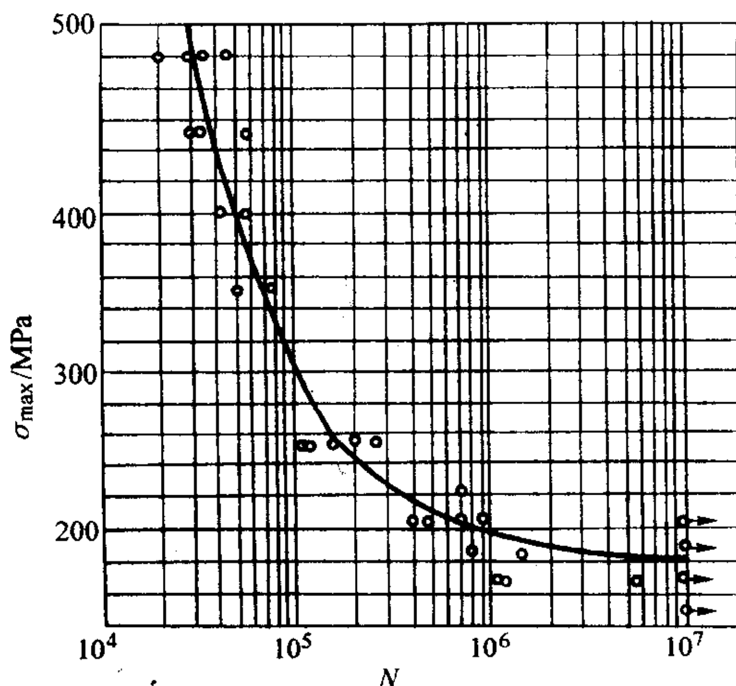


图 6.5-23 30CrMnSiNi2A 钢棒材缺口试样（$\alpha_\sigma=3$）的 $S-N$ 曲线

（棒材 ϕ55 mm）

热处理：900℃淬火，250℃回火

材料 $\sigma_b=1\ 755$ MPa

轴向加载，$r=0.1$

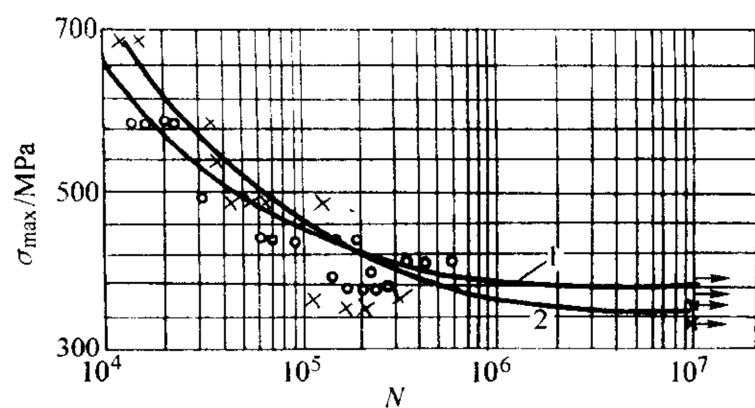


图 6.5-24 30CrMnSiNi2A 钢棒材缺口试样（$\alpha_\sigma=3$）的 $S-N$ 曲线

（棒材 ϕ30 mm）

材料 $\sigma_b=1\ 417$ MPa 1—热处理：900℃淬火，370℃回火

材料 $\sigma_b=1\ 550$ MPa 2—热处理：900℃淬火，320℃回火

轴向加载，$r=0.1$

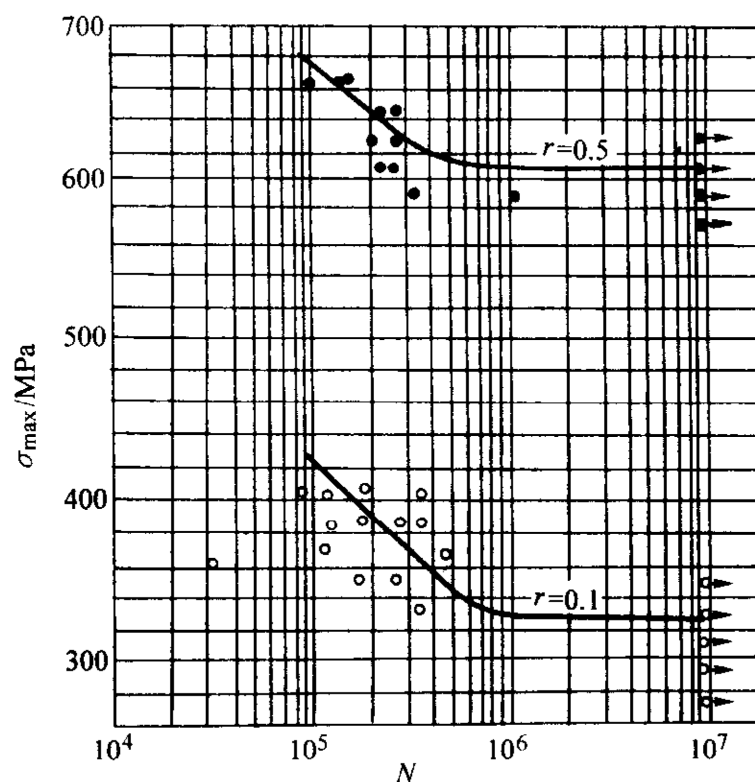


图 6.5-25 30CrMnSiA 钢棒材缺口试样（$\alpha_\sigma=3$）的 $S-N$ 曲线

（棒材 ϕ26 mm）

热处理：890℃油淬火，520℃回火

材料 $\sigma_b=1\ 184$ MPa

轴向加载，$r=0.1$，0.5

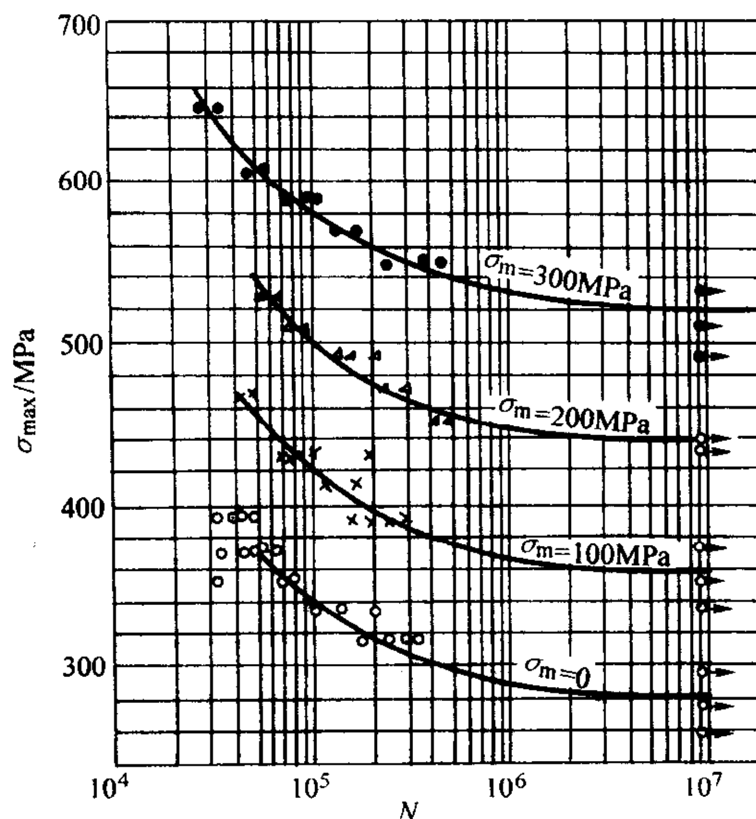


图 6.5-26 45 钢棒材缺口试样（$\alpha_\sigma=2$）的 $S-N$ 曲线

（棒材 ϕ26 mm）

热处理：调质

材料 $\sigma_b=834$ MPa

轴向加载，$\sigma_m=0$，100，200，300 MPa

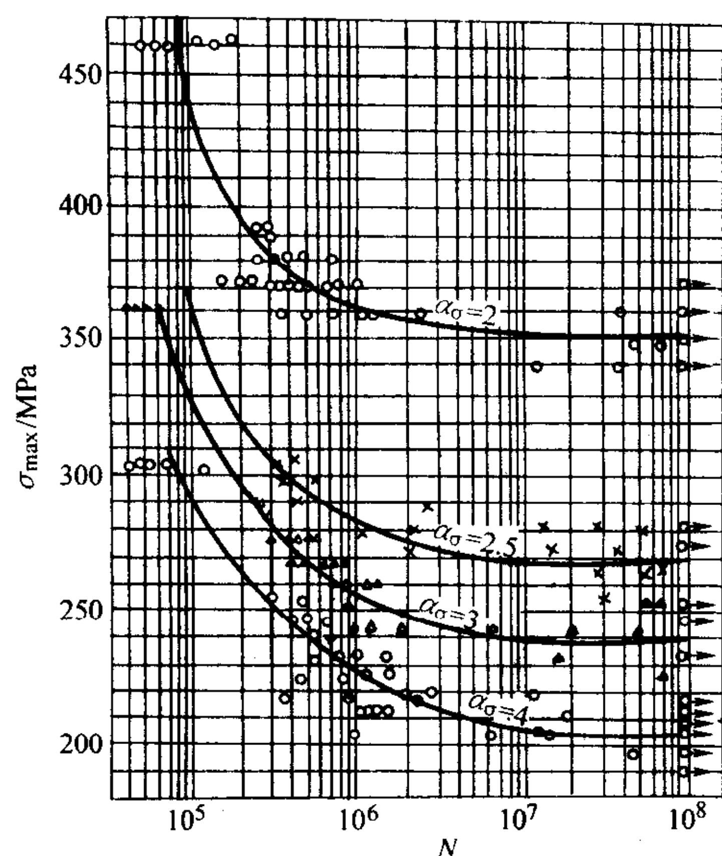


图 6.5-27 30CrMnSiA 钢锻件缺口试样（$\alpha_\sigma=2$、2.5、3.4）的 S－N 曲线

热处理：900℃油淬火，510℃回火

材料 $\sigma_b=1\ 110$ MPa

悬臂旋转弯曲试验，$r=-1$

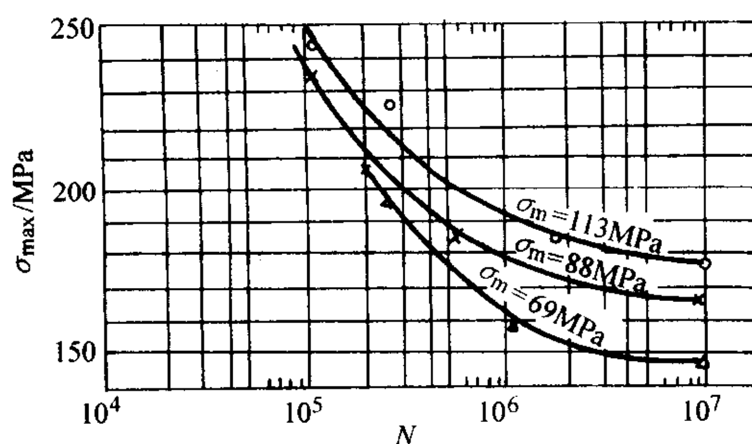


图 6.5-28 2A12－T4 铝合金板材光滑试样的 S－N 曲线

（$\delta=1$ mm）

热处理：T4 状态

材料 $\sigma_b=451$ MPa

轴向加载，$\sigma_m=69$，88，113 MPa

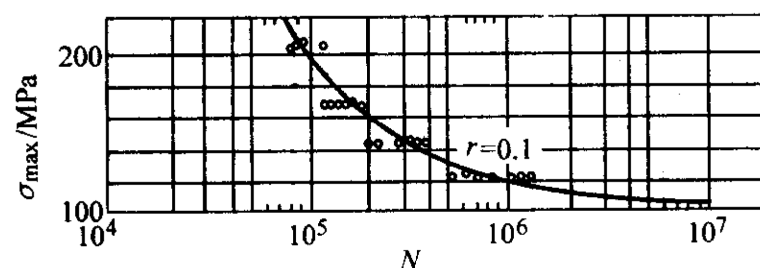


图 6.5-29 2A12－T4 阳极化铝合金板材光滑试样的 S－N 曲线

（$\delta=2.5$ mm）

热处理：T4 状态，无色硬阳极化

材料 $\sigma_b=407$ MPa

轴向加载，$r=0.1$

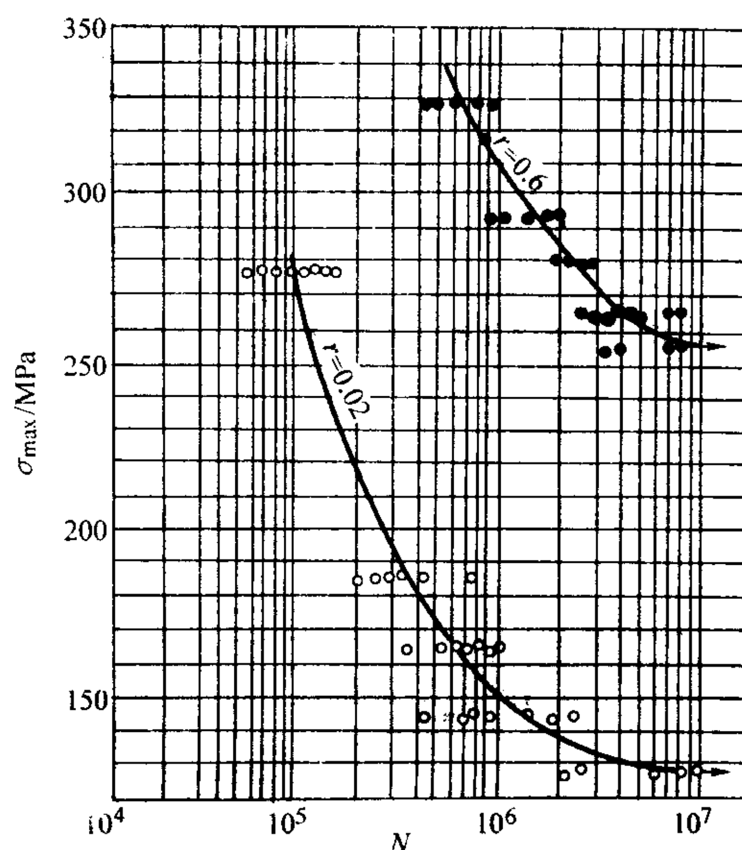


图 6.5-30 2A12－T4 铝合金板材光滑试样的 S－N 曲线

（$\delta=2.5$ mm）

热处理：淬火，自然时效

材料 $\sigma_b=457$ MPa

轴向加载，$r=0.02$，0.6

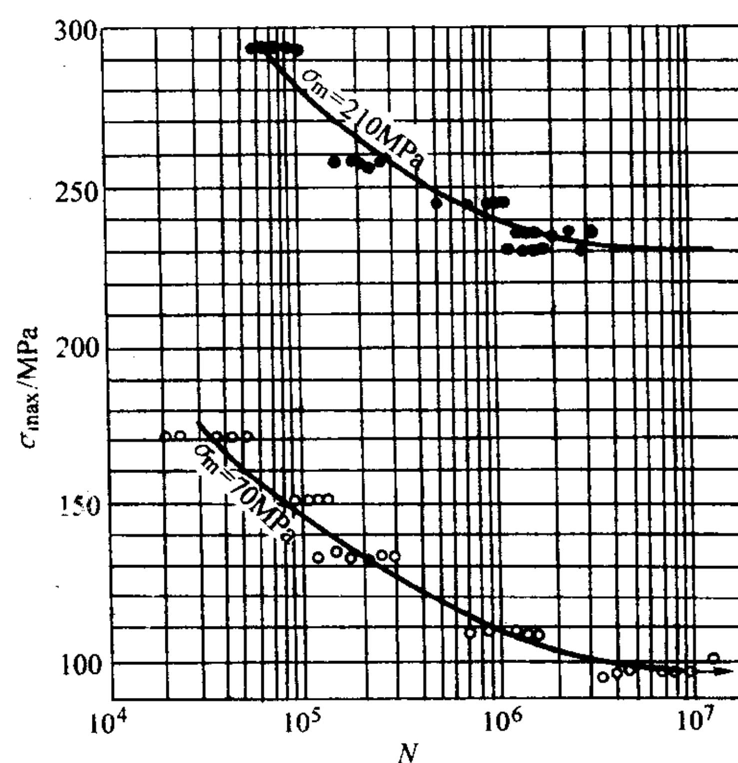


图 6.5-31 2A12－T4 铝合金板材缺口试样（$\alpha_\sigma=2$）的 S－N 曲线

（$\delta=2.5$ mm）

热处理：淬火，自然时效

材料 $\sigma_b=449$ MPa

轴向加载，$\sigma_m=70$，210 MPa

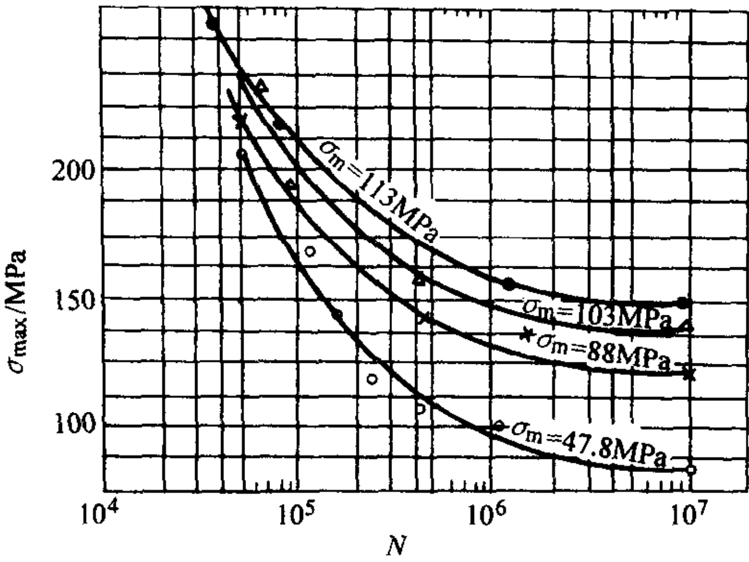


图 6.5-32　2A12 - T4 铝合金板材缺口试样（$\alpha_\sigma = 2.5$）的 $S-N$ 曲线

（$\delta = 1$ mm）

热处理：淬火，自然时效

材料　$\sigma_b = 451$ MPa

轴向加载，$\sigma_m = 47.8$，88，103，113 MPa

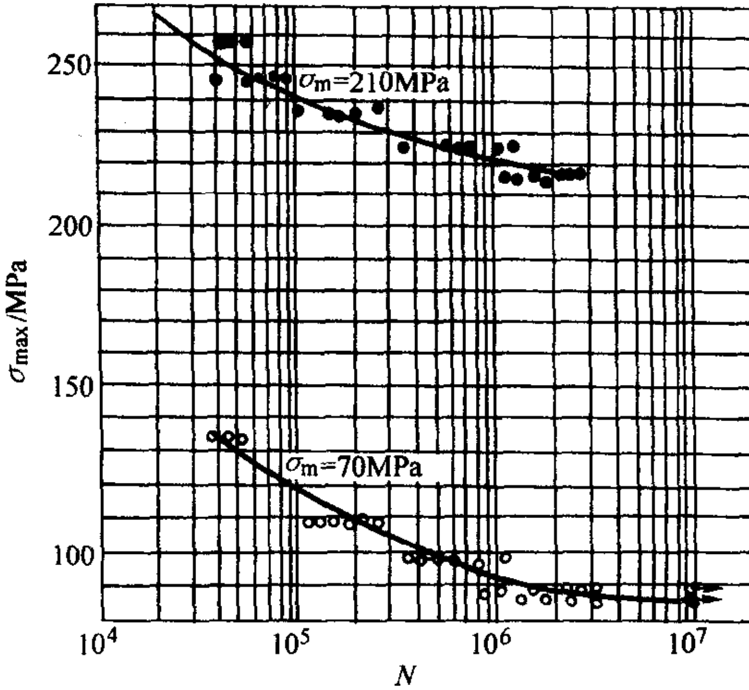


图 6.5-33　2A12 - T4 铝合金板材缺口试样（$\alpha_\sigma = 4$）的 $S-N$ 曲线

（$\delta = 2.5$ mm）

热处理：淬火，自然时效

材料　$\sigma_b = 441$ MPa

轴向加载，$\sigma_m = 70$，210 MPa

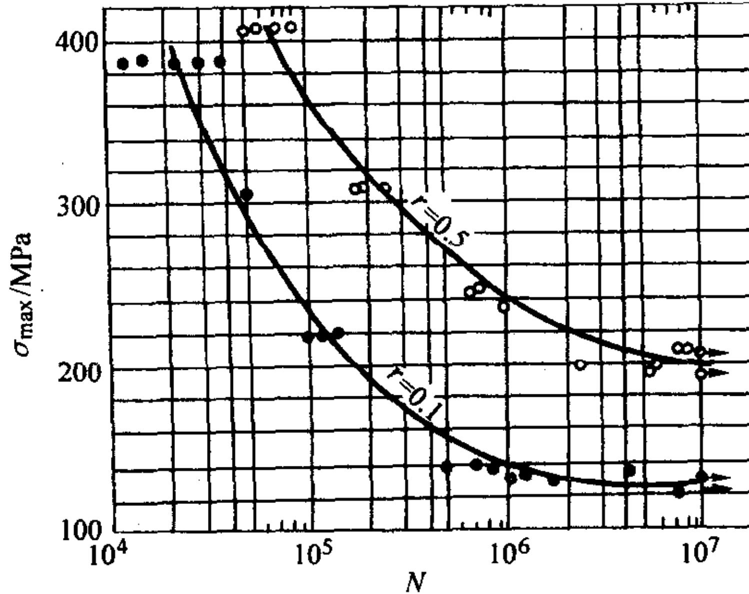


图 6.5-34　2A12 - T6 铝合金板材光滑试样的 $S-N$ 曲线

（$\delta = 2.5$ mm）

热处理：T6 状态

材料　$\sigma_b = 429$ MPa

轴向加载，$r = 0.1$，0.5

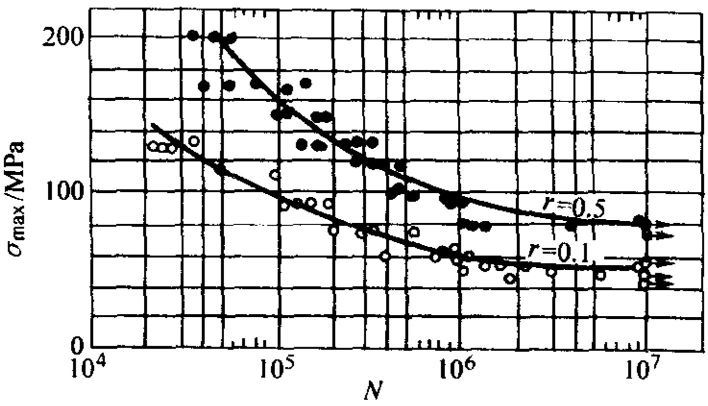


图 6.5-35　2A12 - T6 铝合金板材缺口试样（$\alpha_\sigma = 3$）的 $S-N$ 曲线

（$\delta = 2.5$ mm）

热处理：T6 状态

材料　$\sigma_b = 429$ MPa

轴向加载，$r = 0.1$，0.5

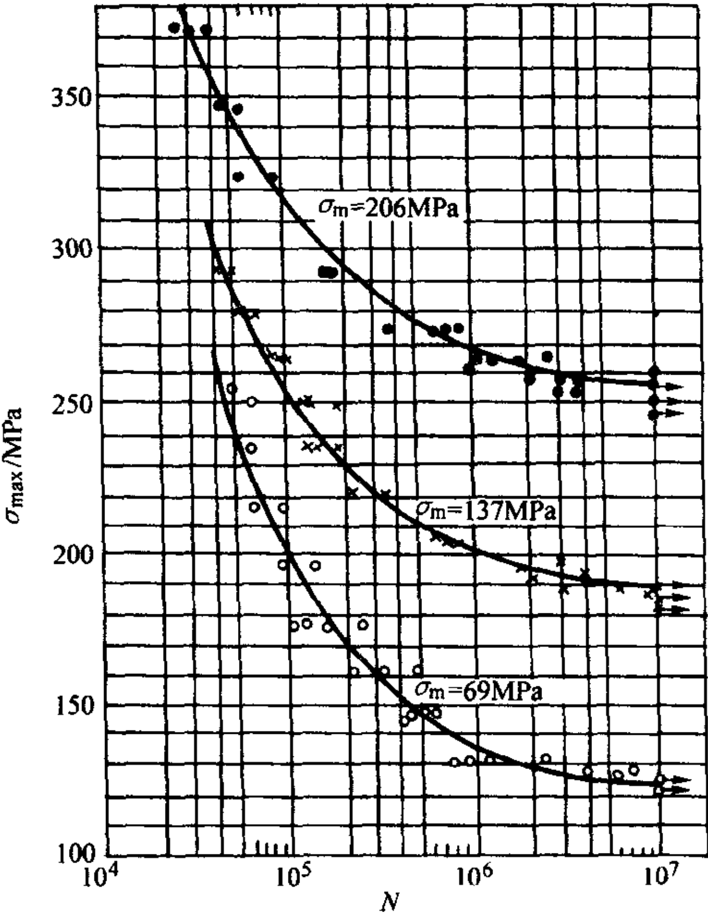


图 6.5-36　7A04 高强度铝合金板材光滑试样的 $S-N$ 曲线

（$\delta = 2.5$ mm）

热处理：T6 状态

材料　$\sigma_b = 538$ MPa

轴向加载，$\sigma_m = 69$，137，206 MPa

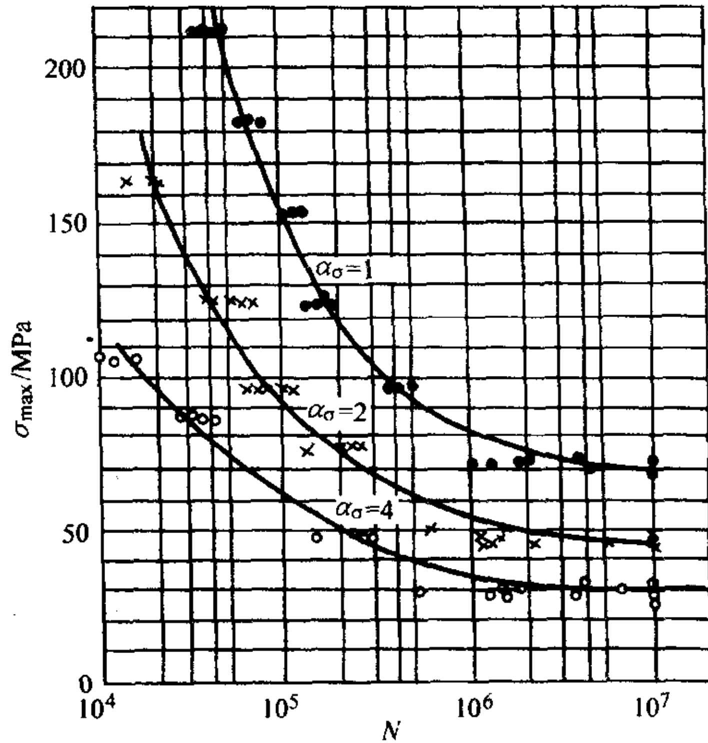


图 6.5-37　7A04 高强度铝合金板材试样（$\alpha_\sigma = 1、2、4$）的 $S-N$ 曲线

（$\delta = 2.5$ mm）

热处理：T6 状态

材料　$\sigma_b = 553$ MPa

轴向加载，$\sigma_m = 0$

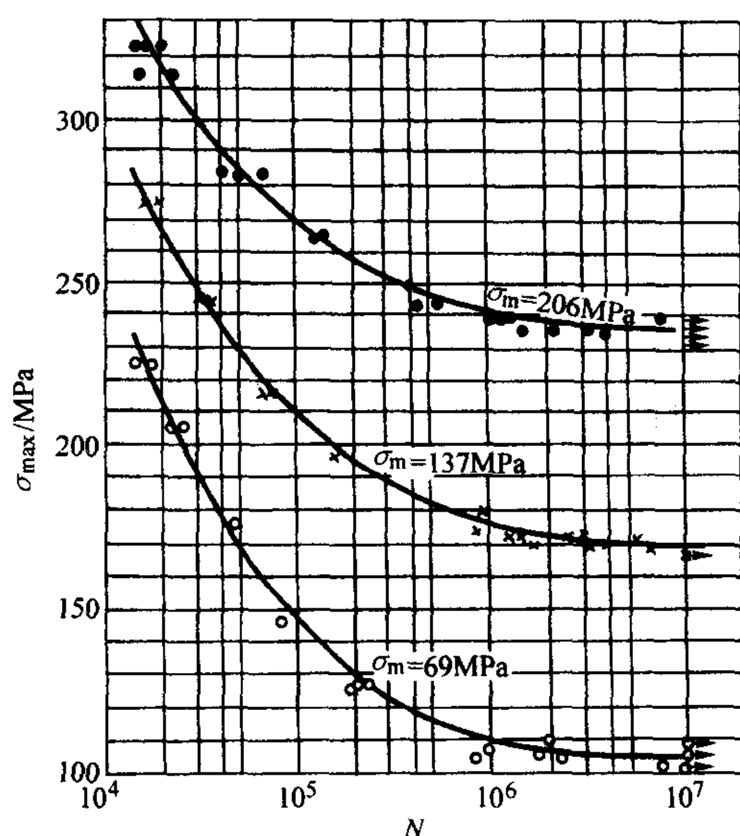


图 6.5-38　7A04 高强度铝合金板材缺口试样（$\alpha_\sigma=2$）的 $S-N$ 曲线

（$\delta=2.5$ mm）

热处理：T6 状态

材料　$\sigma_b=538$ MPa

轴向加载，$\sigma_m=69$，137，206 MPa

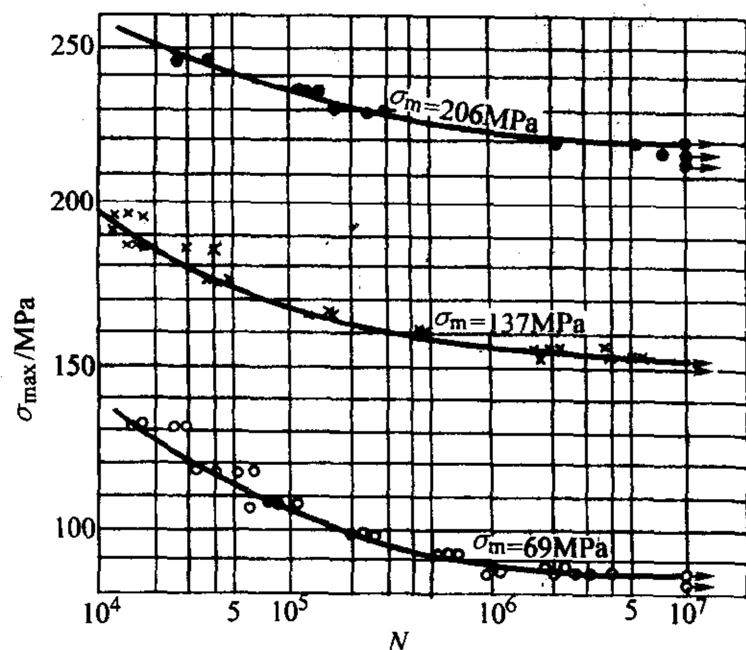


图 6.5-39　7A04 高强度铝合金板材缺口试样（$\alpha_\sigma=4$）的 $S-N$ 曲线

（$\delta=2.5$ mm）

热处理：T6 状态

材料　$\sigma_b=538$ MPa

轴向加载，$\sigma_m=69$，137，206 MPa

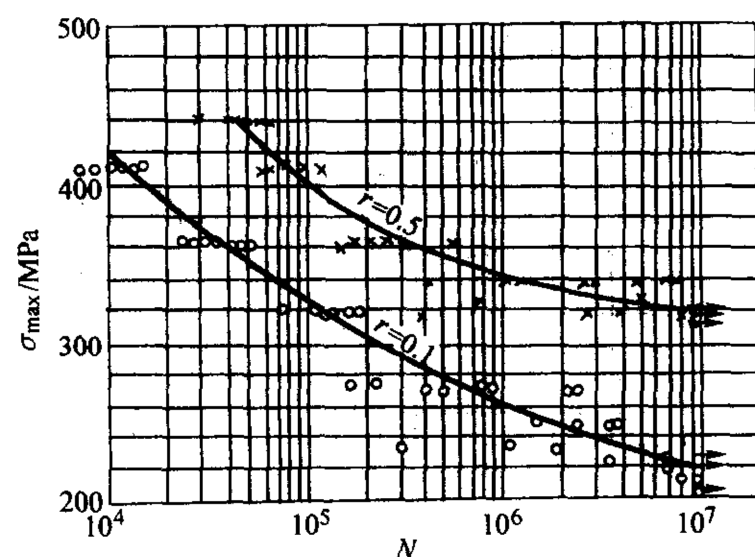


图 6.5-40　2A12B 铝合金预拉伸厚板光滑试样的 $S-N$ 曲线

（$\delta=19$ mm）

热处理：T4 预拉伸

材料　$\sigma_b=455$ MPa

轴向加载，$r=0.1$，0.5

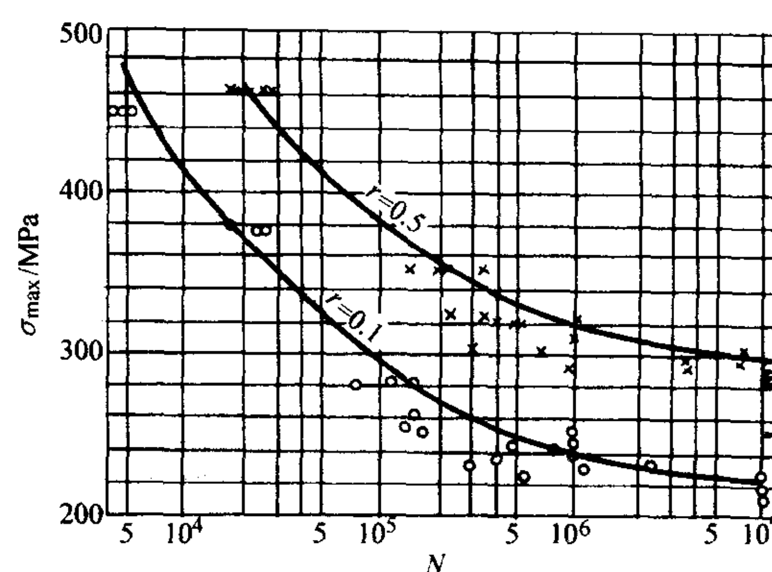


图 6.5-41　2A12B 铝合金预拉伸厚板光滑试样的 $S-N$ 曲线

（$\delta=19$ mm）

热处理：淬火自然时效，预拉伸，190℃　12 h 人工时效

材料　$\sigma_b=481$ MPa

轴向加载，$r=0.1$，0.5

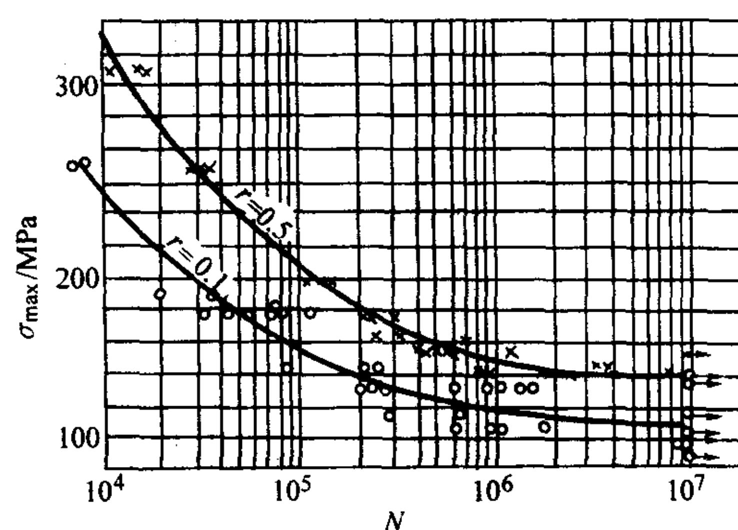


图 6.5-42　2A12B 铝合金预拉伸厚板缺口试样（$\alpha_\sigma=2$）的 $S-N$ 曲线

（$\delta=19$ mm）

热处理：T4 预拉伸

材料　$\sigma_b=455$ MPa

轴向加载，$r=0.1$，0.5

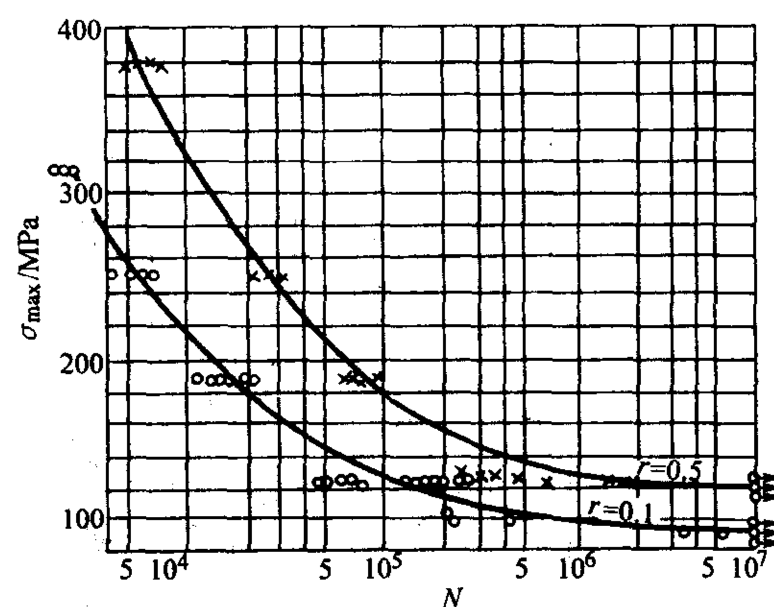


图 6.5-43　2A12B 铝合金预拉伸厚板缺口试样（$\alpha_\sigma=3$）的 $S-N$ 曲线

（$\delta=19$ mm）

热处理：淬火自然时效，预拉伸，190℃　12 h 人工时效

材料　$\sigma_b=481$ MPa

轴向加载，$r=0.1$，0.5

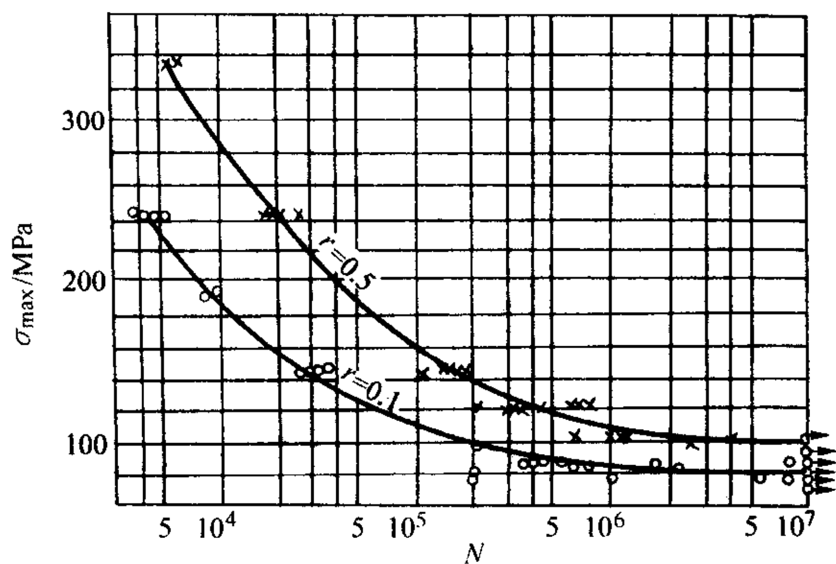


图 6.5-44　2A12B 铝合金预拉伸厚板缺口试样（$\alpha_\sigma=5$）的 $S-N$ 曲线

（$\delta=19$ mm）

热处理：T4 预拉伸

材料　$\sigma_b=455$ MPa

轴向加载，$r=0.1$，0.5

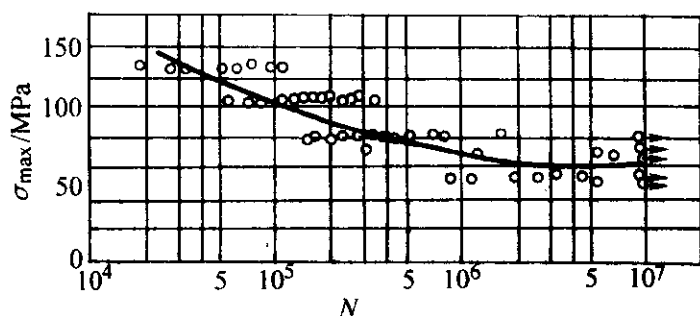


图 6.5-45　2A12B 铝合金预拉伸厚板缺口试样（$\alpha_\sigma=5$）的 $S-N$ 曲线

（$\delta=19$ mm）

热处理：T4 预拉伸

材料　$\sigma_b=455$ MPa

轴向加载，$r=-0.5$

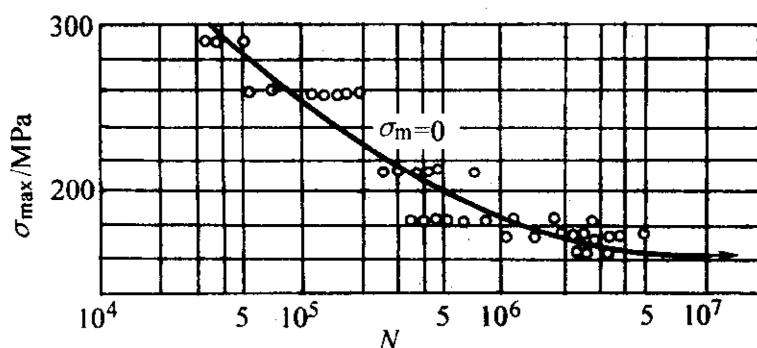


图 6.5-46　7A09 高强度铝合金棒材光滑试样的 $S-N$ 曲线（$\phi 25$ mm）

热处理：T6 状态

材料　$\sigma_b=647$ MPa

轴向加载，$\sigma_m=0$

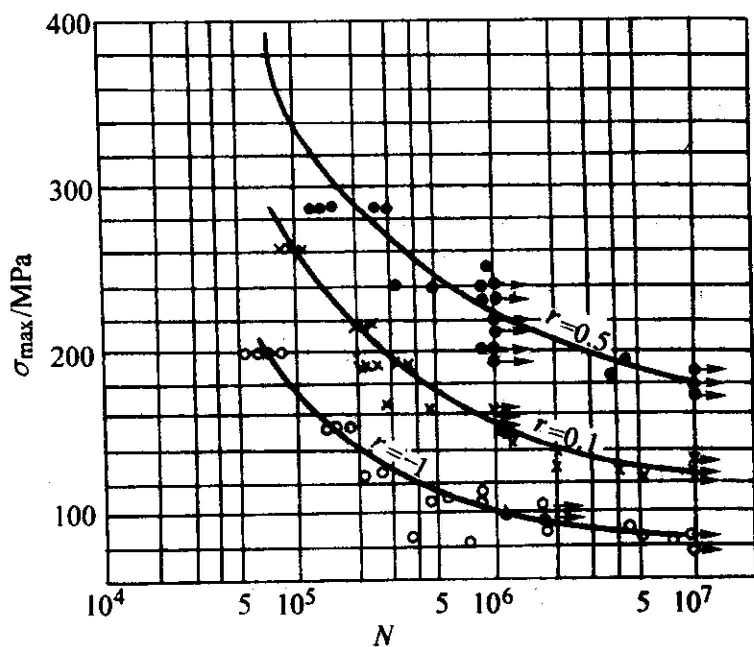


图 6.5-47　7A09 高强度铝合金过时效板材光滑试样的 $S-N$ 曲线

（$\delta=6$ mm）

热处理：460℃淬火，110℃保温，再 160℃保温

材料　$\sigma_b=498$ MPa

轴向加载，$r=-1$，0.1，0.5

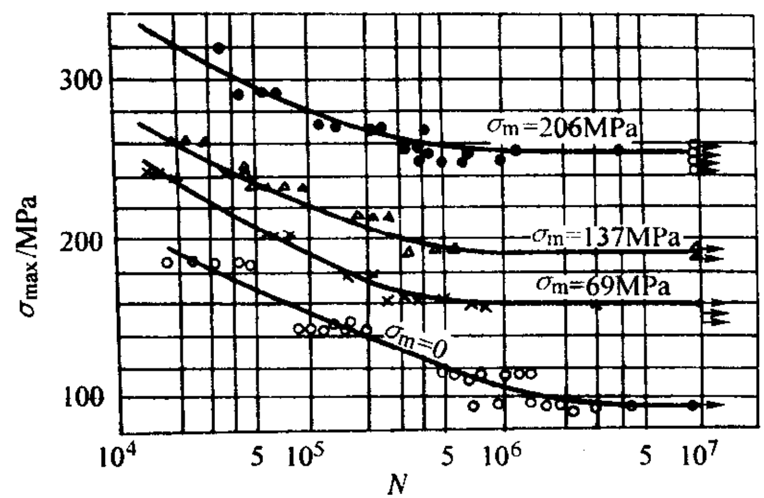


图 6.5-48　7A09 高强度铝合金棒材缺口试样（$\alpha_\sigma=2.4$）的 $S-N$ 曲线（$\phi 25$ mm）

热处理：T6 状态

材料　$\sigma_b=647$ MPa

轴向加载，$\sigma_m=0$，69，137，206 MPa

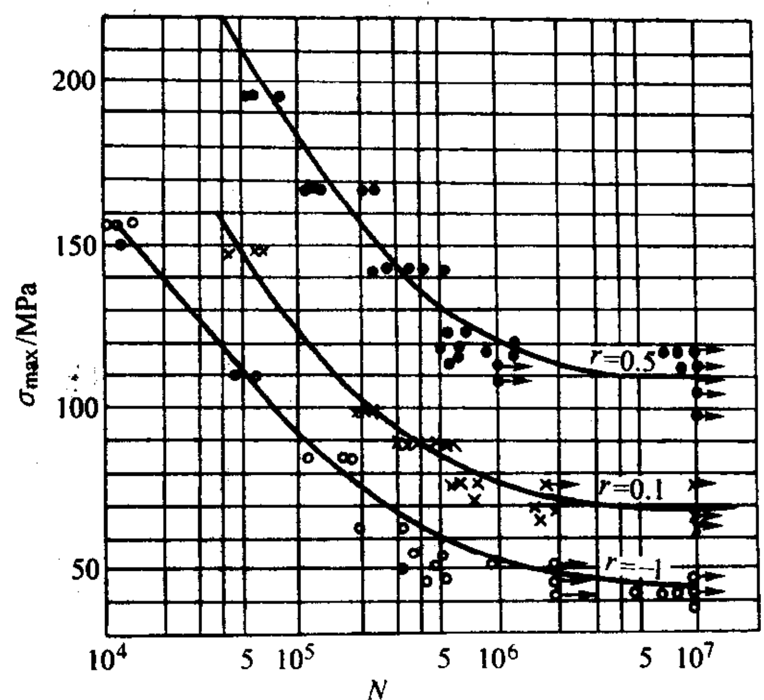


图 6.5-49　7A09 高强度铝合金过时效板材缺口试样（$\alpha_\sigma=3$）的 $S-N$ 曲线（$\delta=6$ mm）

热处理：460℃淬火，110℃保温，再 160℃保温

材料　$\sigma_b=498$ MPa

轴向加载，$r=-1$，0.1，0.5

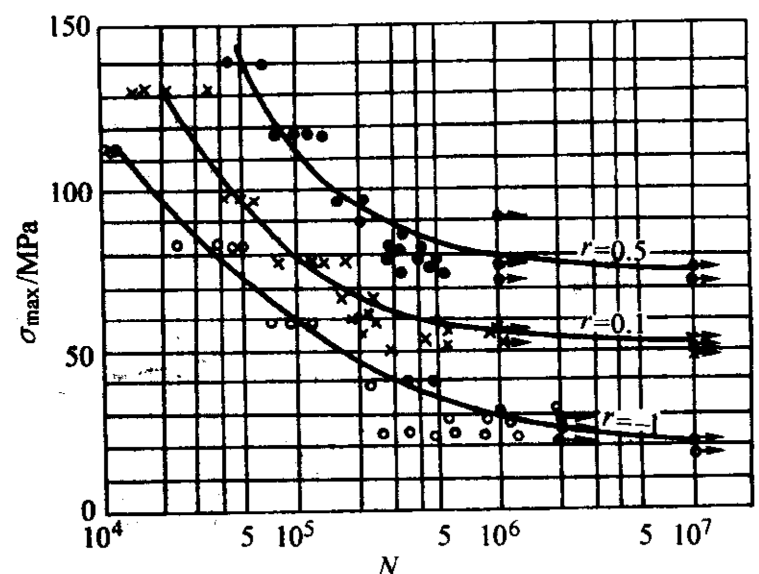


图 6.5-50　7A09 高强度铝合金过时效板材缺口试样（$\alpha_\sigma=5$）的 $S-N$ 曲线（$\delta=6$ mm）

热处理：460℃淬火，110℃保温，再 160℃保温

材料　$\sigma_b=498$ MPa

轴向加载，$r=-1$，0.1，0.5

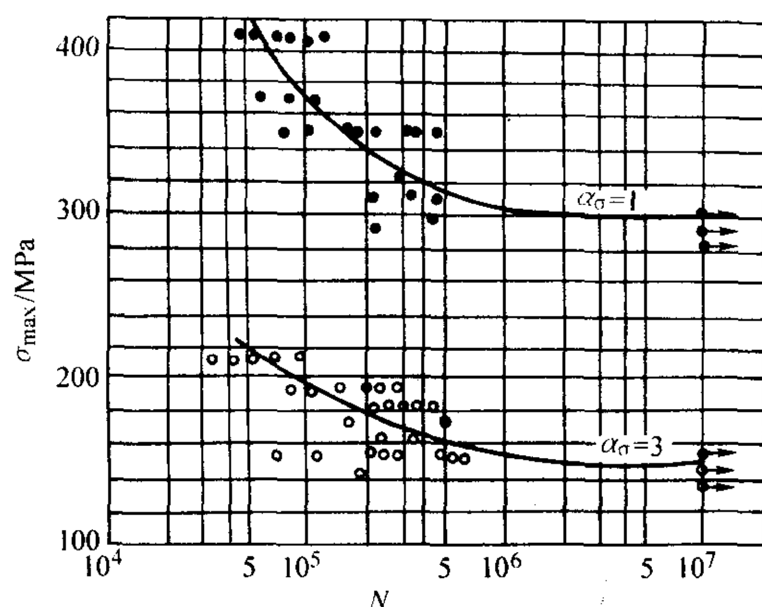


图 6.5-51 2A14 铝合金棒材缺口试样（$\alpha_\sigma=1$、3）的 $S-N$ 曲线
（ϕ25 mm）
热处理：T6 状态
材料 $\sigma_b=541$ MPa
轴向加载，$r=0.1$

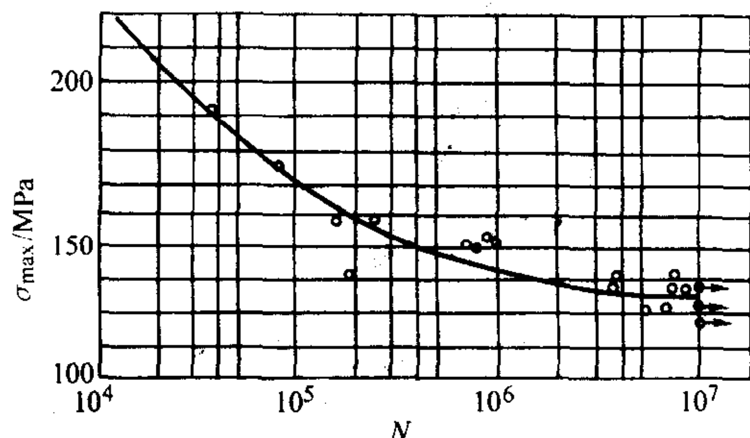


图 6.5-52 MB15 镁合金光滑试样的 $S-N$ 曲线（ϕ20 mm）
热处理：热挤压，人工时效
材料 $\sigma_b=330$ MPa
旋转弯曲试验，$r=-1$

3 影响材料疲劳强度的因素

3.1 应力集中影响

3.1.1 理论应力集中系数

在零件的截面几何形状突然变化处（如轴肩、圆角、沟槽、横孔等），局部应力远大于名义应力，这种现象称为应力集中。在材料的弹性范围内，最大局部应力 σ_{max} 与名义应力 σ_n 的比值 α_σ，称为理论应力集中系数，即

$$\alpha_\sigma=\frac{\sigma_{max}}{\sigma_n} \tag{6.5-10}$$

式（6.5-10）所定义的理论应力集中系数 α_σ 是几何参数，仅由零件的几何形状决定。

对于扭转的理论应力集中系数，定义为

$$\alpha_\tau=\frac{\tau_{max}}{\tau_n} \tag{6.5-11}$$

假设材料是各向同性、均匀的，在材料的弹性极限范围内，局部最大应力 σ_{max}（τ_{max}）可以用弹性力学解析法、光弹法或有限元法求得，从而得到不同几何形状的试样在不同载荷下的理论应力集中系数。

常见的几何形状理论应力集中系数见图 6.5-53～图 6.5-108。

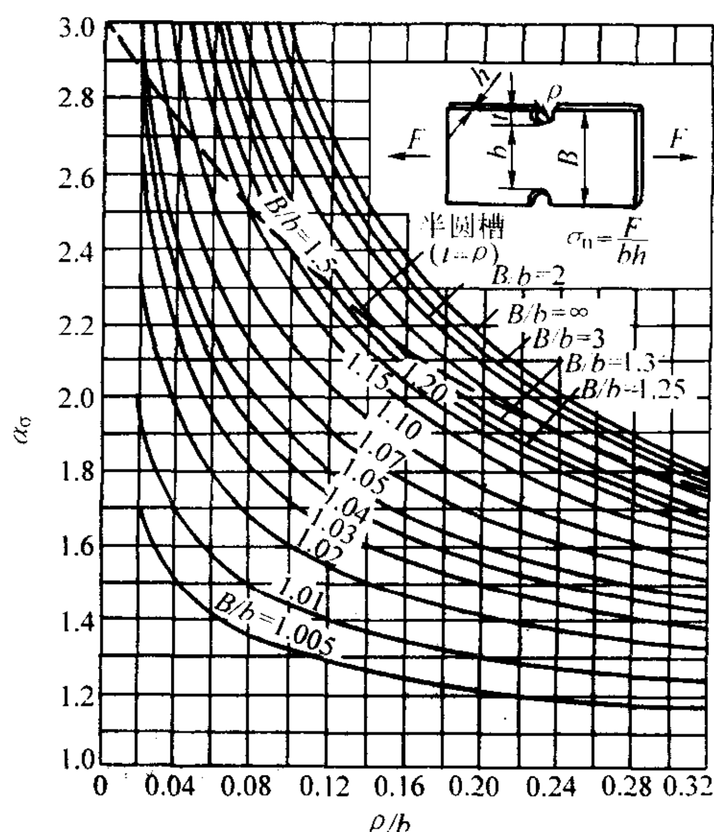


图 6.5-53 有两侧小圆弧槽的平板拉伸时的理论应力集中系数

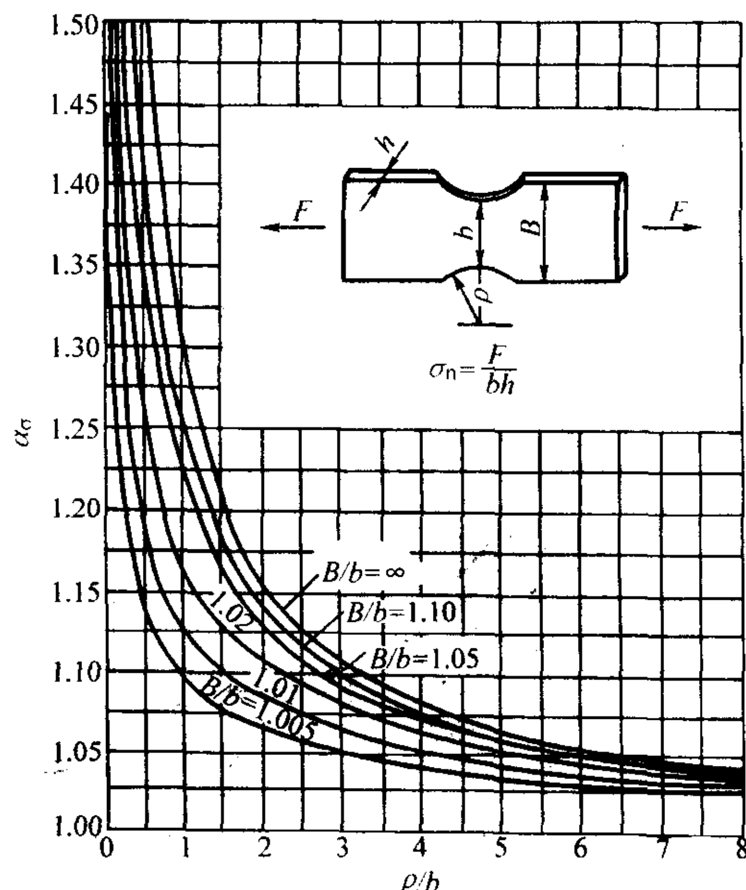


图 6.5-54 有两侧大圆弧槽的平板拉伸时的理论应力集中系数

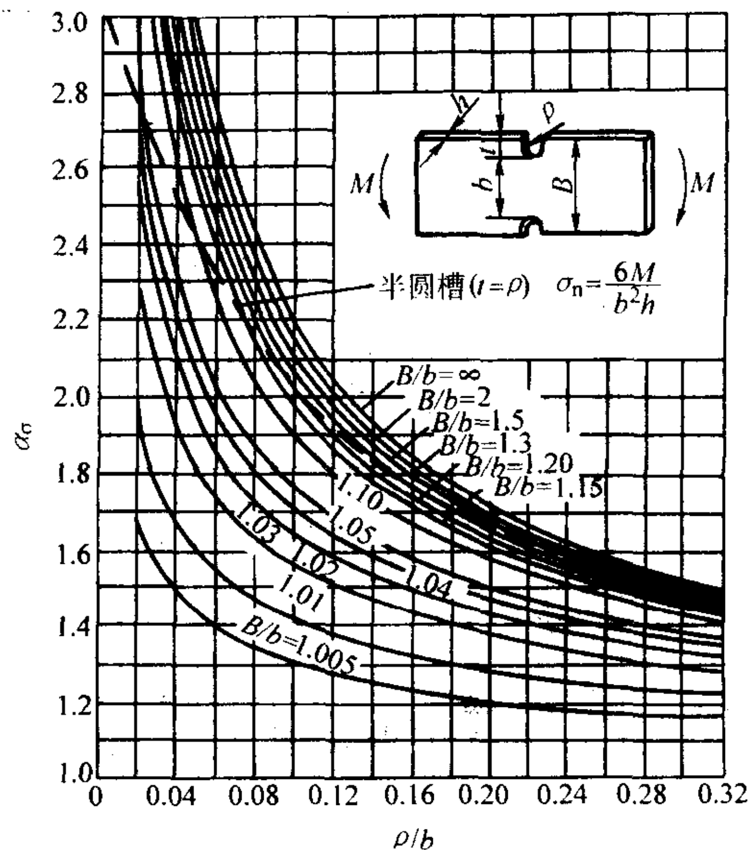


图 6.5-55 有两侧小圆弧槽的平板弯曲时的理论应力集中系数

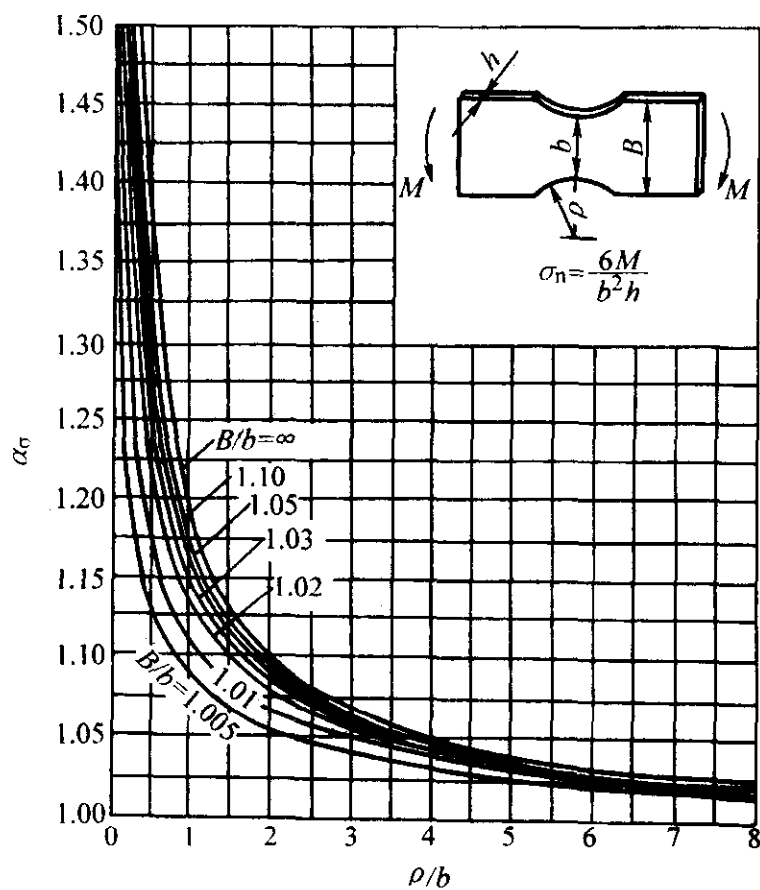


图 6.5-56 有两侧大圆弧槽的平板弯曲时的理论应力集中系数

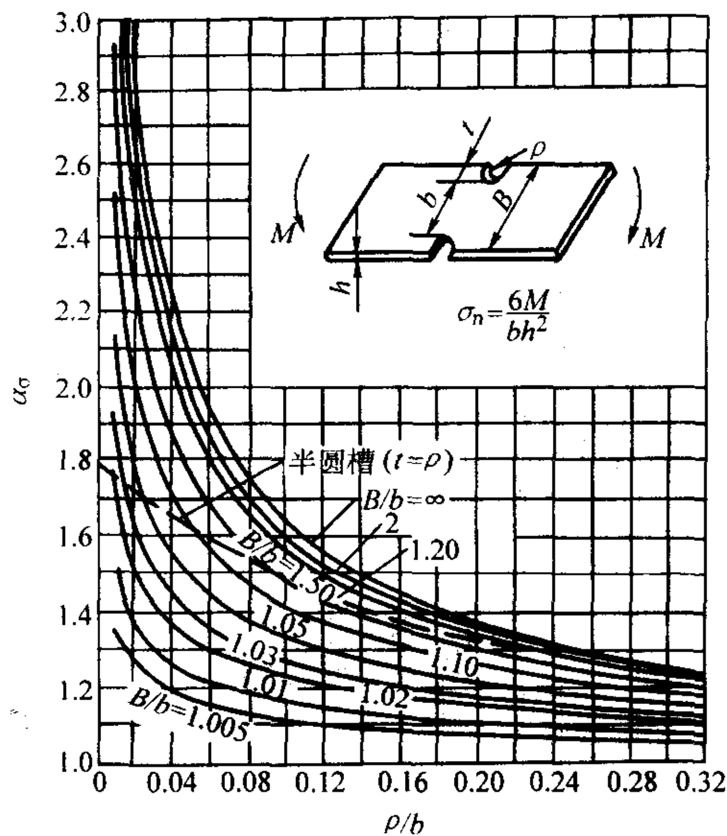


图 6.5-57 有两侧小圆弧槽的平板横向弯曲时的理论应力集中系数

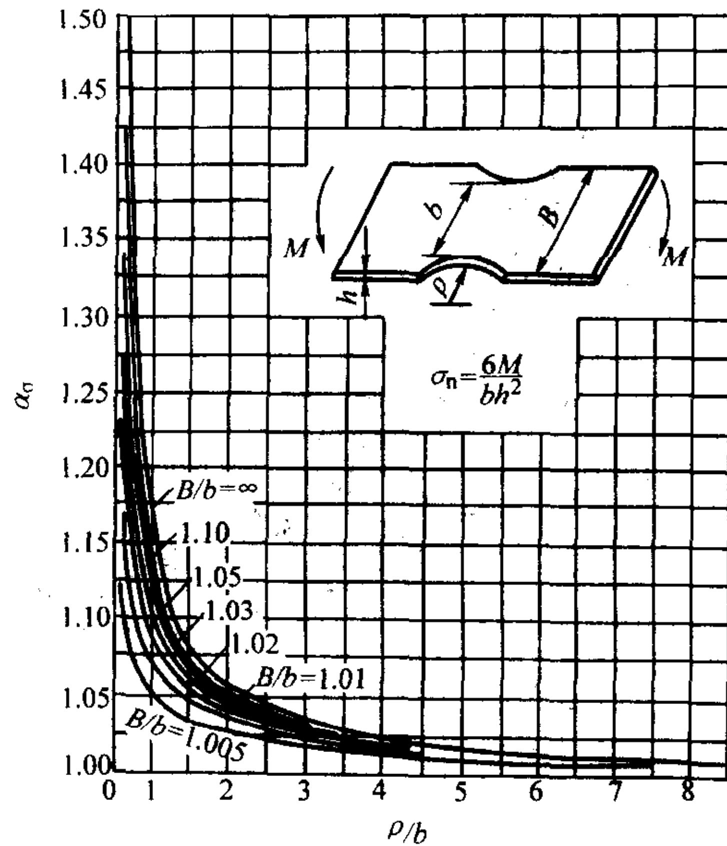


图 6.5-58 有两侧大圆弧槽的平板横向弯曲时的理论应力集中系数

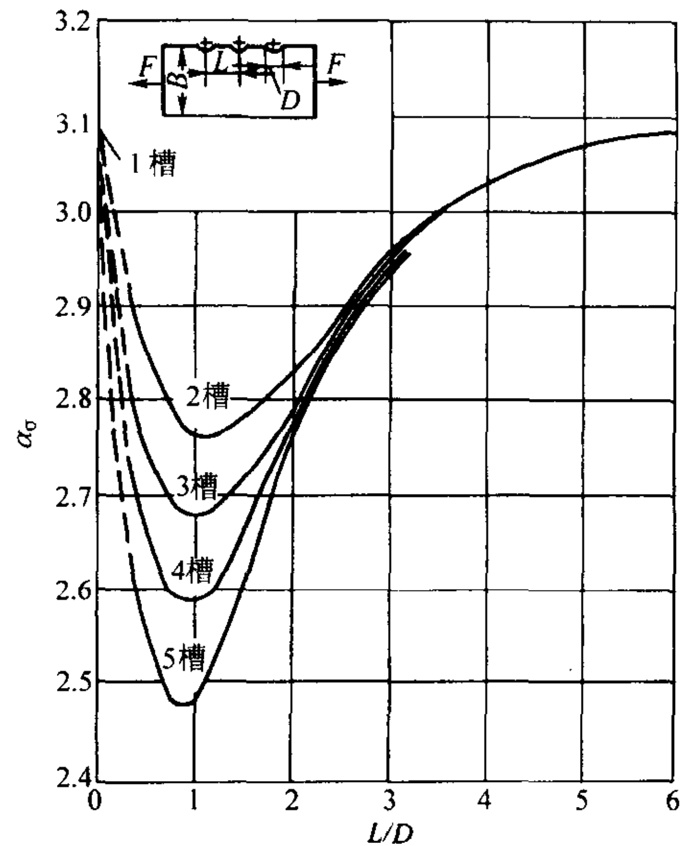


图 6.5-59 有单侧半圆槽的平板拉伸的理论应力集中系数

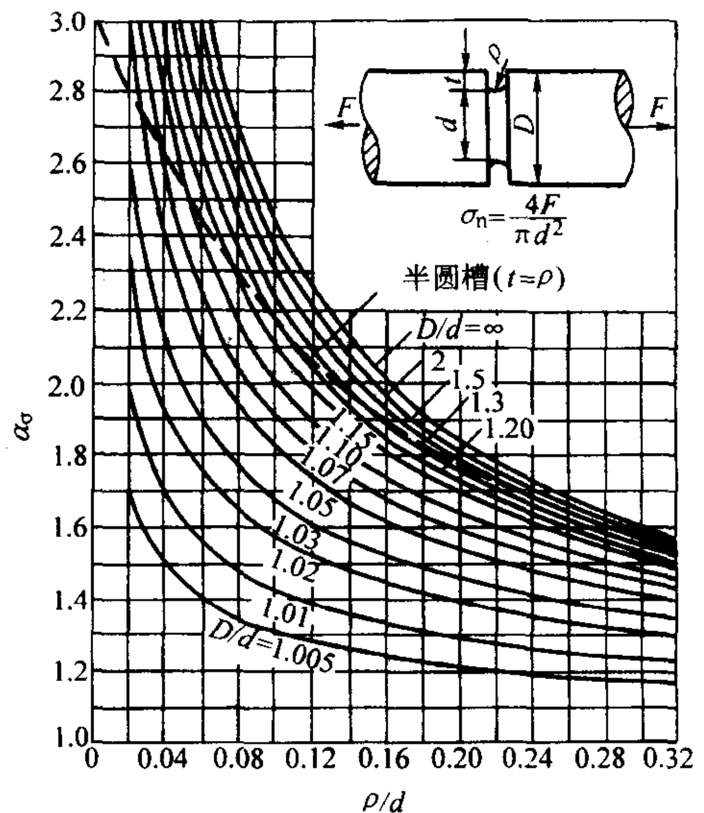


图 6.5-60 有小环形槽的轴拉伸时的理论应力集中系数

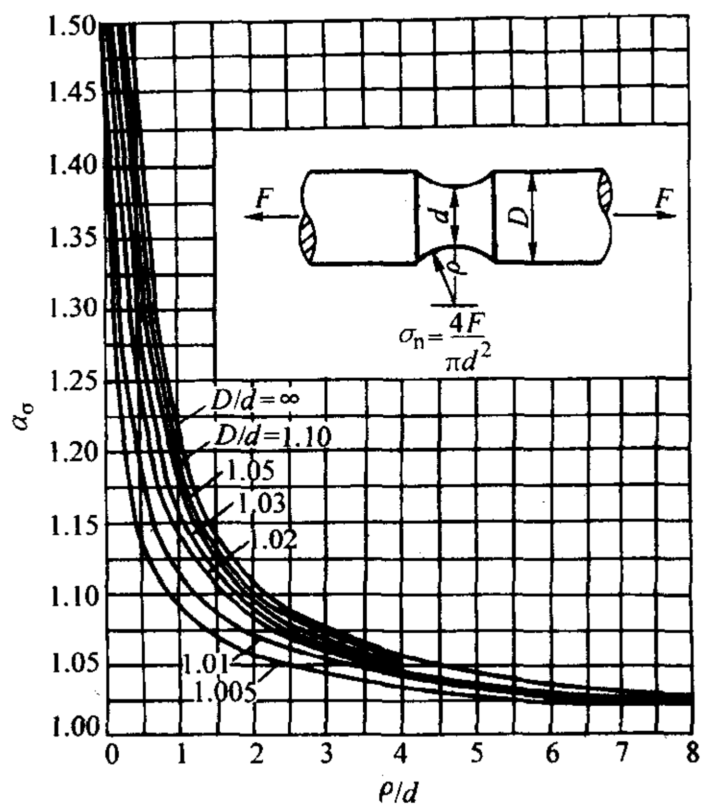


图 6.5-61 有大环形槽的轴拉伸时的理论应力集中系数

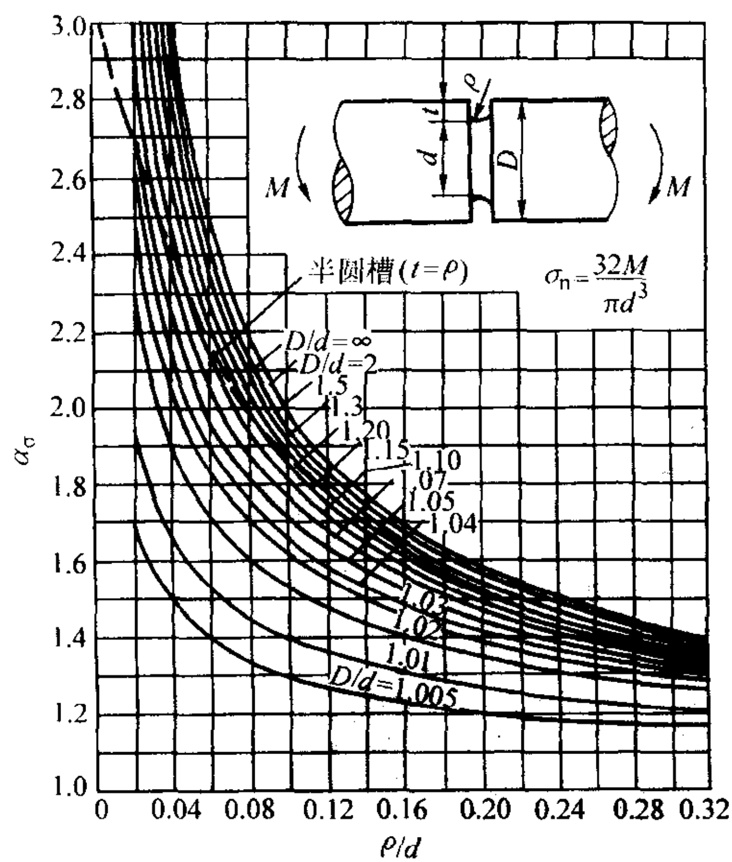


图 6.5-62　有小环形槽的轴弯曲时的理论应力集中系数

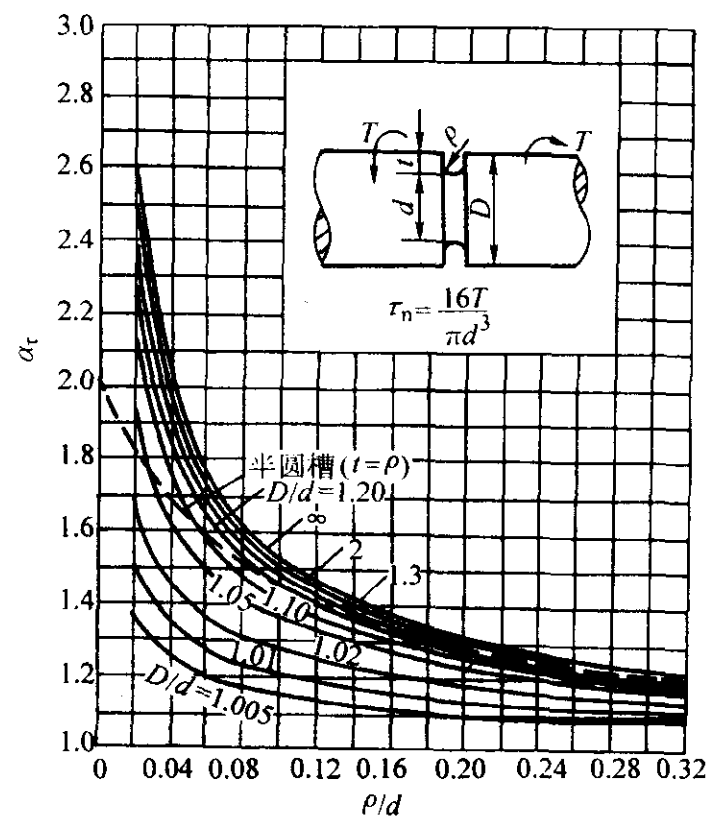


图 6.5-64　有小环形槽的轴扭转时的理论应力集中系数

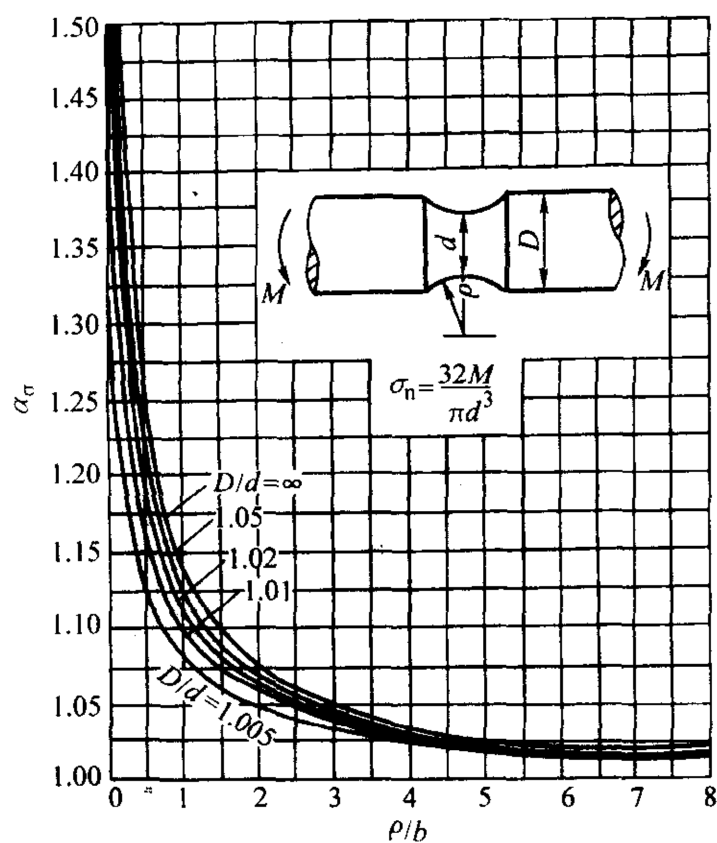


图 6.5-63　有大环形槽的轴弯曲时的理论应力集中系数

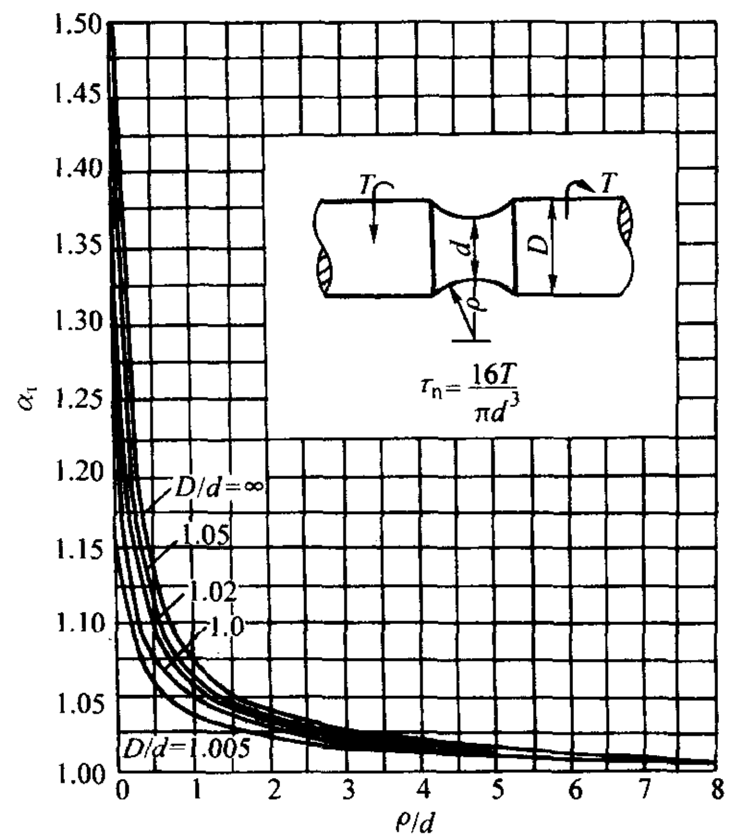


图 6.5-65　有大环形槽的轴扭转时的理论应力集中系数

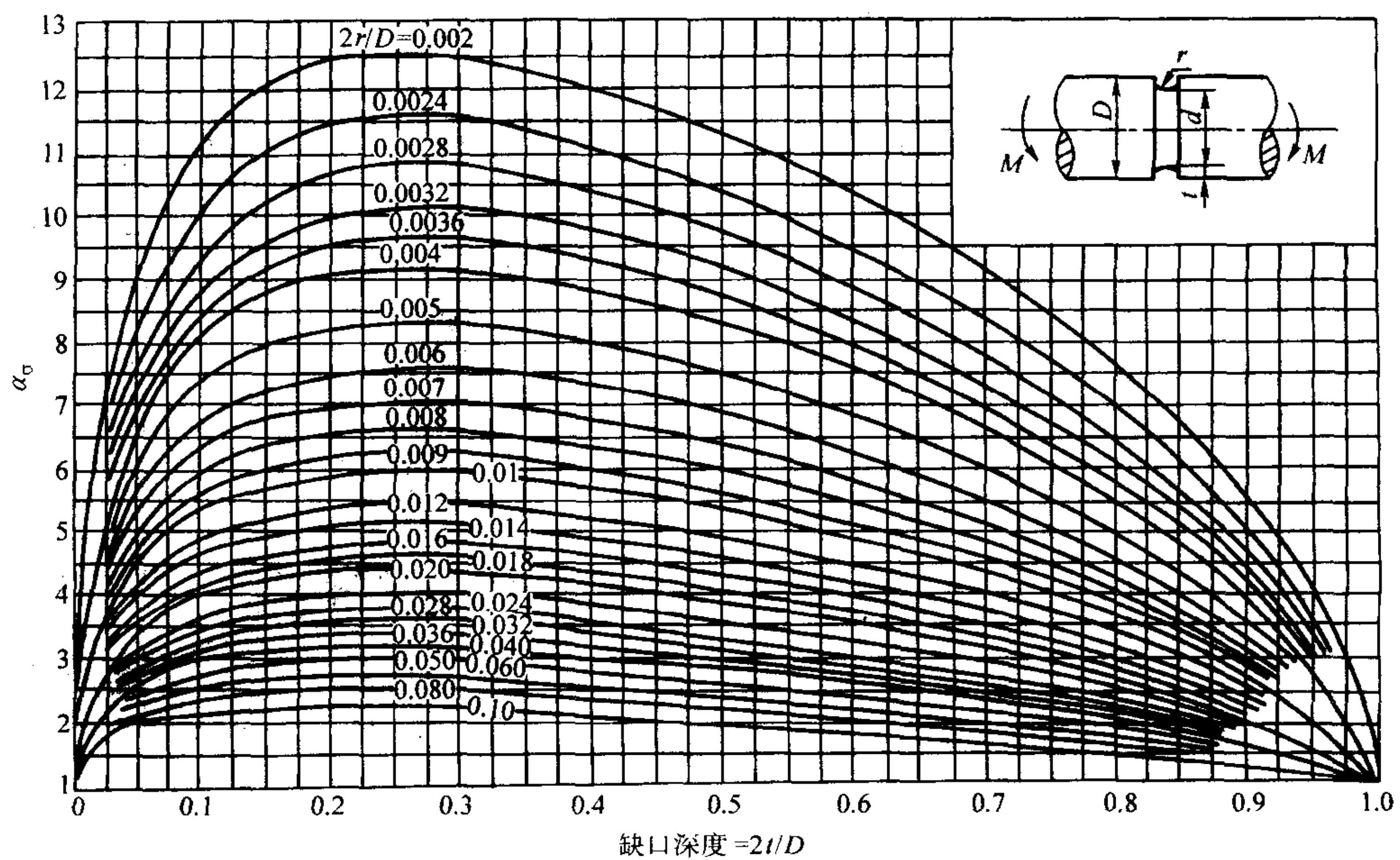


图 6.5-66　有小环形槽的轴弯曲时的理论应力集中系数

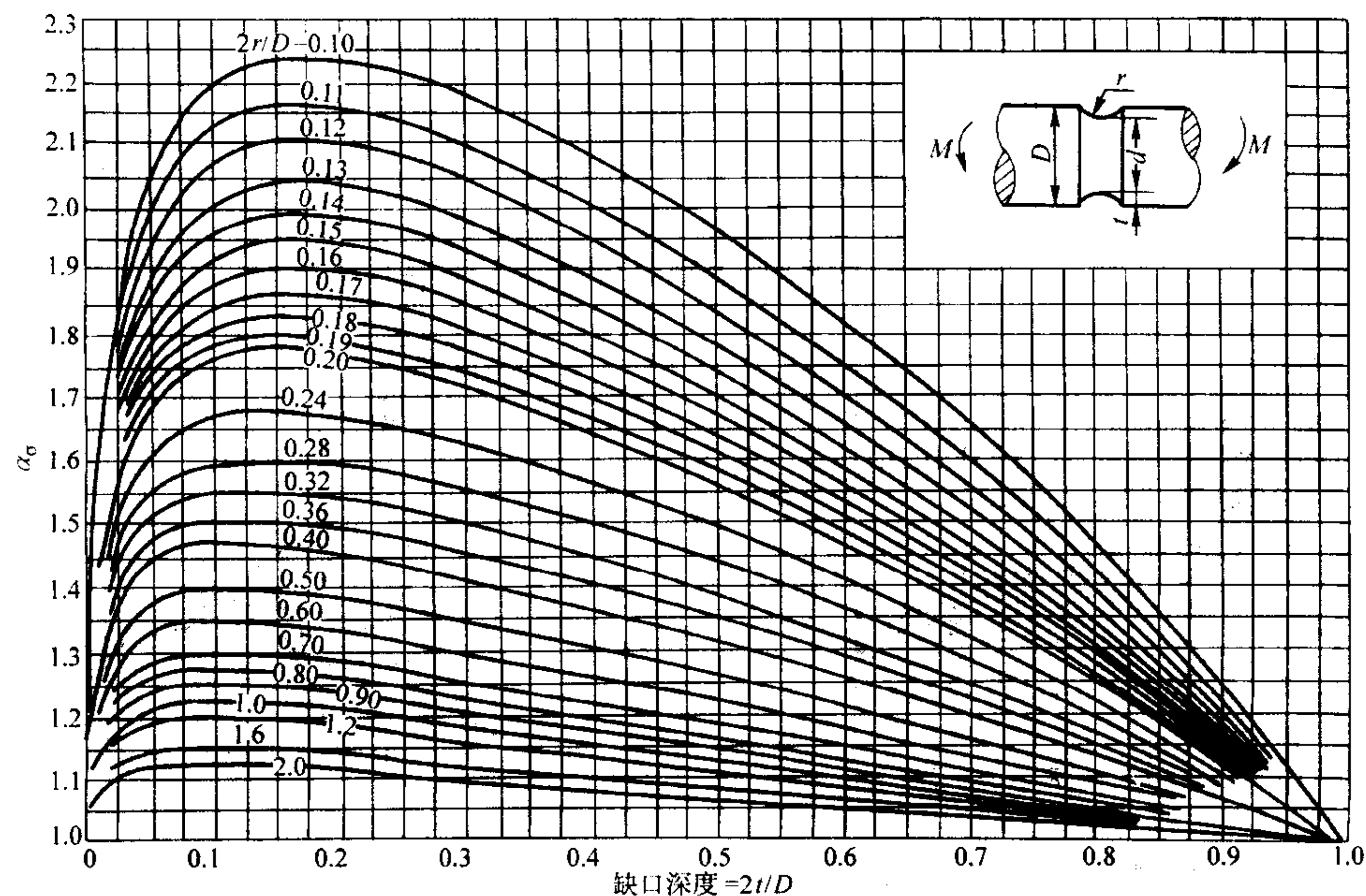


图 6.5-67　有大环形槽的轴弯曲时的理论应力集中系数

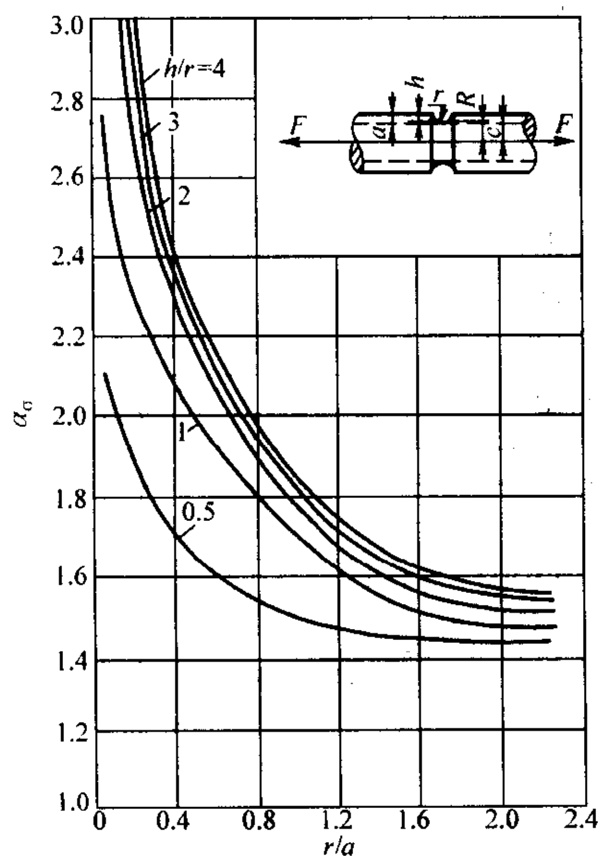


图 6.5-68　有环形槽的空心轴拉伸时的理论应力集中系数

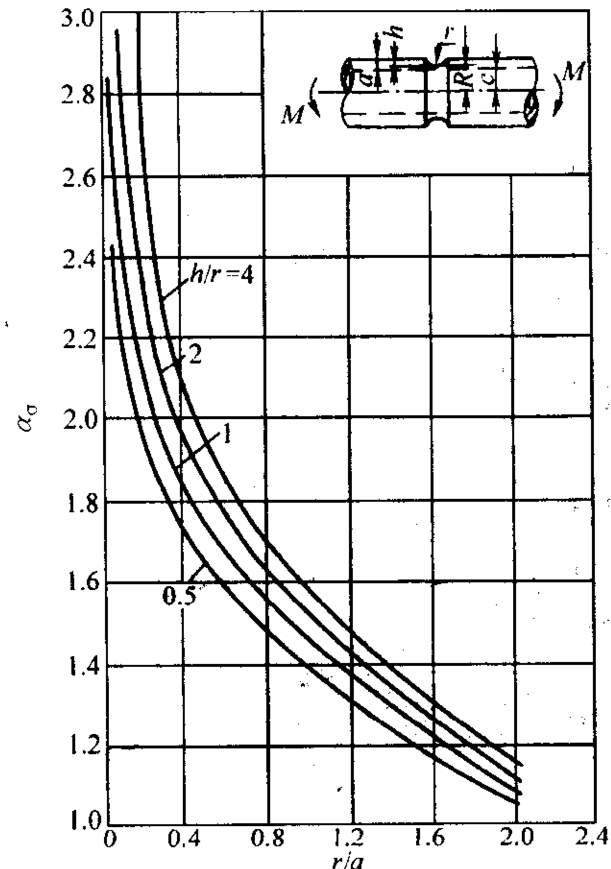


图 6.5-69　有环形槽的空心轴弯曲时的理论应力集中系数

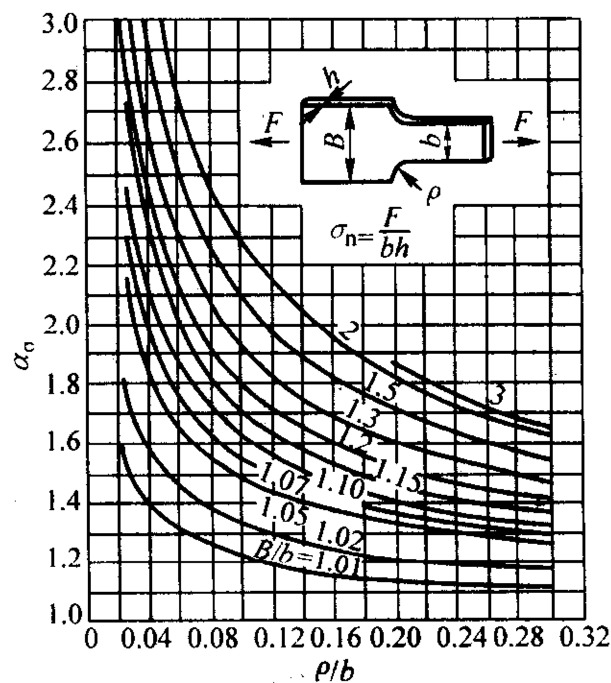


图 6.5-70　有肩板拉伸时的理论应力集中系数

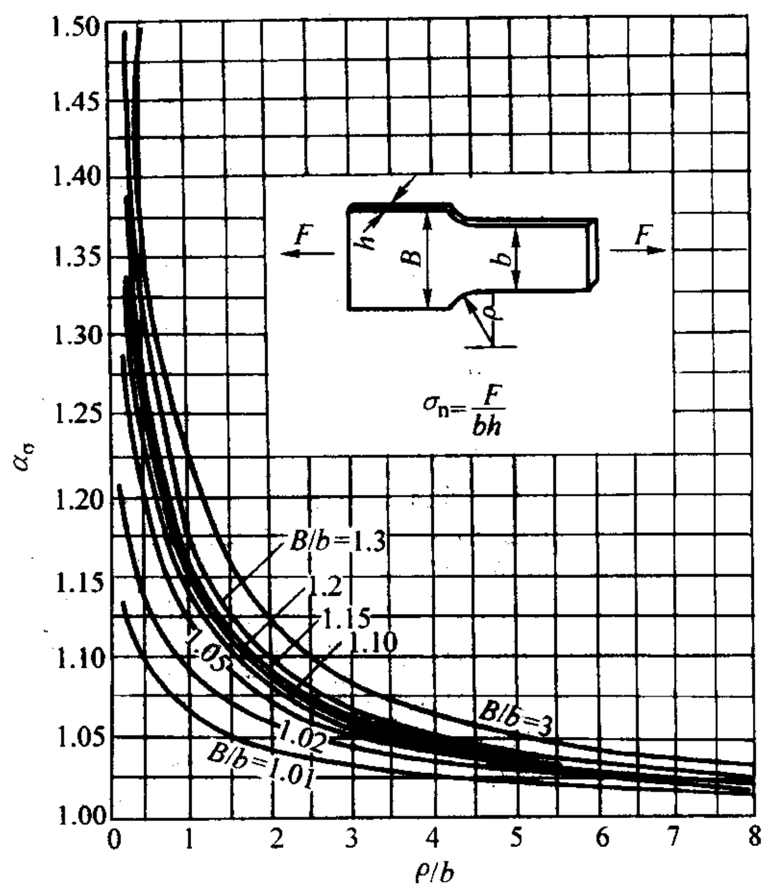


图 6.5-71　有肩板拉伸时的理论应力集中系数

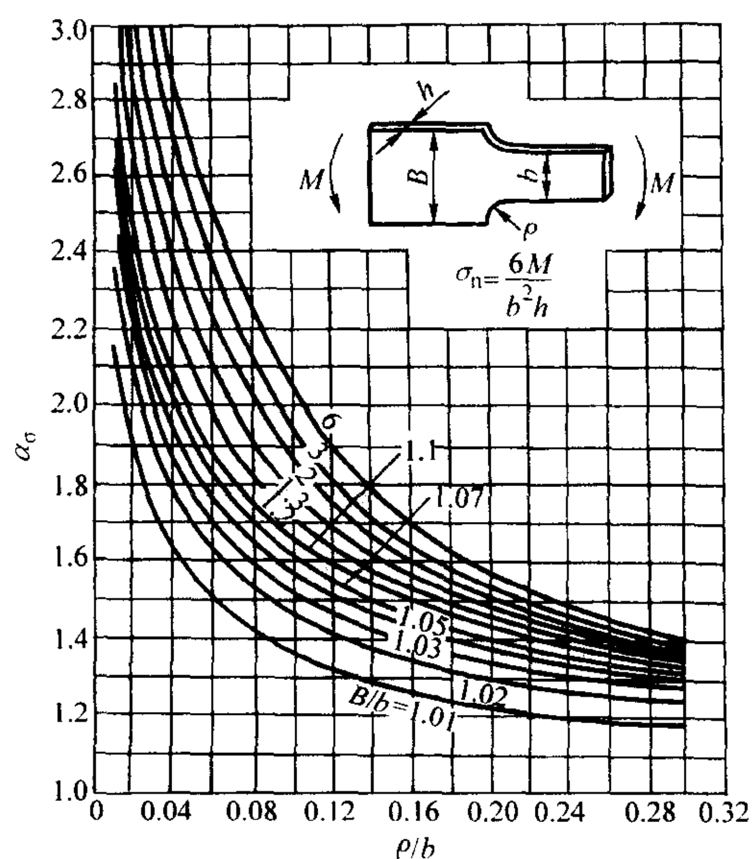


图 6.5-72　有肩板弯曲时的理论应力集中系数

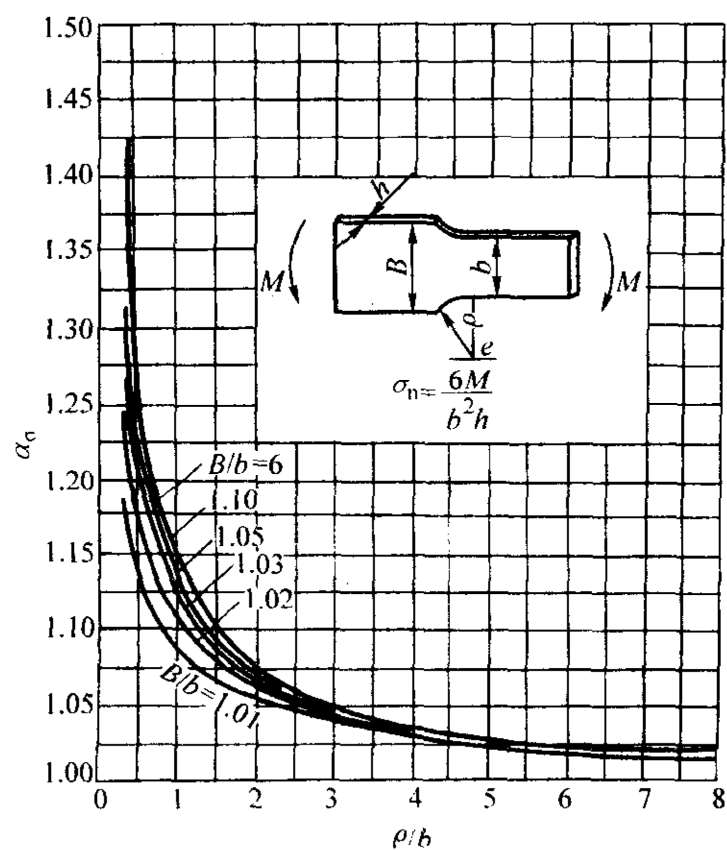


图 6.5-73　有肩板弯曲时的理论应力集中系数

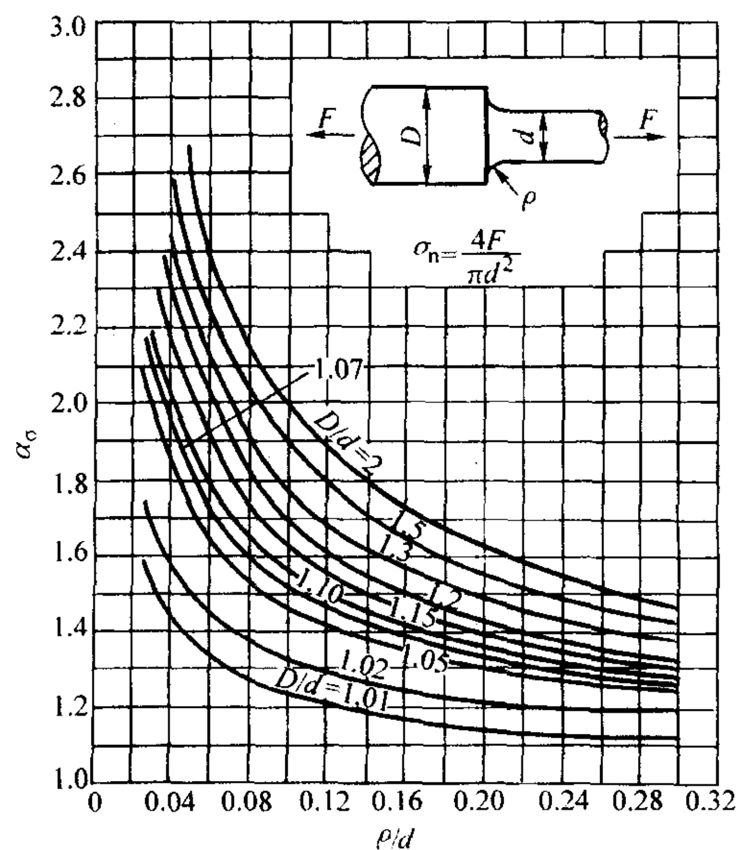


图 6.5-74　阶梯轴拉伸时的理论应力集中系数

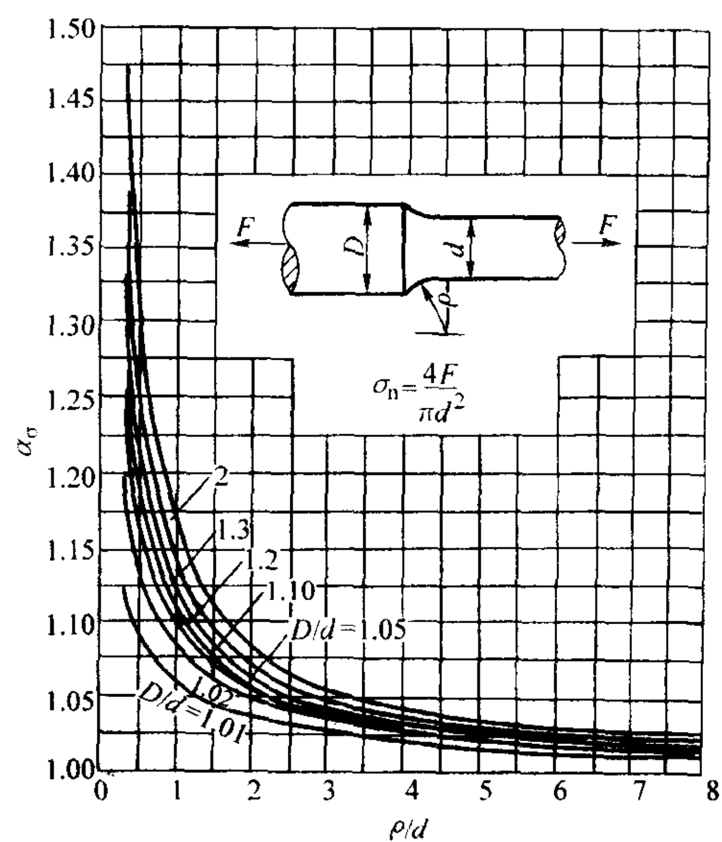


图 6.5-75　阶梯轴拉伸时的理论应力集中系数

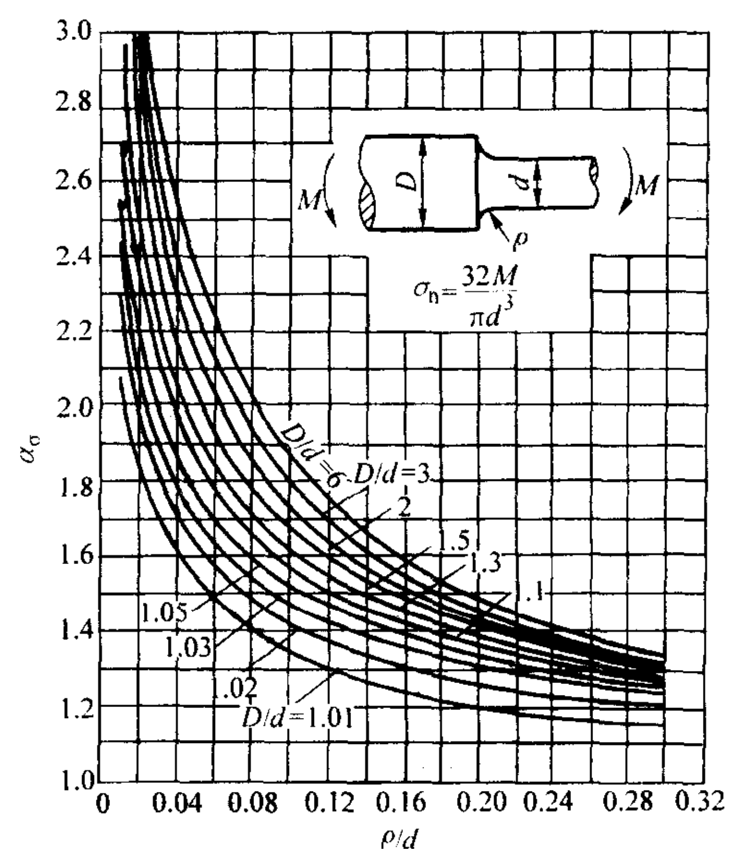


图 6.5-76　阶梯轴弯曲时的理论应力集中系数

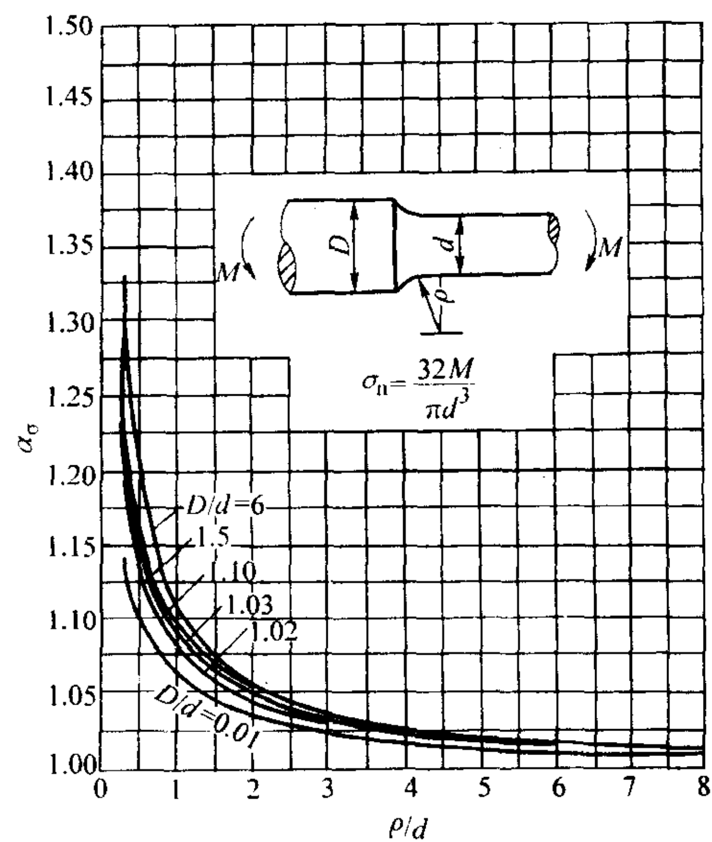


图 6.5-77　阶梯轴弯曲时的理论应力集中系数

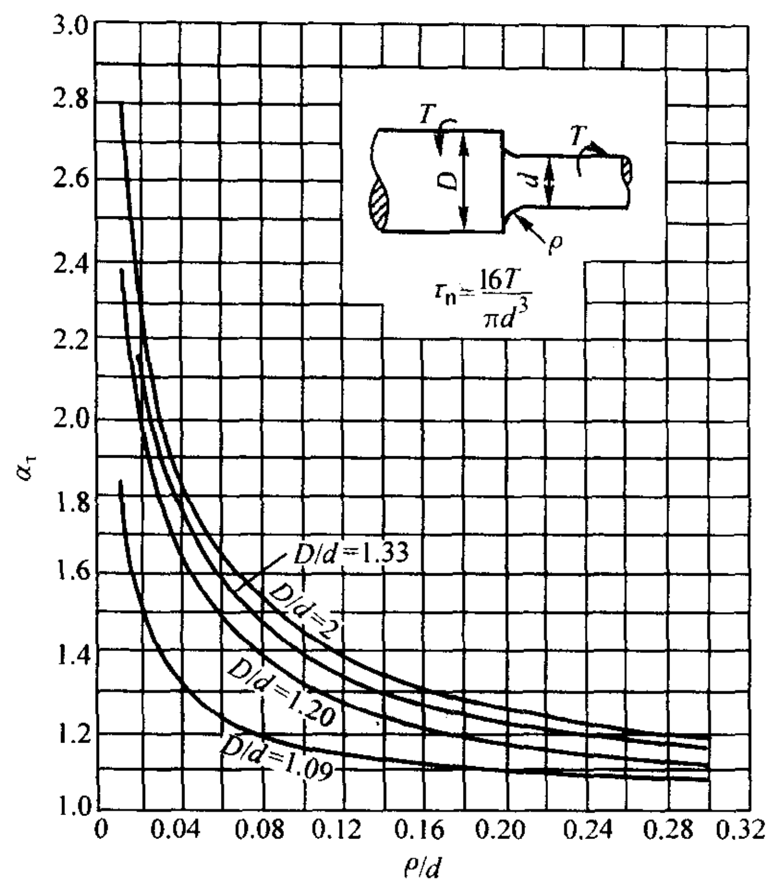


图 6.5-78 阶梯轴扭转时的理论应力集中系数

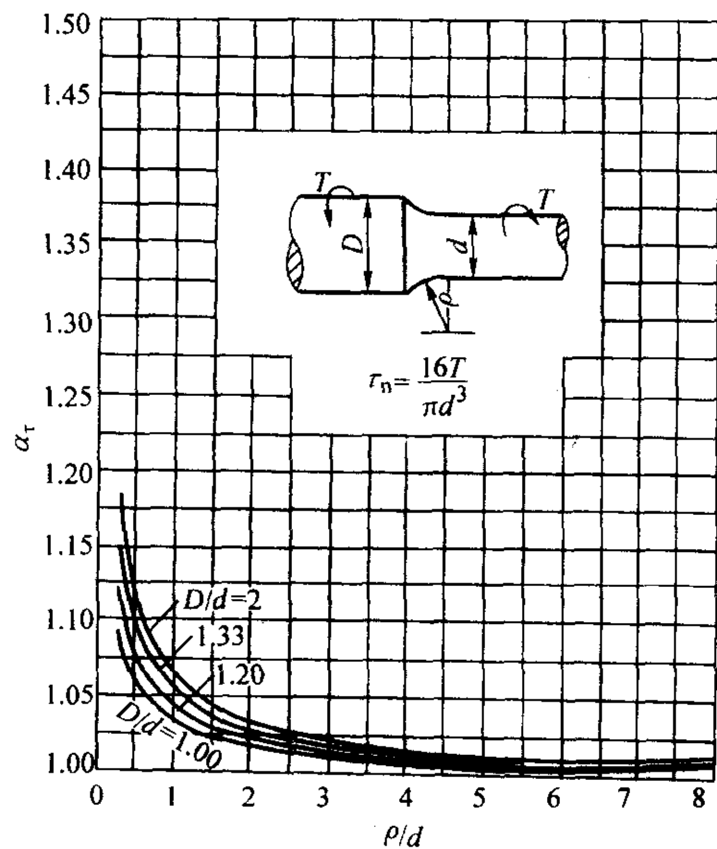


图 6.5-79 阶梯轴扭转时的理论应力集中系数

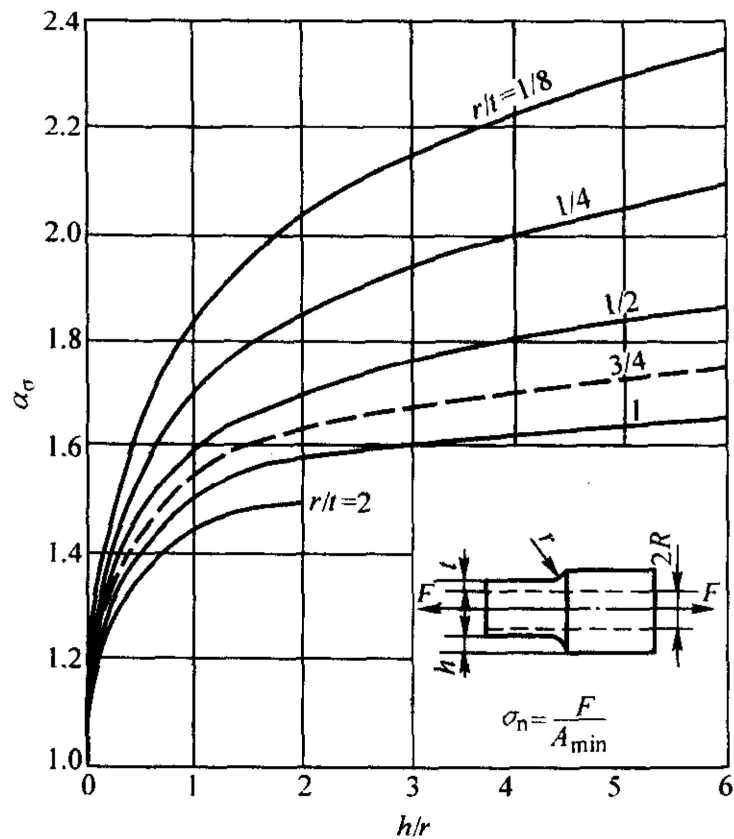


图 6.5-80 空心阶梯轴拉伸时的理论应力集中系数

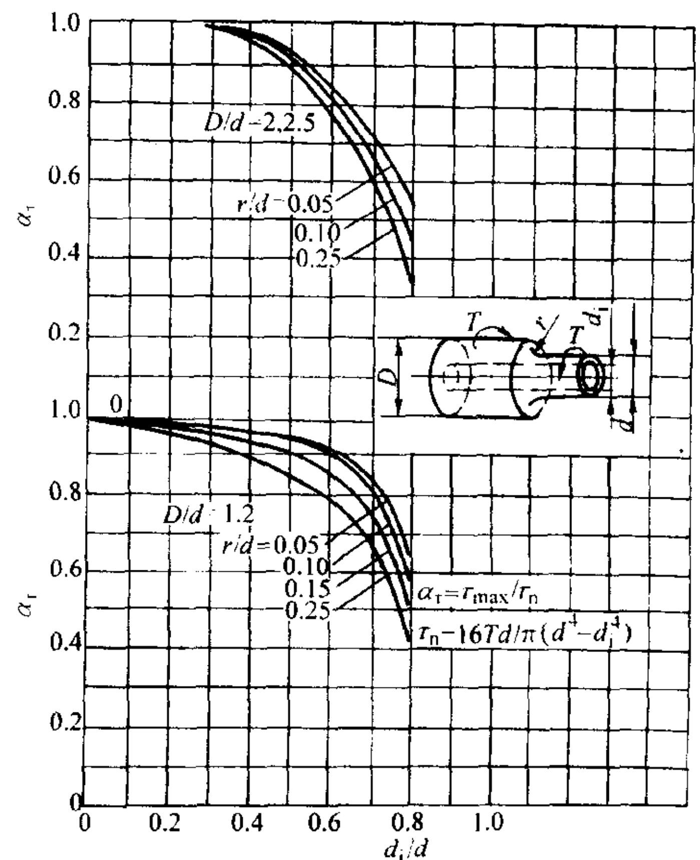


图 6.5-81 空心阶梯轴扭转时的理论应力集中系数

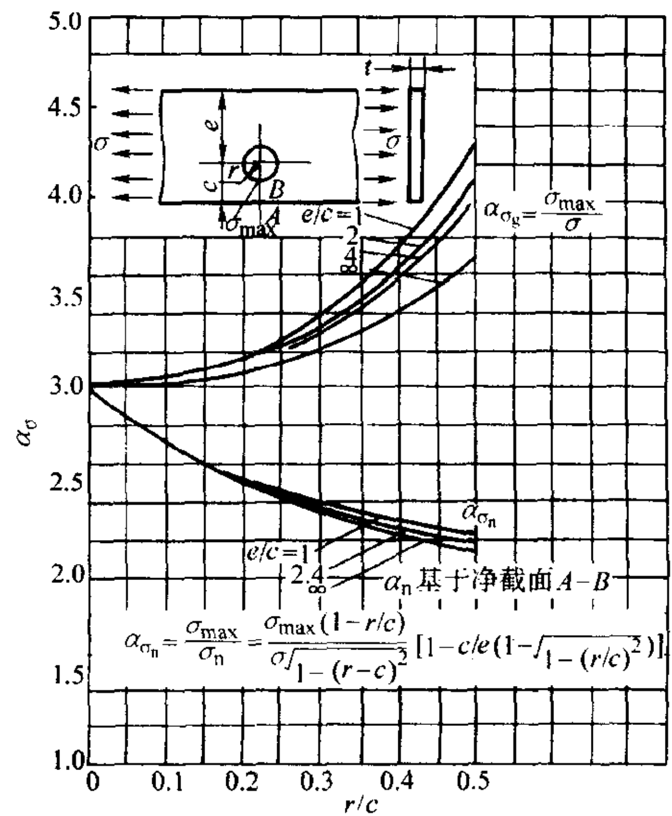


图 6.5-82 带偏心圆孔的受拉扁杆的理论应力集中系数

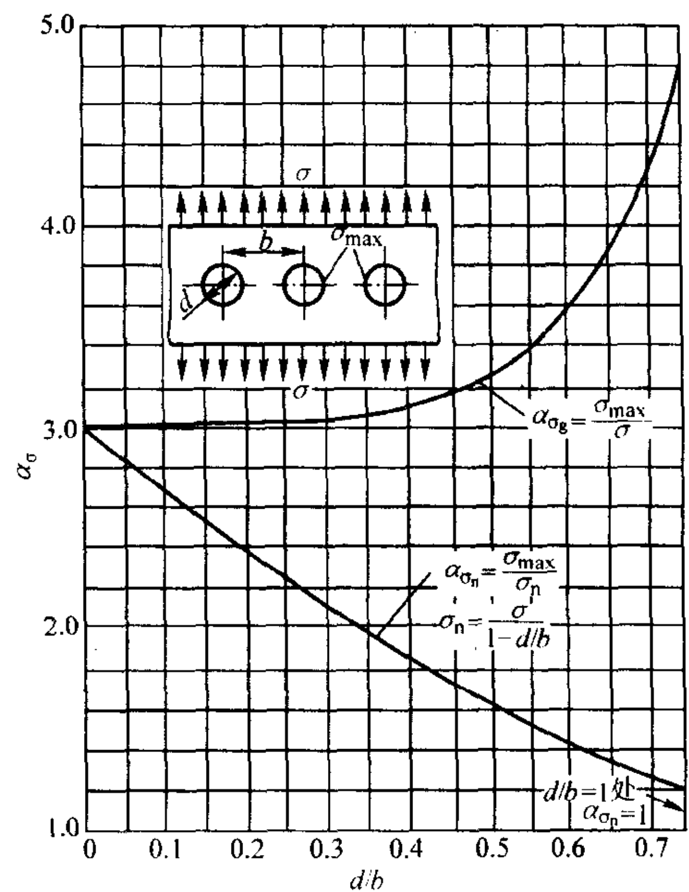


图 6.5-83 多孔受拉板（应力方向与孔的轴线垂直）的理论应力集中系数

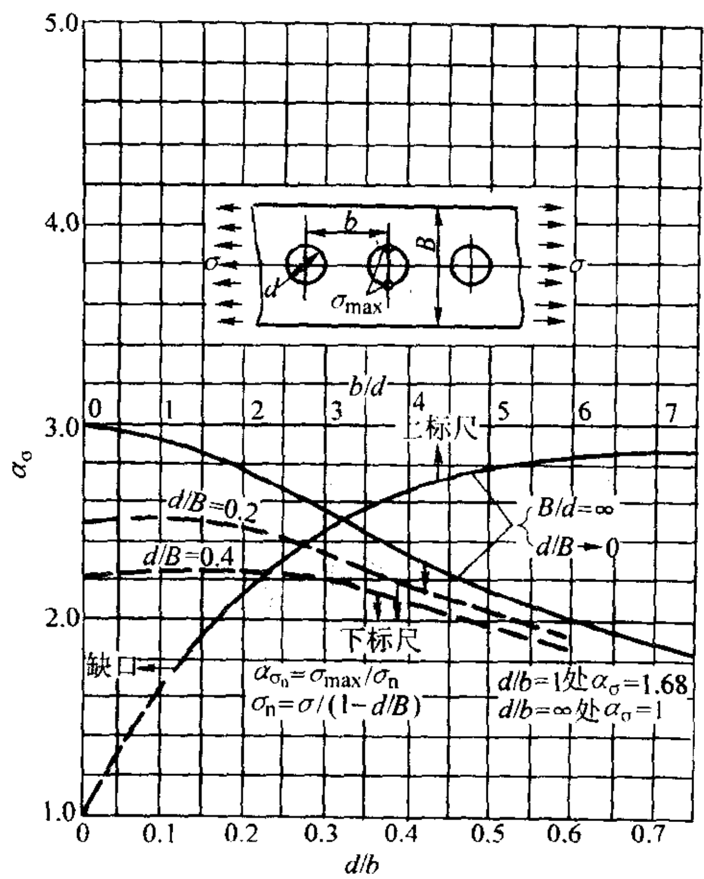


图 6.5-84　多孔受拉板（应力方向与孔的轴线平行）**的理论应力集中系数**

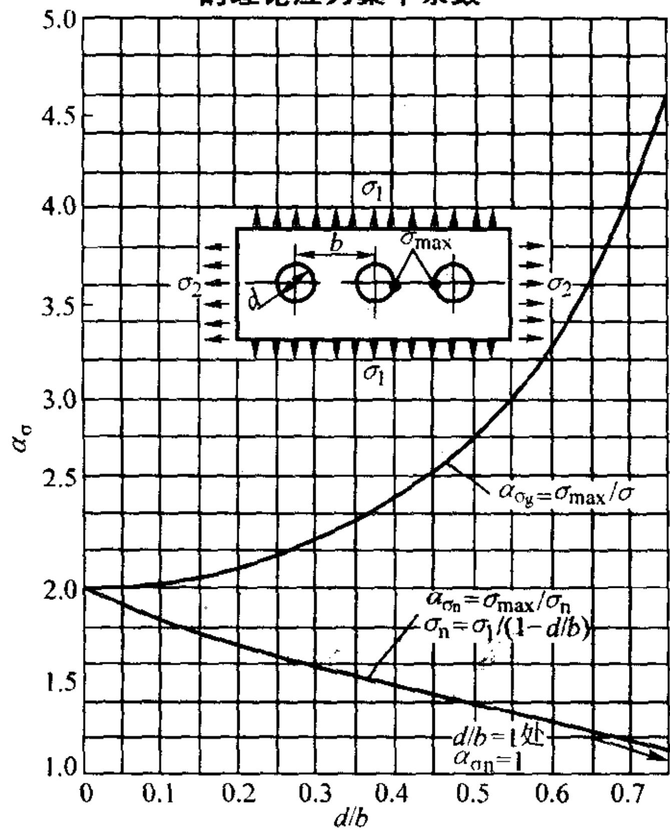


图 6.5-85　受双方拉伸的单排多孔板的理论应力集中系数

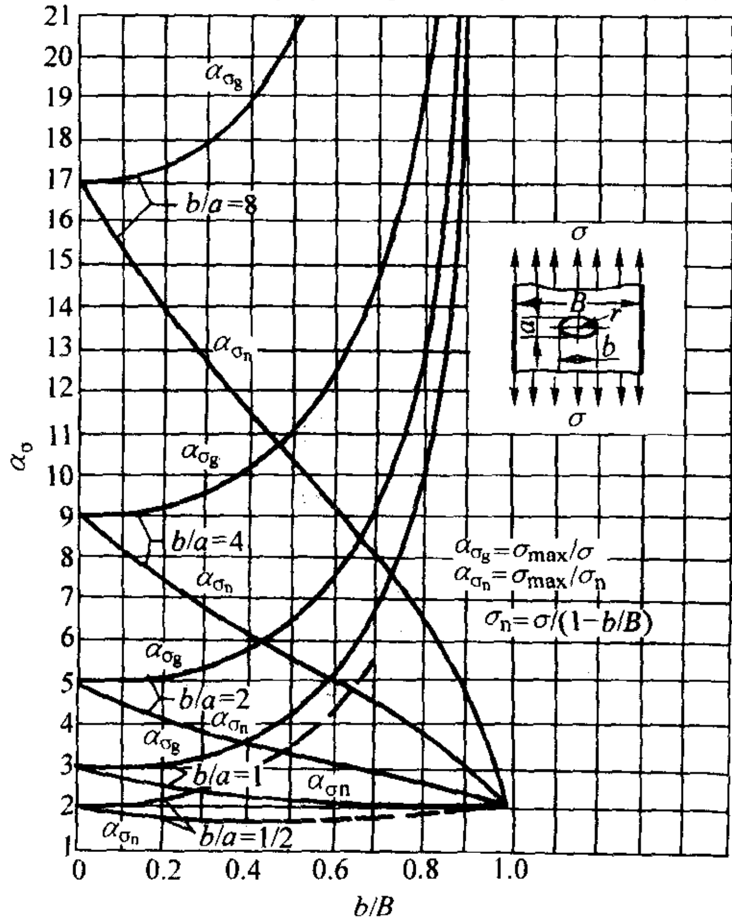


图 6.5-86　带椭圆孔的有限宽受拉板的理论应力集中系数

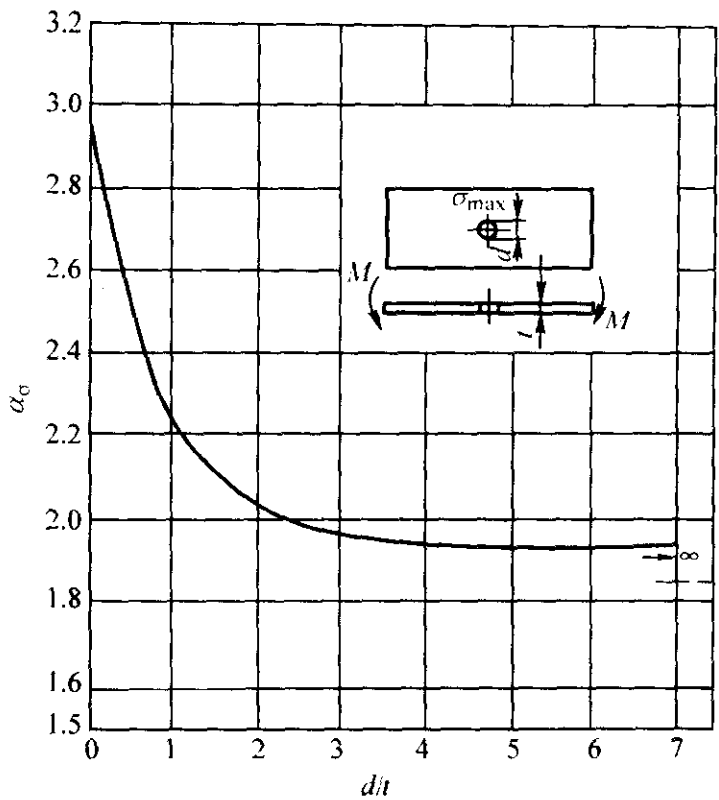


图 6.5-87　中央有孔的板弯曲的理论应力集中系数

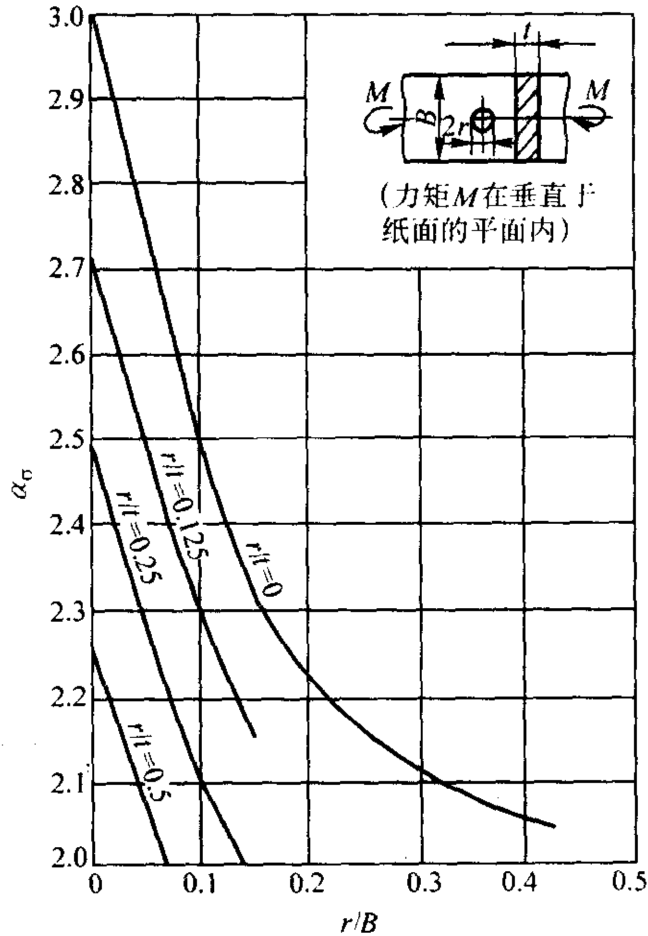


图 6.5-88　中央有孔的板弯曲的理论应力集中系数

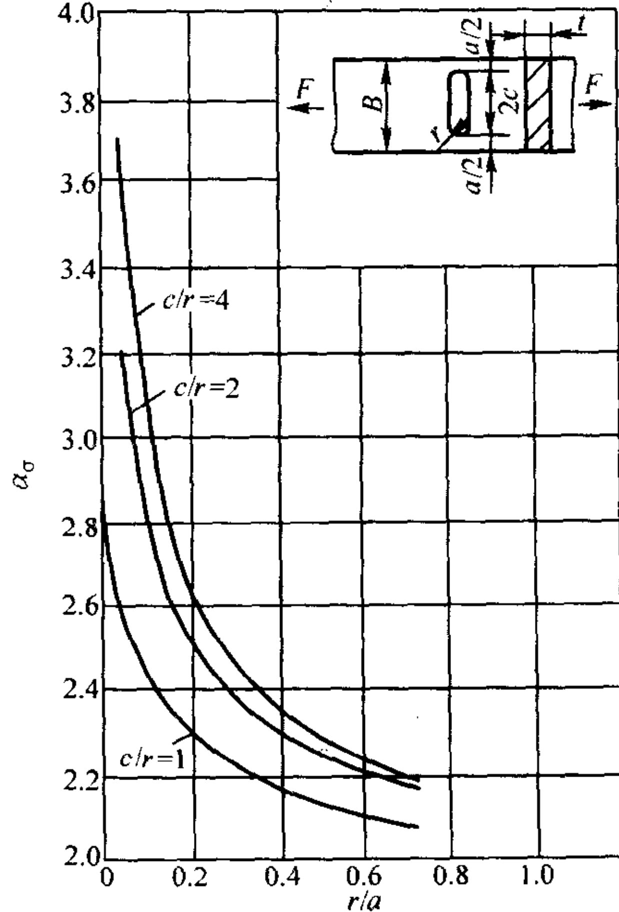


图 6.5-89　有长孔的板拉伸的理论应力集中系数

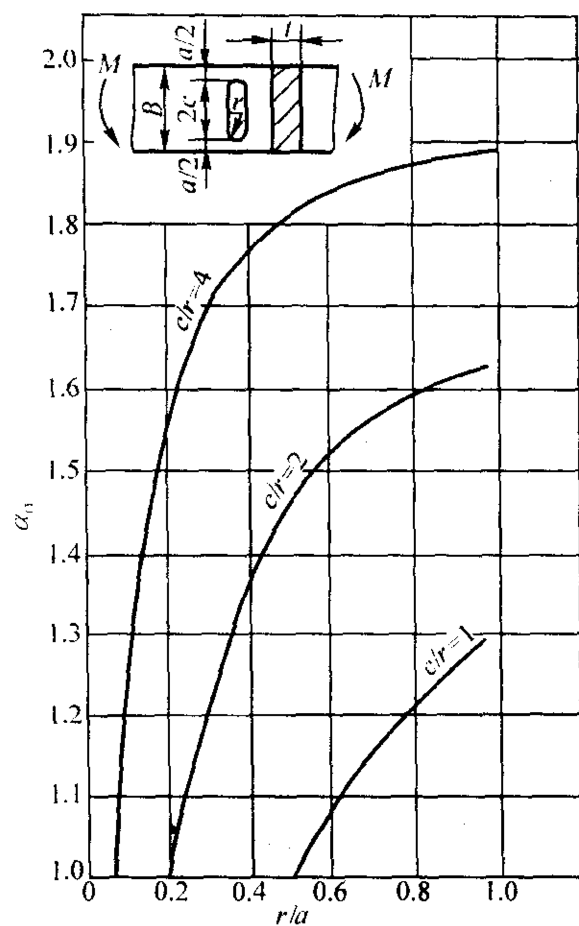


图 6.5-90　有长孔的板弯曲的理论应力集中系数

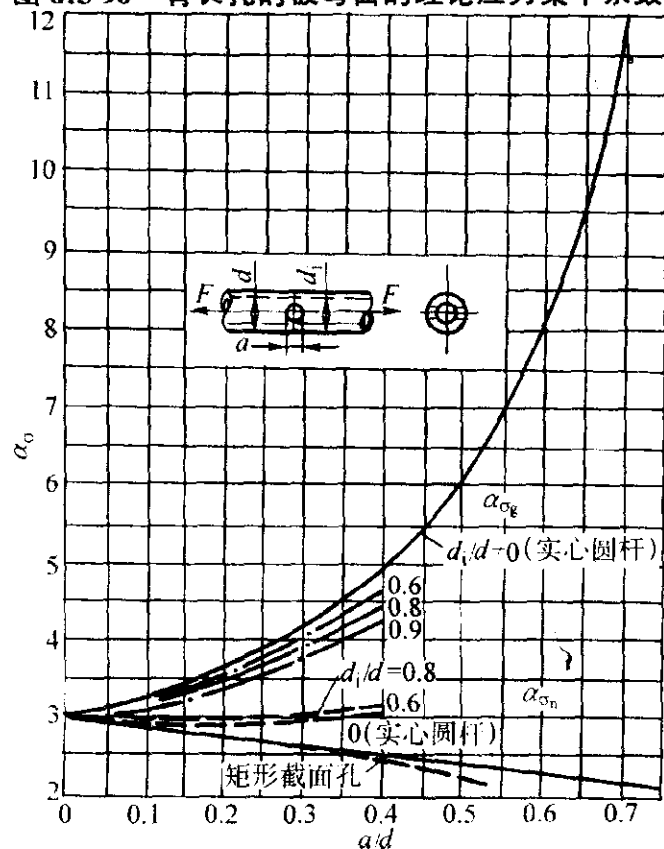


图 6.5-91　带通孔的受拉圆杆（管）的理论应力集中系数

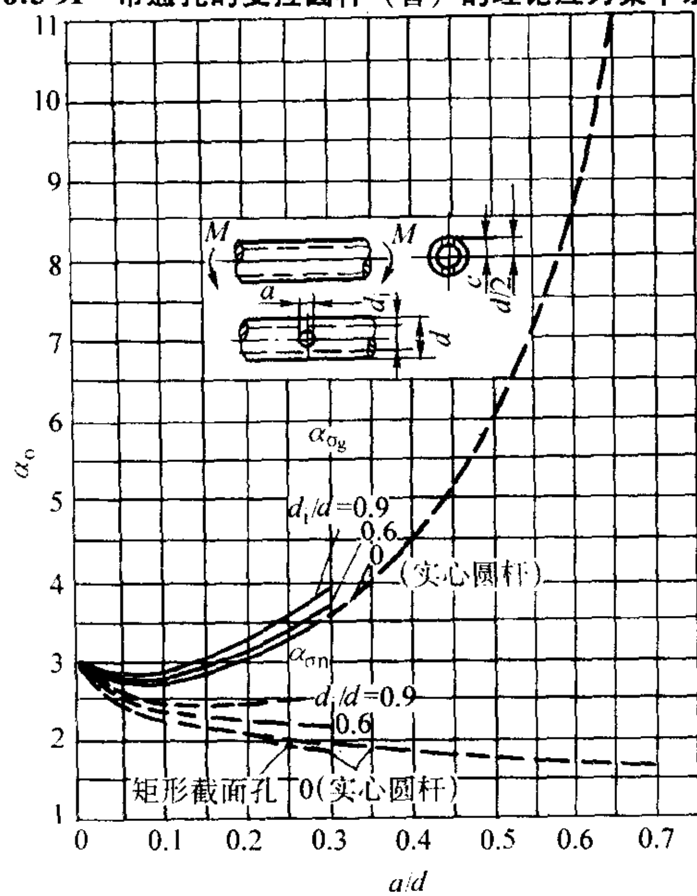


图 6.5-92　带通孔的受弯圆杆（管）的理论应力集中系数

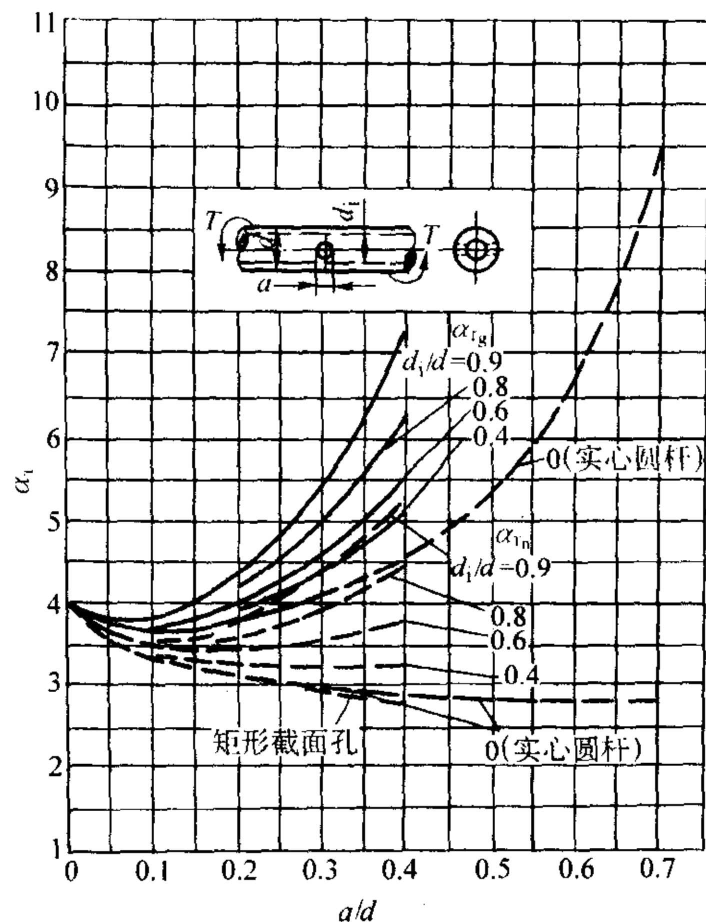


图 6.5-93　常通孔的受扭圆杆（管）的理论应力集中系数

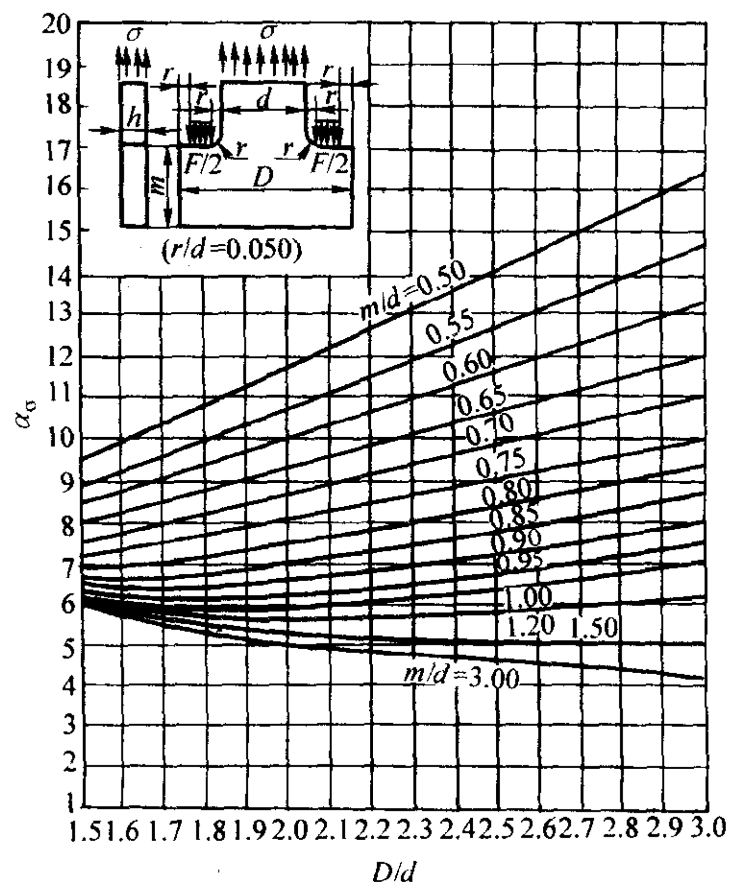


图 6.5-94　有肩板受均布力的理论应力集中系数

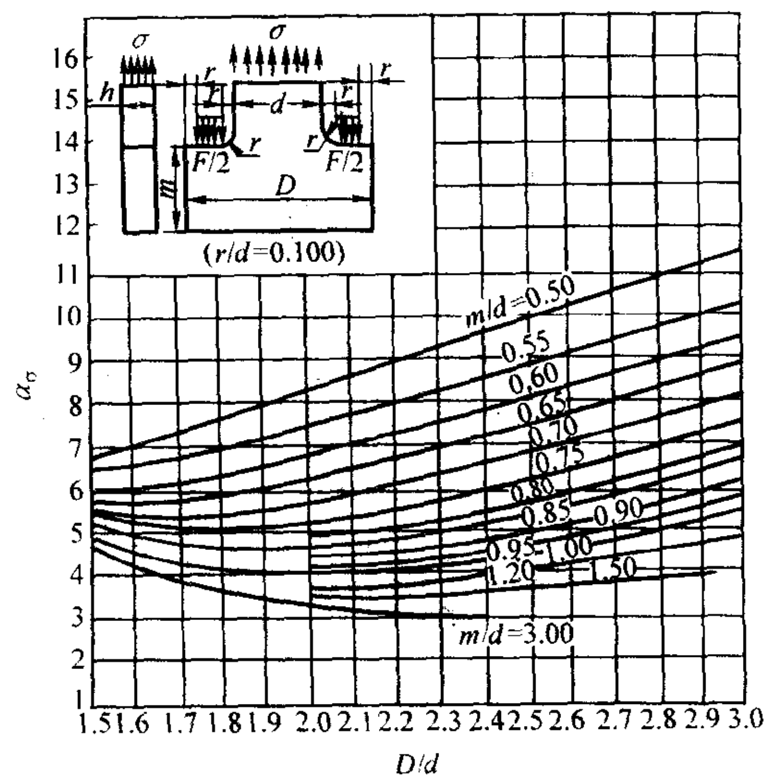


图 6.5-95　有肩板受均布力的理论应力集中系数

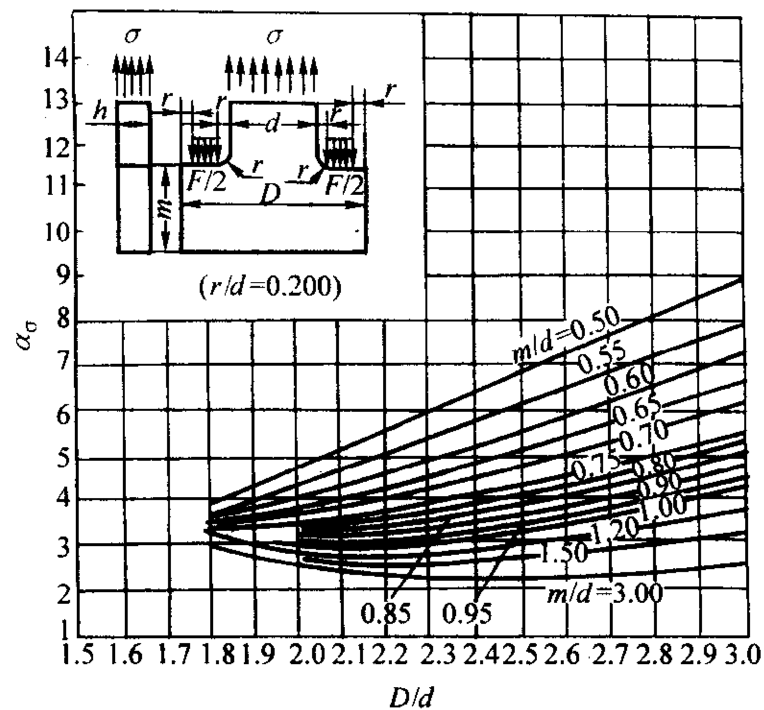


图 6.5-96　有肩板受均布力的理论应力集中系数

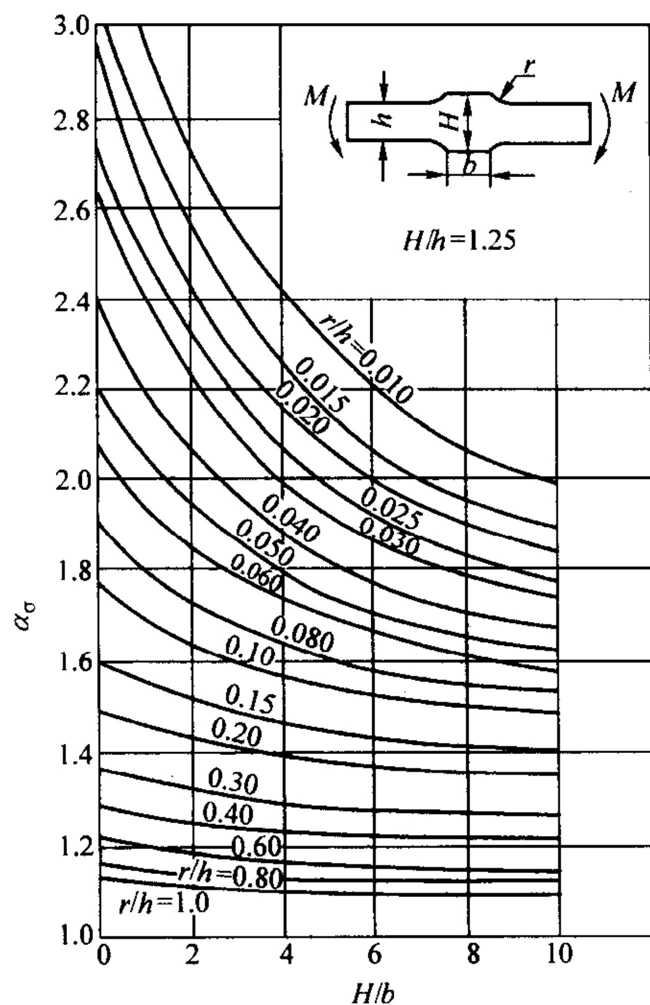


图 6.5-97　有凸台的板弯曲的理论应力集中系数

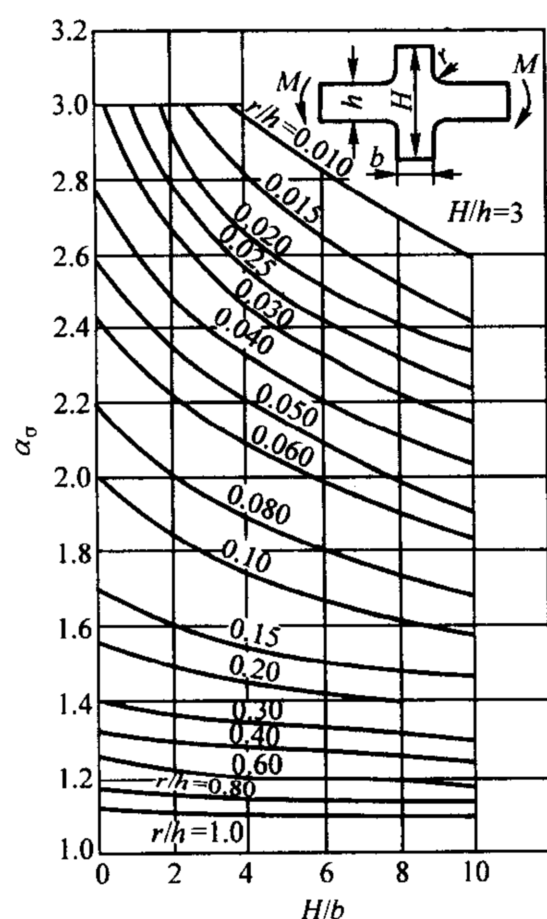


图 6.5-98　有凸台的板弯曲的理论应力集中系数

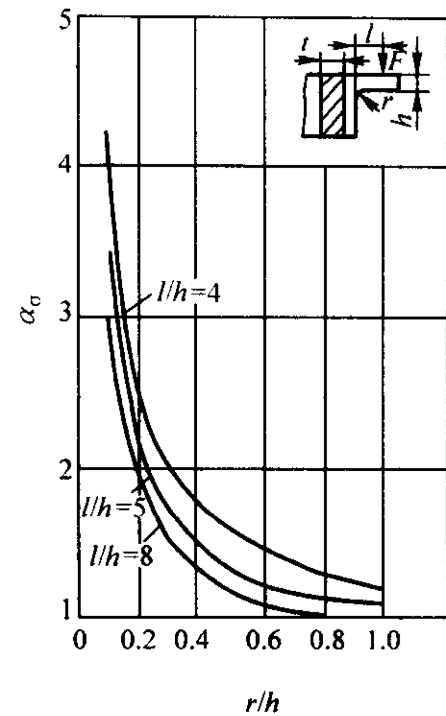


图 6.5-99　L 形截面受集中力弯曲的理论应力集中系数

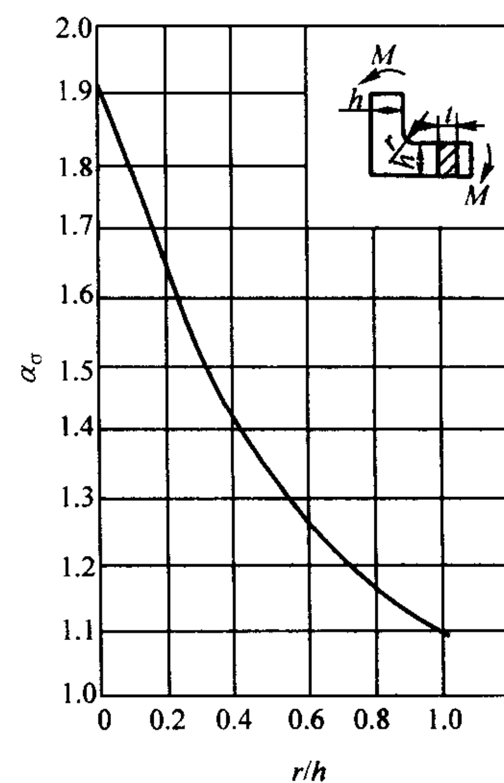
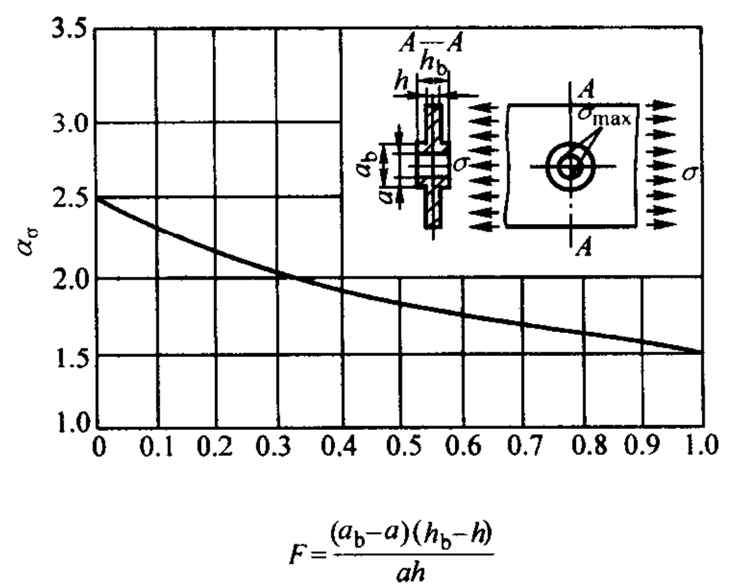


图 6.5-100　L 形截面受弯矩的理论应力集中系数



$$F=\frac{(a_b-a)(h_b-h)}{ah}$$

图 6.5-101　有凸台的板拉伸的理论应力集中系数

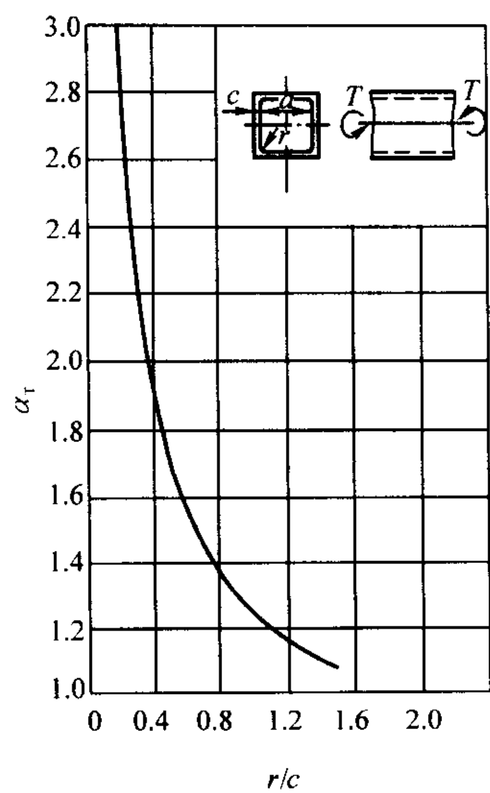


图 6.5-102 箱形截面杆扭转的理论应力集中系数

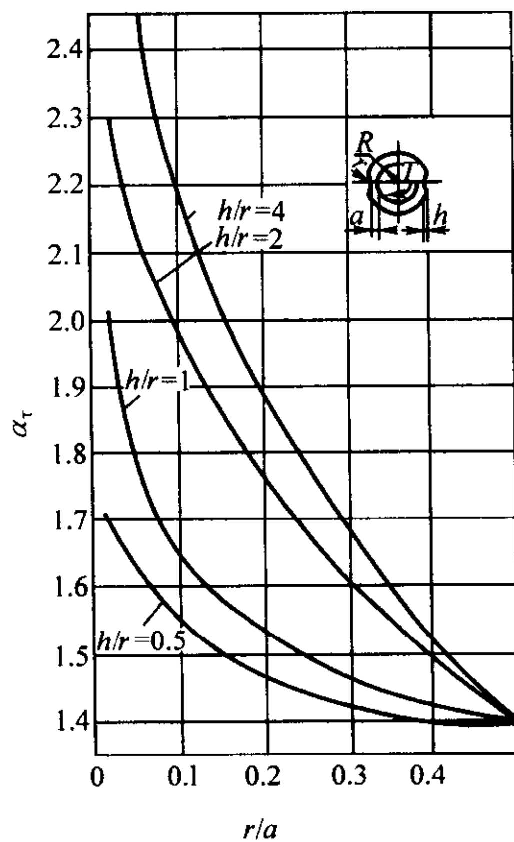


图 6.5-103 有两纵向圆槽的空心轴扭转的理论应力集中系数

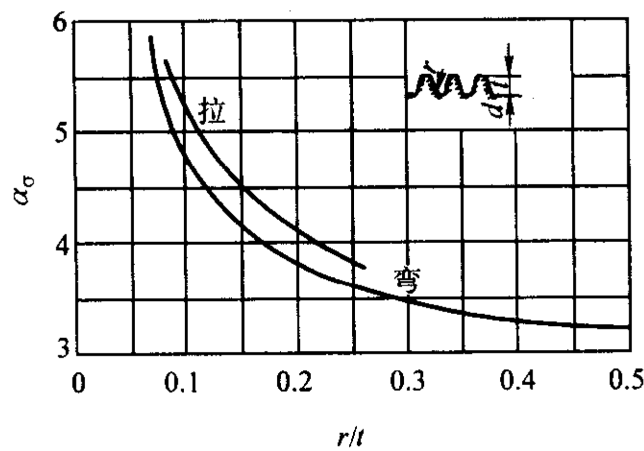


图 6.5-104 螺纹受拉伸或弯曲的理论应力集中系数

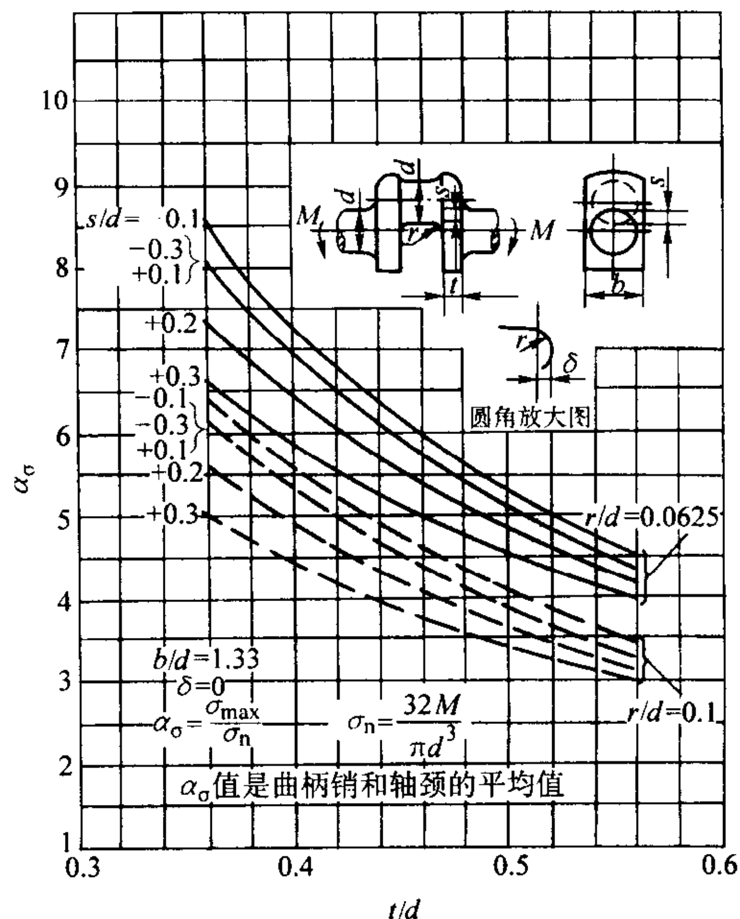


图 6.5-105 曲轴的弯曲的理论应力集中系数

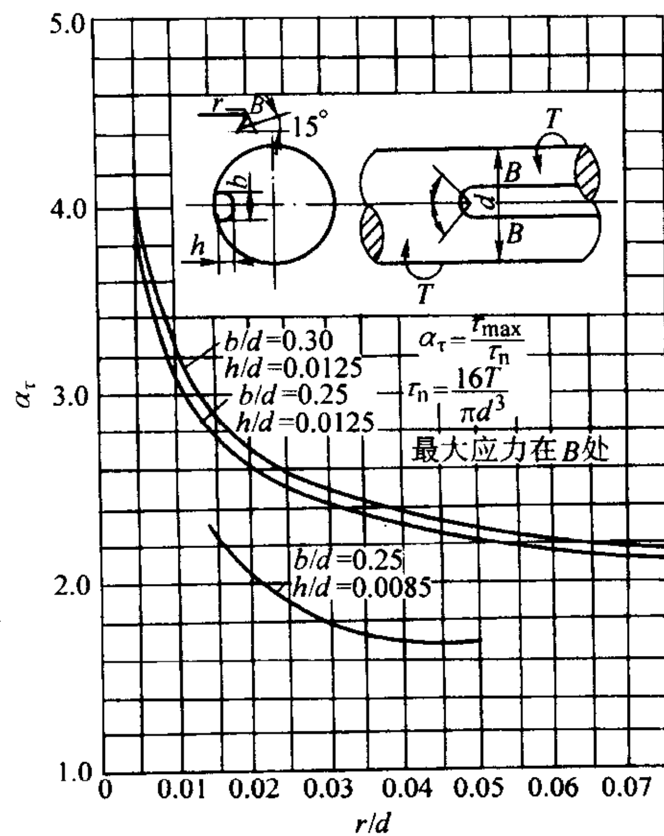


图 6.5-106 有端部半圆形键槽的受扭轴的理论应力集中系数

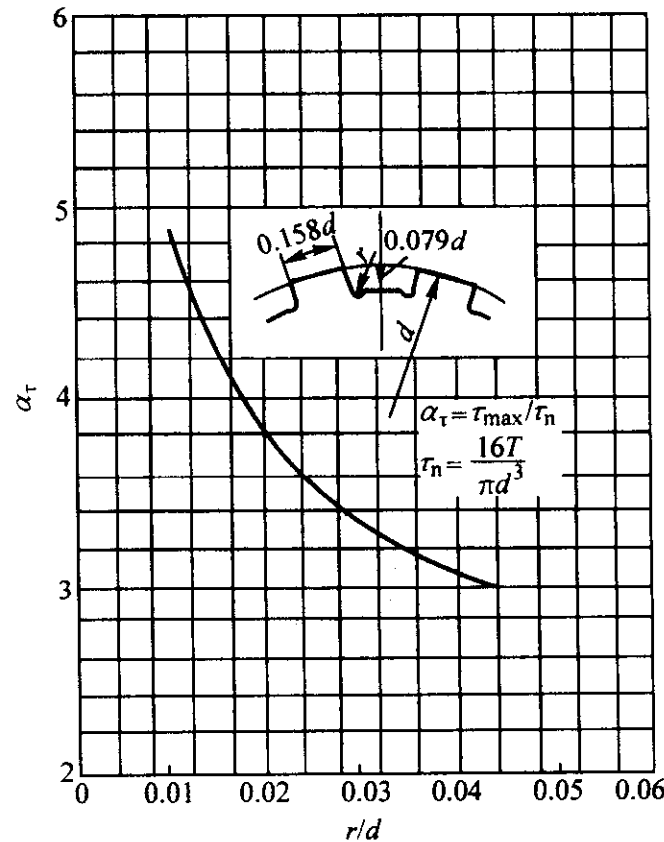


图 6.5-107 花键轴扭转的理论应力集中系数

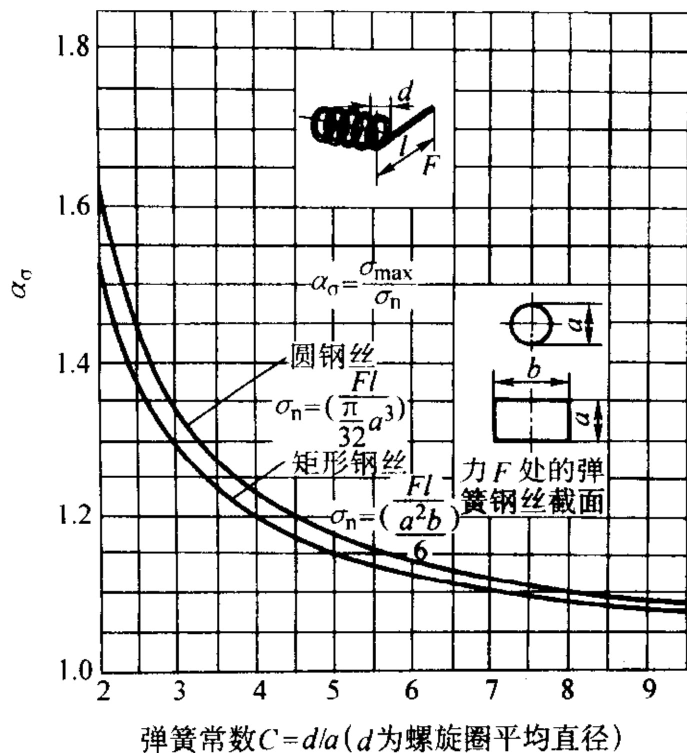


图 6.5-108 螺旋弹簧的扭转的理论应力集中系数

3.1.2 有效应力集中系数

理论应力集中系数的大小，不能作为由于存在局部峰值应力而使疲劳强度降低的标准。而真实材料的内部是存在着各种各样的缺陷和不同的晶粒分布情况的。同时在应力集中区的局部峰值应力常超过屈服点，使部分材料产生塑性变形，从而使应力重新分配。这就使得零件的疲劳强度不仅要由零件的几何形状所决定，而且还与零件的材料性质以及载荷类型等因素有关。

在循环应力条件下，把实际衡量应力集中对疲劳强度影响的系数，称为有效应力集中系数 K_σ 或 K_τ。在载荷条件和绝对尺寸相同时，循环应力下的有效应力集中系数，等于光滑试样与有效应力集中试样的疲劳极限之比，即

$$K_\sigma=\frac{\sigma_{-1}}{(\sigma_{-1})_K} \text{或} K_\tau=\frac{\tau_{-1}}{(\tau_{-1})_K} \quad (6.5\text{-}12)$$

式中，σ_{-1}和 τ_{-1}为光滑试样对称循环弯曲（或拉压）的疲劳极限和对称循环扭转的疲劳极限；$(\sigma_{-1})_K$ 和 $(\tau_{-1})_K$ 为有应力集中试样对称循环弯曲（或拉压）的疲劳极限和对称循环扭转的疲劳极限。

有效应力集中系数 K，总是小于理论应力集中系数 α。为了在数量上估计 K 与 α 之间的差别，引入材料对应力集中的敏性系数 q，它们之间的关系为

对弯曲或拉压：

$$q_\sigma=\frac{K_\sigma-1}{\alpha_\sigma-1}$$

对扭转：

$$q_\tau=\frac{K_\tau-1}{\alpha_\tau-1}$$

或写成

$$\left.\begin{aligned}K_\sigma&=1+q_\sigma(\alpha_\sigma-1)\\K_\tau&=1+q_\tau(\alpha_\tau-1)\end{aligned}\right\} \quad (6.5\text{-}13)$$

如 $q_\sigma=0$ 和 $q_\tau=0$，则 $K_\sigma=1$ 和 $K_\tau=1$，没有应力集中产生，即材料对应力集中不敏感。如 $q_\sigma=1$ 和 $q_\tau=1$，则 $K_\sigma=\alpha_\sigma$ 和 $K_\tau=\alpha_\tau$，即材料对应力集中十分敏感。q 值一般在 0 与 1 之间，在实际应用中，常设 $q_\sigma=q_\tau=q$。

求有效应力集中系数有两种方法：一是直接用零部件在特定材料及形状下试验求得；另一种按照式（6.5-13）的关系，由零件的几何形状查得相应的理论应力集中系数 α，当该材料与有关尺寸确定的敏性系数 q 已知时，即可求得有效应力集中系数。前者最能表征实际情况，所以在疲劳强度设计中，应尽可能建议采用。

钢材的敏性系数 q，可查图 6.5-109。

某些典型的零件结构的有效应力集中系数如图 6.5-110～图 6.5-135 及表 6.5-6、表 6.5-7 所示。

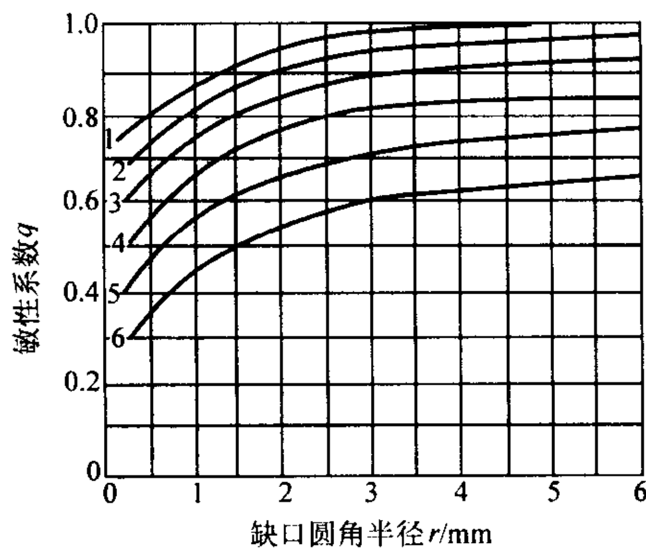


图 6.5-109 钢的应力集中敏性系数与材料的力学性能和缺口圆角半径的关系

1—σ_b = 1 300 MPa；2—σ_b = 1 200 MPa；3—σ_b = 1 000 MPa；4—σ_b = 800 MPa；5—σ_b = 600 MPa；6—σ_b = 400 MPa

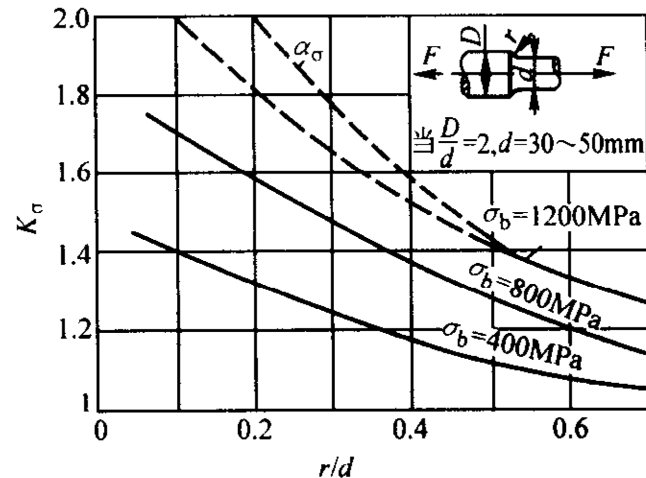


图 6.5-110 阶梯钢轴的对称拉压的有效应力集中系数（实线）

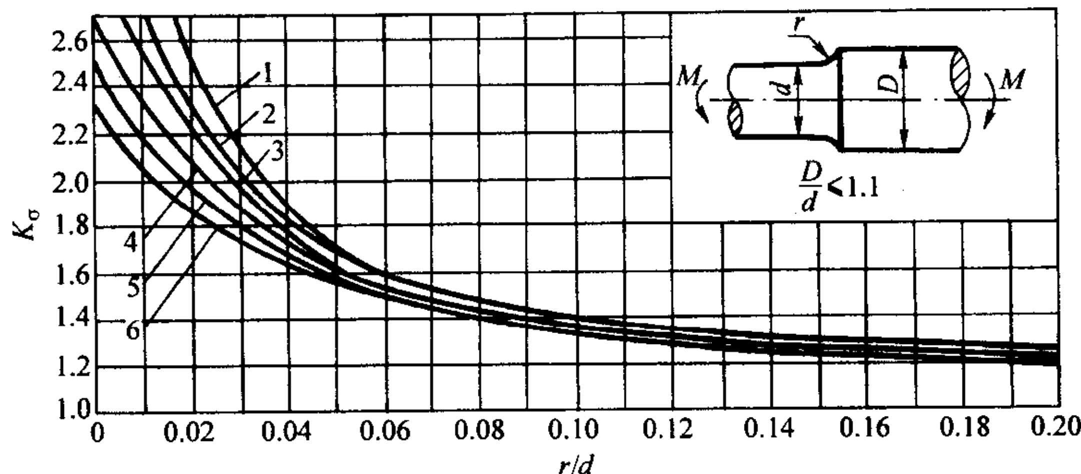


图 6.5-111 阶梯钢轴的弯曲的有效应力集中系数

1—σ_b≥1 000 MPa；2—σ_b = 900 MPa；3—σ_b = 800 MPa；4—σ_b = 700 MPa；5—σ_b = 600 MPa；6—σ_b≤500 MPa

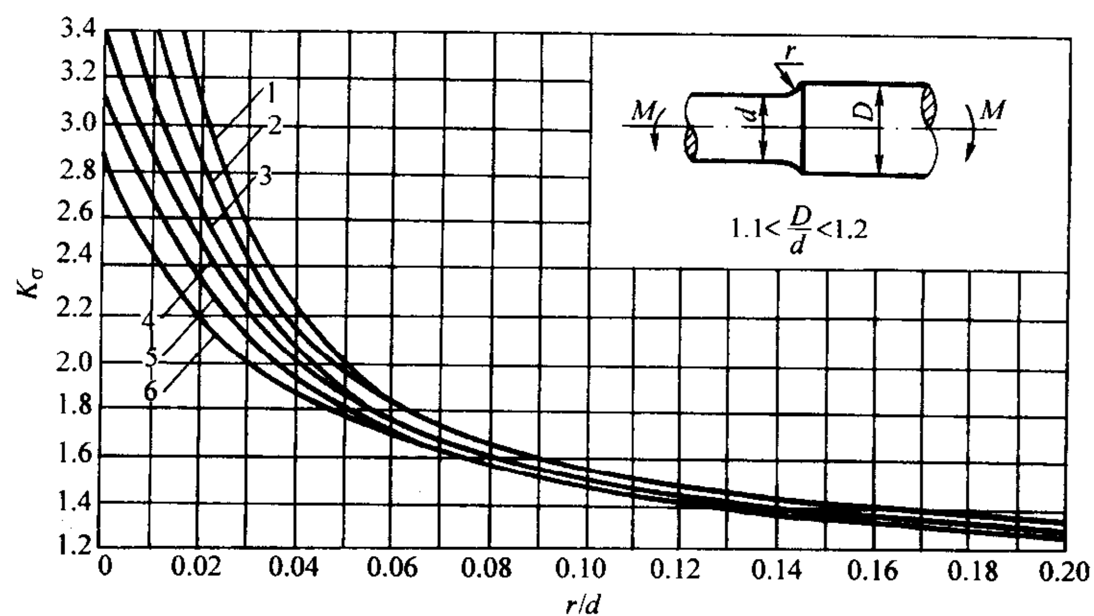


图 6.5-112　阶梯钢轴的弯曲的有效应力集中系数

1—$\sigma_b \geqslant 1\ 000$ MPa；2—$\sigma_b = 900$ MPa；3—$\sigma_b = 800$ MPa；
4—$\sigma_b = 700$ MPa；5—$\sigma_b = 600$ MPa；6—$\sigma_b \leqslant 500$ MPa

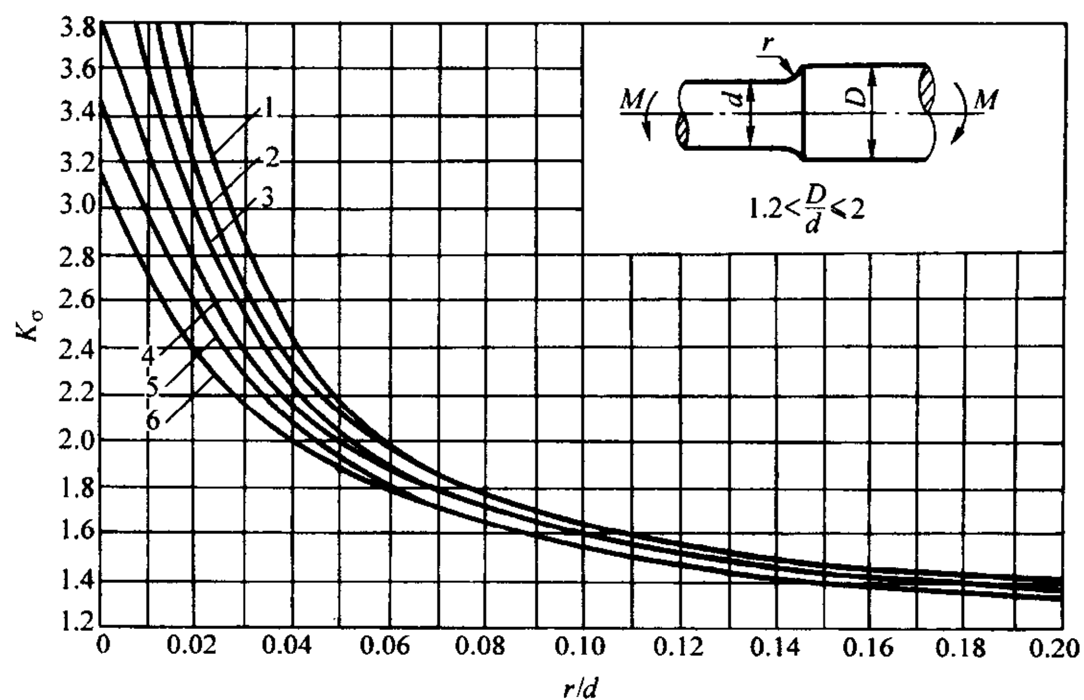


图 6.5-113　阶梯钢轴的弯曲的有效应力集中系数

1—$\sigma_b \geqslant 1\ 000$ MPa；2—$\sigma_b = 900$ MPa；3—$\sigma_b = 800$ MPa；
4—$\sigma_b = 700$ MPa；5—$\sigma_b = 600$ MPa；6—$\sigma_b \leqslant 500$ MPa

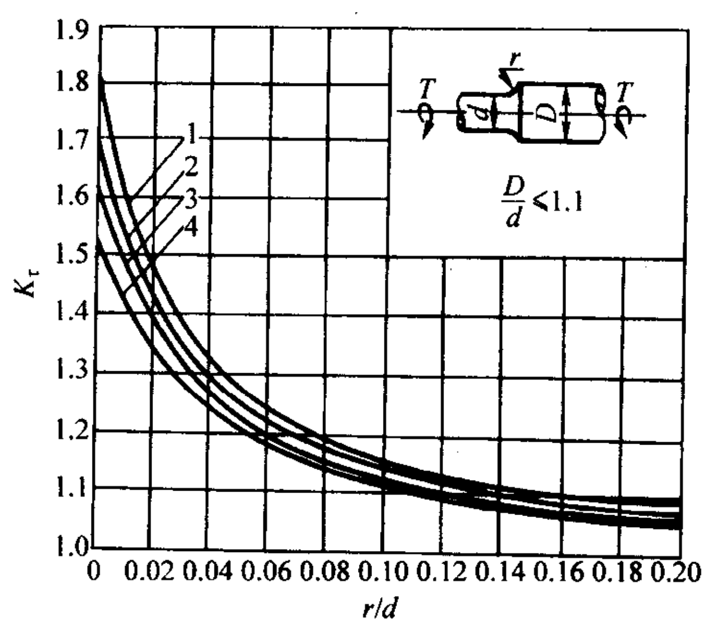


图 6.5-114　阶梯钢轴的扭转的有效应力集中系数

1—$\sigma_b \geqslant 1\ 000$ MPa；2—$\sigma_b = 900$ MPa；
3—$\sigma_b = 800$ MPa；4—$\sigma_b \leqslant 700$ MPa

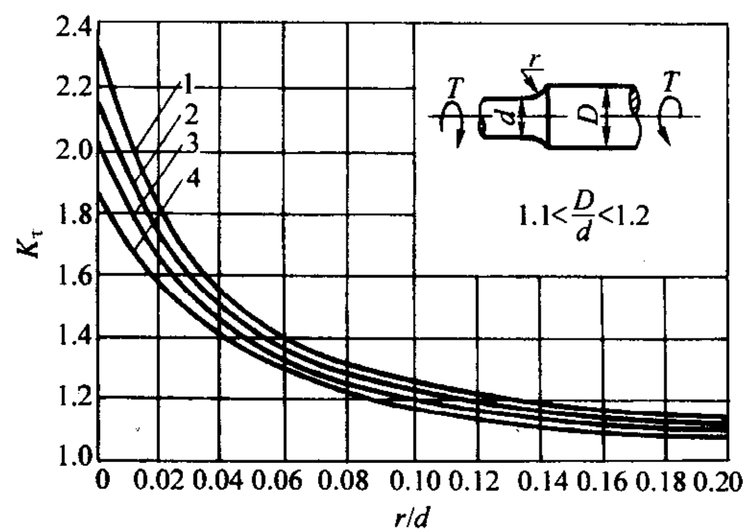


图 6.5-115　阶梯钢轴的扭转的有效应力集中系数

1—$\sigma_b \geqslant 1\ 000$ MPa；2—$\sigma_b = 900$ MPa；
3—$\sigma_b = 800$ MPa；4—$\sigma_b \leqslant 700$ MPa

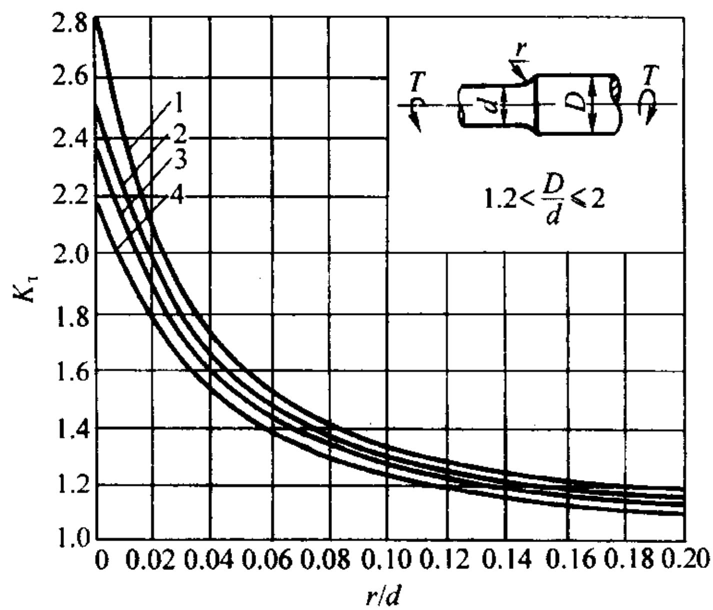


图 6.5-116 阶梯钢轴的扭转的有效应力集中系数

1—$\sigma_b \geqslant 1\ 000$ MPa; 2—$\sigma_b = 900$ MPa;

3—$\sigma_b = 800$ MPa; 4—$\sigma_b \leqslant 700$ MPa

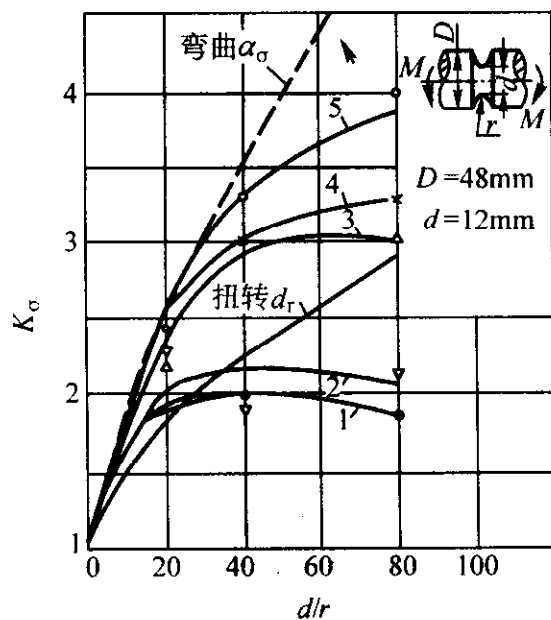


图 6.5-117 有环形深槽钢轴的旋转弯曲的有效应力集中系数

（虚线为理论应力集中系数）

1—$w_C = 0.25\%$; 2—$w_C = 0.38\%$;

3—$w_C = 0.75\%$; 4—Ni - Cr 钢; 5—Ni - Cr 钢

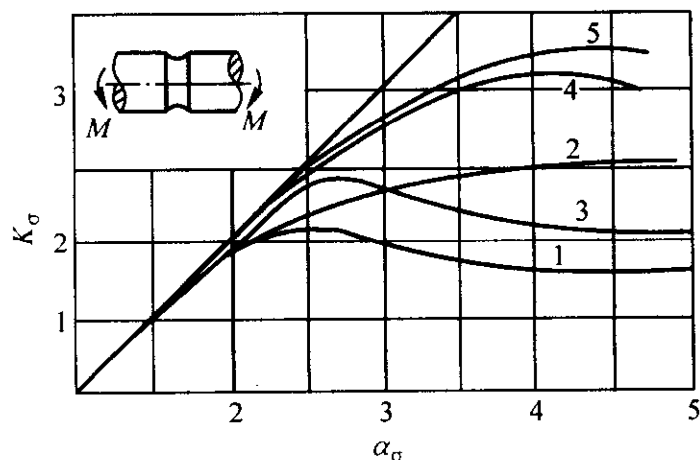
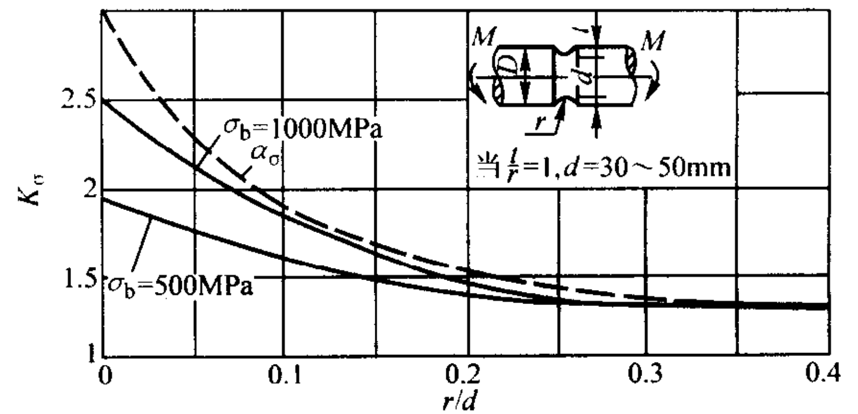


图 6.5-118 有环形槽钢轴的旋转弯曲的有效应力集中系数

1—$w_C = 0.22\%$; 2—$w_C = 0.25\%$;

3—$w_C = 0.38\%$; 4—$w_C = 0.76\%$;

5—$w_{Ni} = 2.8\%$; $w_{Cr} = 0.7\%$

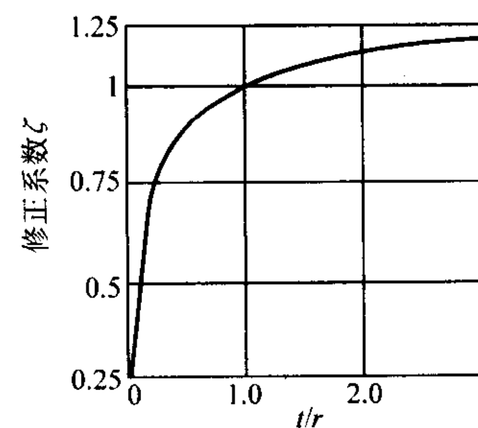


(a) 有环形槽钢轴(当$\frac{t}{r}$=1时)的对称弯曲的有效应力集中系数

(虚线为理论应力集中系数)

当$\frac{t}{r}\neq 1$时的有效应力集中系数的计算式为

$K_\sigma = 1 + \xi[(K_\sigma)_0 - 1]$



(b) 有环形槽钢轴当$\frac{D}{d}<2$时的有效应力集中系数的修正系数ξ

图 6.5-119 有环形槽钢轴的对称弯曲的有效应力集中系数

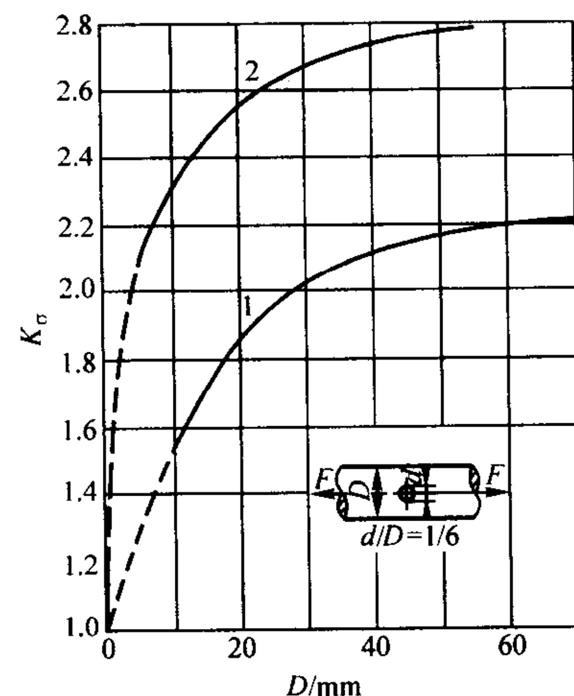


图 6.5-120 有横孔钢轴的拉压的有效应力集中系数

1—$w_C = 0.07\%$低碳钢，$\sigma_b = 330$ MPa;

2—Ni - Cr - Mo 钢（$w_{Co} = 0.43\%$,

$w_{Ni} = 2.64\%$, $w_{Cr} = 0.75\%$,

$w_{Mn} = 0.65\%$, $w_{Mo} = 0.58\%$,

$w_V = 0.05\%$）

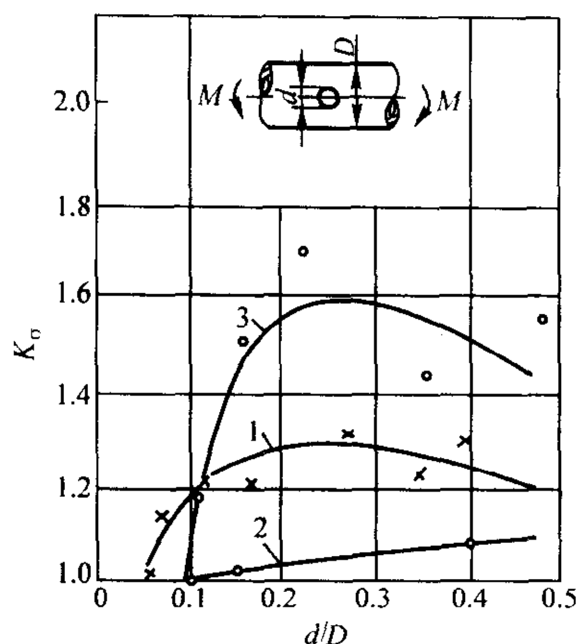


图 6.5-121 有横孔的空心铸铁圆棒的旋转弯曲的有效应力集中系数

1—球墨铸铁，$D=23$ mm；

2—孕育铸铁，$D=12$ mm；

3—孕育铸铁，$D=23$ mm

（铁素体包围的片状石墨的铸铁称孕育铸铁）

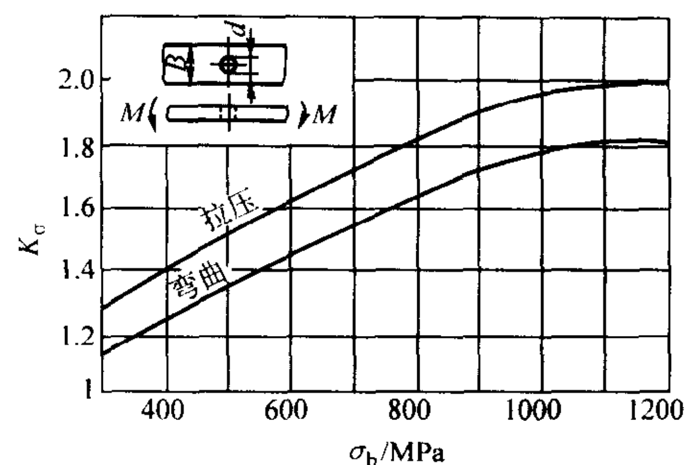


图 6.5-122 有孔钢板的有效应力集中系数

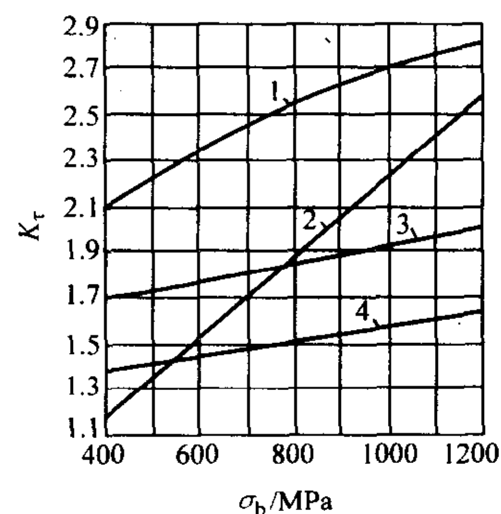
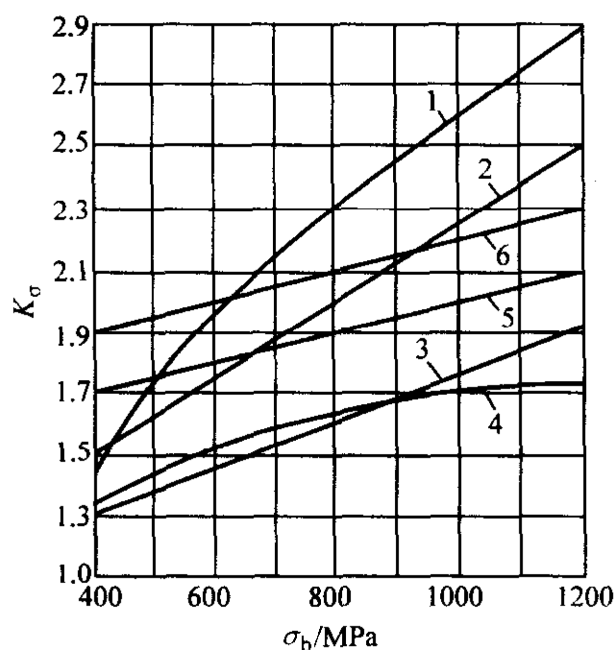


图 6.5-123 有键槽、横孔的钢轴扭转的有效应力集中系数

1—矩形花键；2—渐开线花键；3—键槽；

4—横孔 $\frac{d}{D}=0.05\sim0.25$



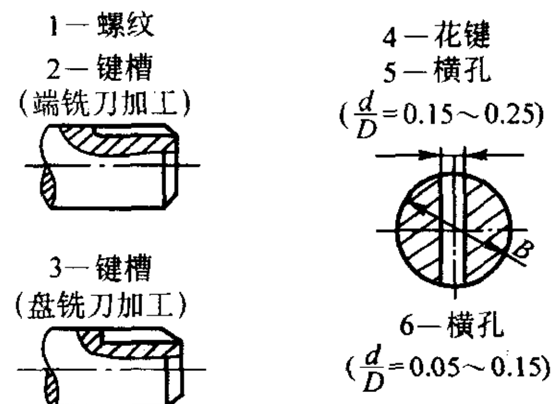
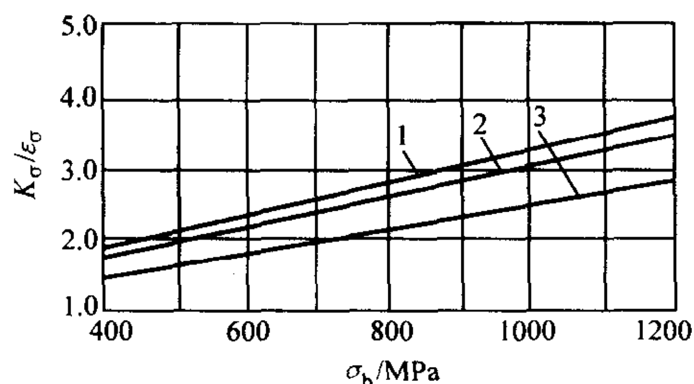
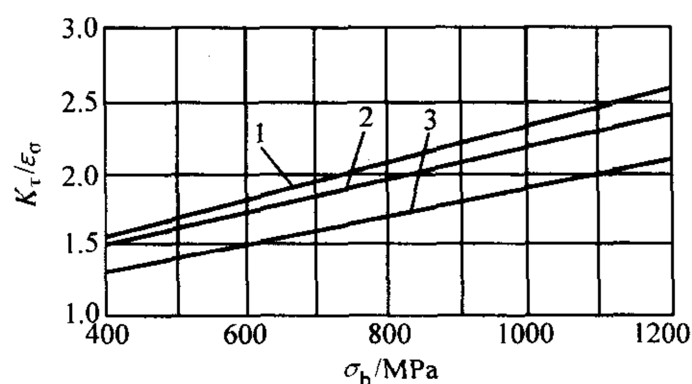


图 6.5-124 有螺纹、键槽、横孔的钢零件的弯曲（拉伸）的有效应力集中系数



(a) 弯曲和拉压



(b) 扭转

图 6.5-125 钢轴上配合件（间隙配合 $\frac{H7}{h6}$）的有效应力集中系数与尺寸系数的比值

1—$d\geqslant100$ mm；2—$d=50$ mm；3—$d\leqslant30$ mm

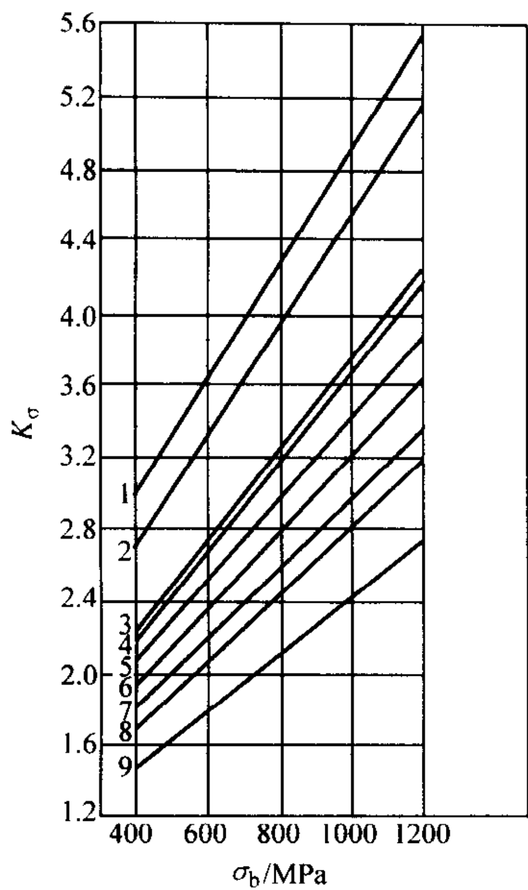


图 6.5-126 压力配合钢轴弯曲的有效应力集中系数

1—过盈配合$\frac{H7}{s6}$，$d>100$ mm；2—过盈配合$\frac{H7}{s6}$，$d=50$ mm；3—过盈配合$\frac{H7}{s6}$，$d=30$ mm；4—过盈配合$\frac{H7}{r5}$，$d>100$ mm；5—过盈配合$\frac{H7}{r5}$，$d=50$ mm；6—间隙配合$\frac{H7}{h6}$，$d>100$ mm；7—间隙配合$\frac{H7}{h6}$，$d=50$ mm；8—过盈配合$\frac{H7}{r5}$，$d=30$ mm；9—间隙配合$\frac{H7}{h6}$，$d=30$ mm

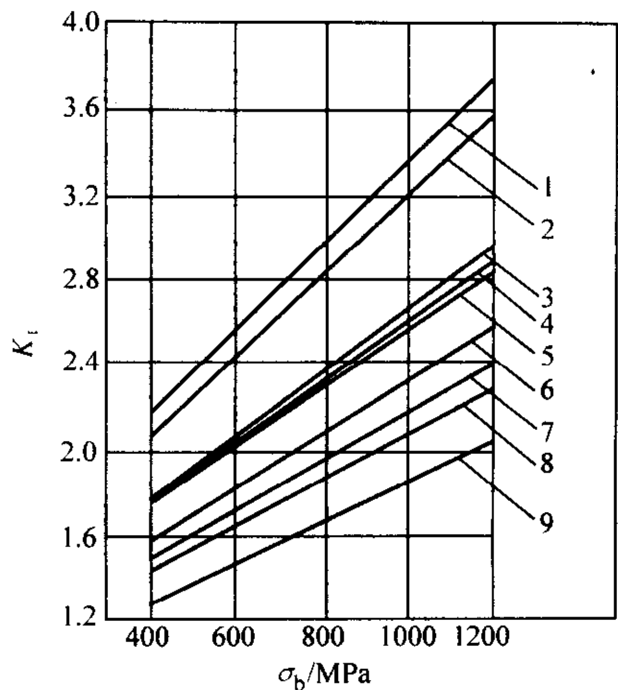
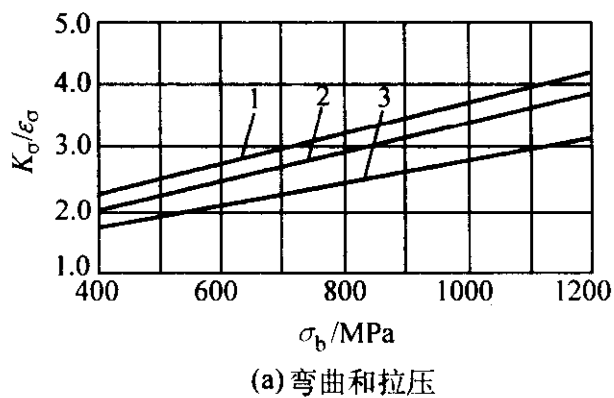
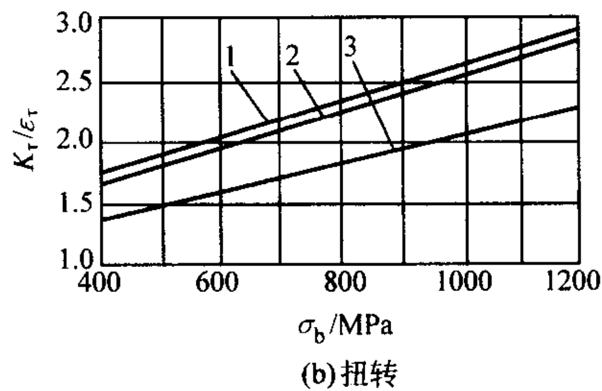


图 6.5-127 压力配合钢轴的扭转的有效应力集中系数

1—过盈配合$\frac{H7}{s6}$，$d>100$ mm；2—过盈配合$\frac{H7}{s6}$，$d=50$ mm；3—过盈配合$\frac{H7}{s6}$，$d=30$ mm；4—过盈配合$\frac{H7}{r5}$，$d>100$mm；5—过盈配合$\frac{H7}{r5}$，$d=50$ mm；6—间隙配合$\frac{H7}{h6}$，$d>100$ mm；7—间隙配合$\frac{H7}{h6}$，$d=50$ mm；8—过盈配合$\frac{H7}{r5}$，$d=30$ mm；9—间隙配合$\frac{H7}{h6}$，$d=30$ mm



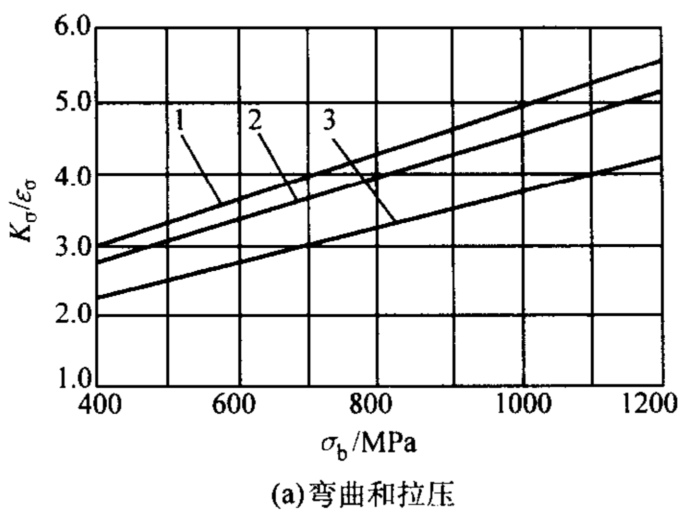
(a) 弯曲和拉压



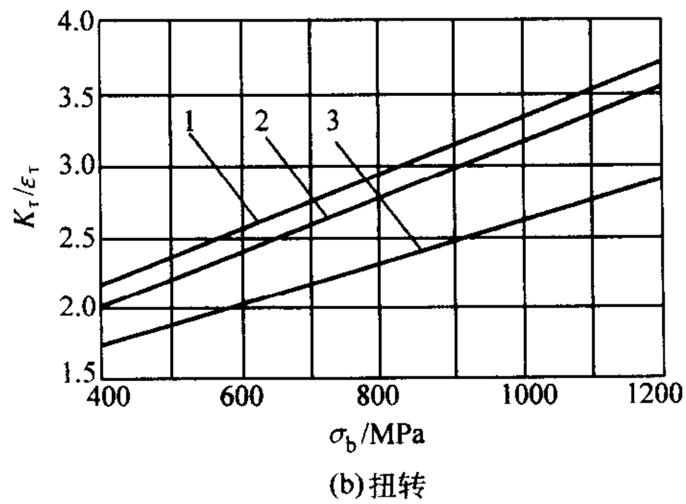
(b) 扭转

图 6.5-128 钢轴上配合件（过渡配合$\frac{H7}{k6}$）的有效应力集中系数与尺寸系数的比值

1—$d\geqslant100$ mm；2—$d=50$ mm；3—$d\leqslant30$ mm



(a) 弯曲和拉压



(b) 扭转

图 6.5-129 钢轴上配合件（过盈配合$\frac{H7}{s6}$）的有效应力集中系数与尺寸系数的比值

1—$d\geqslant100$ mm；2—$d=50$ mm；3—$d\leqslant30$ mm

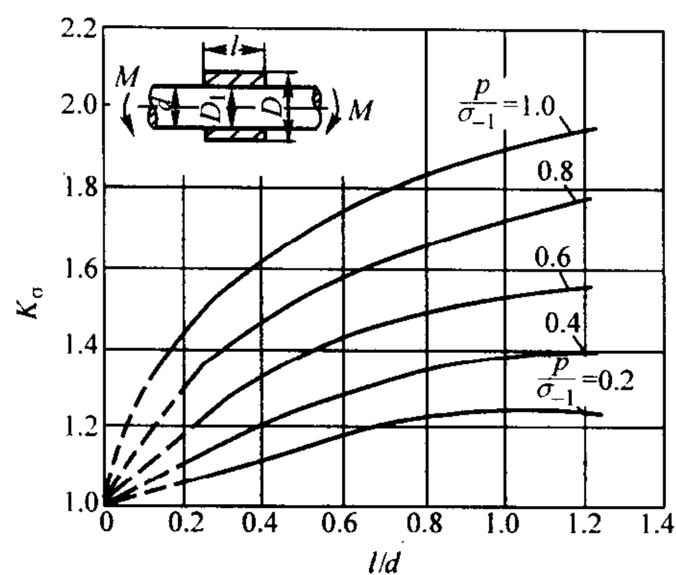


图 6.5-130 压入的过盈配合钢轴的弯曲的有效应力集中系数

$$p=\frac{E\ (d-D_1)\ (D^2-d^2)}{2dD^2}$$

p—径向压力，MPa；E—弹性模量，MPa；

D_1—轴套的内径，mm；D—轴套的外径，mm

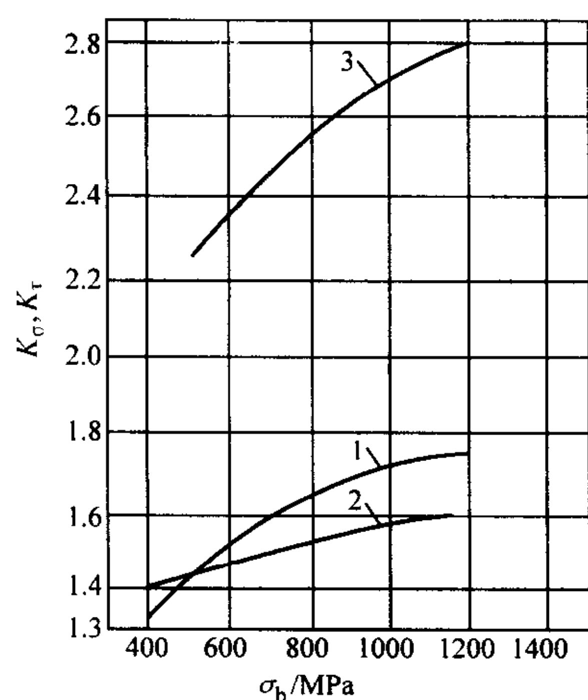


图 6.5-131 花键钢轴的有效应力集中系数

1—渐开线花键轴，弯曲；2—渐开线花键轴，扭转；

3—矩形花键轴，扭转

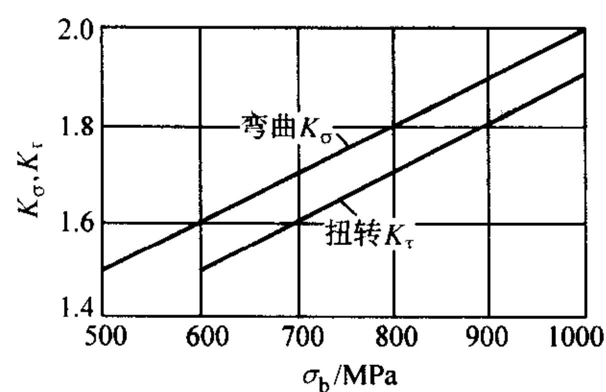


图 6.5-132 有单键或双键槽钢轴的有效应力集中系数

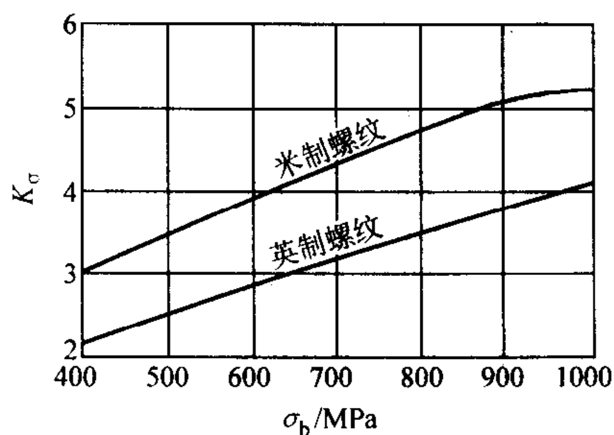
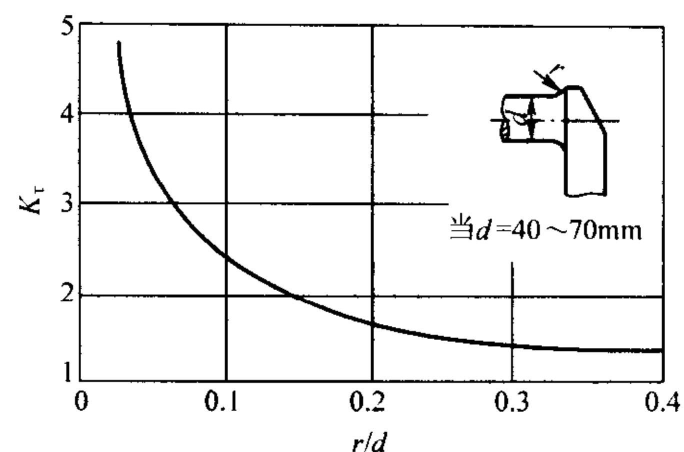
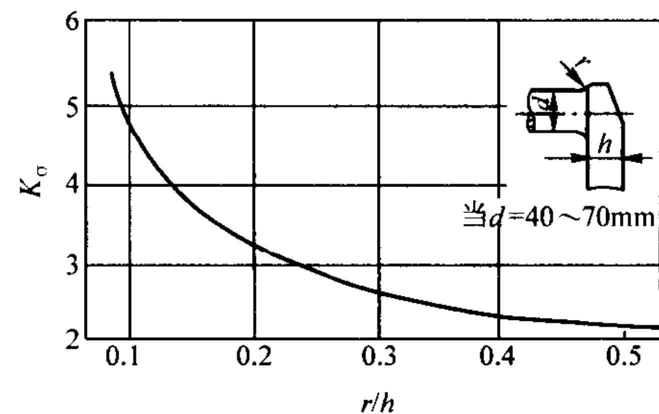


图 6.5-133 螺纹联接拉压的有效应力集中系数（钢件）



(a) 扭转



(b) 弯曲

图 6.5-134 钢曲轴的有效应力集中系数

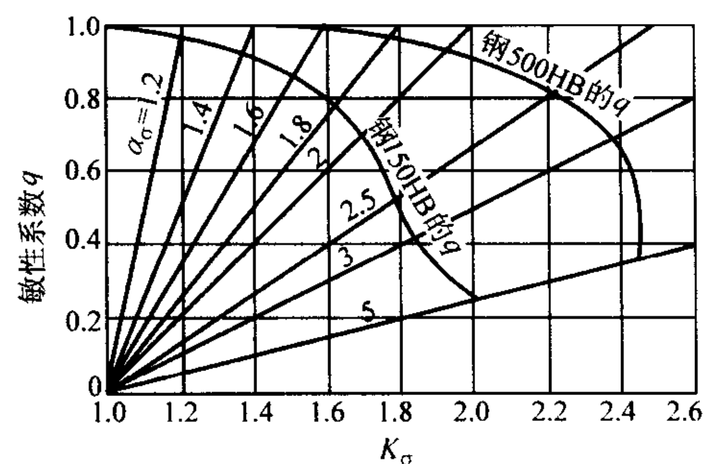


图 6.5-135 由理论应力集中系数及钢材的硬度确定敏性系数 q 或有效应力集中系数 K_σ

表 6.5-6 螺纹联接中的有效应力集中系数

钢号	光滑试样的疲劳极限 σ_{-1l}/MPa	螺纹的疲劳极限 σ'_{-1l}/MPa		有效应力集中系数 K_σ	
		切削螺纹	辊压螺纹	切削螺纹	辊压螺纹
35	176	49	63	2.7	2.1
45	215	58	78	2.8	2.1
38CrA	294	73	98	3.0	2.3
30CrMnSiA	294	73	98	3.0	2.3
40CrNiMoA	431	93	122	3.5	2.6
18Cr2Ni4VA	441	98	127	3.4	2.6

注：本表适用于 $d\leqslant 16$ mm 的米制螺纹，对于大尺寸的螺纹，应考虑尺寸系数。表中的疲劳极限是拉压疲劳试验得到的数值。

表 6.5-7　有键槽钢轴的有效应力集中系数

钢轴型式	钢种	力学性能		有效应力集中系数 K	
		σ_b/MPa	σ_{-1}/MPa	弯曲 K_σ	扭转 K_τ
3 个键槽 4.5 mm×10 mm d=30 mm	C C 3.5Ni	430 590 820	190 240 370	1.75 1.85 2.50	— — —
2 个键槽 4.5 mm×10 mm d=30 mm	C C 3.5Ni	430 590 820	190 240 370	— — —	2.40① 3.20① 4.35①
2 个键槽 5 mm×12 mm d=30 mm	C C C C 1 045 1.25Ni	430 560 650 880 562 725	190 240 — — 260 406	— — — — 1.32 1.61	1.55 1.75 1.85 2.25 — —

① 在装有配合件情况下试验。

3.2　尺寸的影响

在疲劳试验机上试验所用的试样直径通常为 6～10 mm，而一般零件的尺寸与试样有很大差别。尺寸因素对疲劳极限主要有如下影响：

1）尺寸增大时，材料的疲劳极限降低；

2）强度高的合金钢比强度低的合金钢尺寸影响大；

3）应力分布不均匀性增大时，尺寸影响大。

为在设计中计入这种影响，引入尺寸系数 ε。

尺寸系数的定义为当应力集中情况相同时，尺寸为 d 的零件的疲劳极限与标准试样的疲劳极限之比值，即

$$\left.\begin{aligned}&\text{弯曲时}\quad \varepsilon_\sigma=\frac{(\sigma_{-1})_d}{\sigma_{-1}}\\&\text{扭转时}\quad \varepsilon_\tau=\frac{(\tau_{-1})_d}{\tau_{-1}}\end{aligned}\right\}\tag{6.5-14}$$

式中，$(\sigma_{-1})_d$、$(\tau_{-1})_d$ 相应是尺寸为 d 的零件对称循环弯曲疲劳极限和对称循环扭转疲劳极限；σ_{-1}、τ_{-1} 相应是标准直径试样的对称循环弯曲疲劳极限和对称循环扭转疲劳极限。

尺寸系数 ε 的数据很分散，对于重型及一般机械设计，推荐图 6.5-136，这是锻钢的尺寸系数值；对于铸钢，应将图 6.5-136 的数据再降低 5%～10%；对于制造质量控制严的锻钢件，尺寸系数可适当提高。对于低合金结构钢，建议用碳素钢这条曲线。

表 6.5-8～表 6.5-10 分别给出钢试样的尺寸系数 ε 的统计参数，绝对尺寸影响系数 ε_σ、ε_τ 和光滑钢轴和阶梯钢轴对称循环下的弯曲疲劳试验结果。

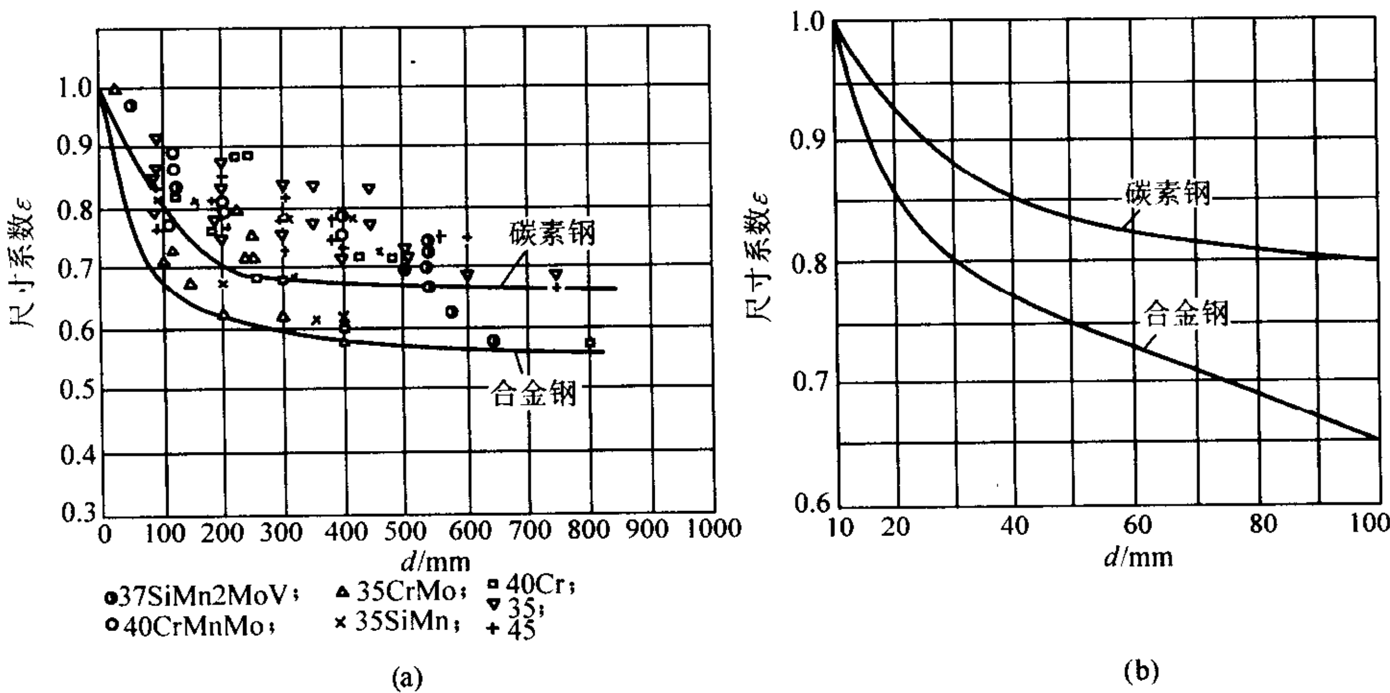


图 6.5-136　锻钢疲劳极限的尺寸系数 ε

表 6.5-8　钢试样的尺寸系数 ε 的统计参数

钢种	尺寸 d /mm	试样数 n	均值 $\bar{\varepsilon}$	标准差 S_ε	变异系数 $\nu=S_\varepsilon/\bar{\varepsilon}$	钢种	尺寸 d /mm	试样数 n	均值 $\bar{\varepsilon}$	标准差 S_ε	变异系数 $\nu=S_\varepsilon/\bar{\varepsilon}$
碳素钢	30～150	8	0.856 2	0.088 95	0.103 88	合金钢	30～150	11	0.79	0.069 0	0.087 34
	150～250	8	0.802 5	0.047 73	0.059 48		150～250	12	0.766 7	0.074 87	0.097 65
	250～350	9	0.791 1	0.034 44	0.043 53		250～350	5	0.678	0.068 34	0.100 79
	350 以上	14	0.73	0.041 88	0.057 37		350 以上	22	0.671 8	0.072 02	0.107 20

表 6.5-9　绝对尺寸影响系数 ε_σ、ε_τ

直径 d/mm		>20～30	>30～40	>40～50	>50～60	>60～70	>70～80	>80～100	>100～120	>120～150	>150～500
ε_σ	碳素钢	0.91	0.88	0.84	0.81	0.78	0.75	0.73	0.70	0.68	0.60
	合金钢	0.83	0.77	0.73	0.70	0.68	0.66	0.64	0.62	0.60	0.54
ε_τ	各种钢	0.89	0.81	0.78	0.76	0.74	0.73	0.72	0.70	0.68	0.60

表 6.5-10 光滑钢轴和阶梯钢轴对称循环下的弯曲疲劳试验结果

钢			d/mm	σ_{-1d}/MPa	σ_{-1kd}/MPa	α_σ	K_σ	q	ε_σ	加载条件
牌号	σ_b/MPa	$\sigma_{-1(10)}$/MPa								
(1) 碳素钢										
Q235A	402	185	190	125	—	—	—	—	0.68	平面弯曲
22 g	445	205	20	185	—	—	—	—	—	弯曲，试样静止
			200	165	—	—	—	—	—	
			150	137	—	—	—	—	0.67	
45	580	267	75	195	115	2.0	1.7	0.7	0.59	平面弯曲
45	584	269	42	245	120	2.4	—	—	0.91	弯曲，试样静止
			180	200	130	2.4	1.5	0.4	0.74	
40	711	327	135	200	106	2.2	1.9	0.7	0.61	平面弯曲
			135	—	87	3.4	2.3	0.5	—	
45	700	322	135	191	110	2.2	1.7	0.6	0.59	平面弯曲
			135	—	76	3.4	2.5	0.6	—	
ZG270－500	485	155	200	75	—	—	—	—	0.48	弯曲，试样静止
(2) 合金钢										
34CrNi3Mo	820	377	20	355	215	1.6	1.6	1.0	0.94	悬臂旋转弯曲
34CrNi3Mo	820	377	170	—	145	1.6	1.6	1.0	0.94	平面弯曲
	997	558	160	245	190	1.6	1.3	0.5	0.51	
	888	440	20	440	295	1.6	1.5	0.8	1.00	
15MnNi4Mo	888	440	170	255	185	1.6	1.4	0.7	0.63	平面弯曲
40Cr	910	311	65	345	235	1.8	1.5	0.6	0.86	
40CrNi	838	385	65	305	185	1.8	1.6	0.7	0.79	
40Cr	805	390	20	365	195	2.3	1.9	0.7	0.94	悬臂旋转弯曲
	805	390	160	330	175	2.4	1.9	0.6	0.85	弯曲，试样静止
40CrNi	821	390	20	390	195	2.3	2.0	0.8	1.00	悬臂旋转弯曲
	821	390	160	335	165	2.4	2.0	0.7	0.88	弯曲，试样静止
34CrNiMo	810	373	135	290	152	2.2	1.9	0.8	0.73	平面弯曲
	810	373	135	—	88	3.4	3.3	1.0	—	
34CrNiMo	850	391	160	300	—	—	—	—	0.77	平面弯曲
25CrMoV	912	420	20	410	175	2.6	2.3	0.8	0.97	悬臂旋转弯曲
	912	420	160	310	125	2.6	2.2	0.8	0.74	平面弯曲
25CrNi3-MoVA	817	376	280	—	77	3.1	—	—	—	平面弯曲
	823	379	18	305	—	—	—	—	0.81	悬臂旋转弯曲

3.3 表面状况影响

3.3.1 表面加工状况

疲劳试验的标准试样表面都经过磨光，而实际零件的表面加工方法则多种多样，表面加工粗糙相当于存在很多微缺口，在零件承受载荷时就产生应力集中。不管零件承受弯曲或扭转或二者联合作用的载荷，都是零件表面应力最大，所以疲劳源多从表面开始。因此表面质量不同，其疲劳强度也不同。粗糙表面导致疲劳强度降低。为了计入这一影响，在疲劳强度计算中引入了表面加工系数 β_1，其定义为

$$\beta_1 = \frac{(\sigma_{-1})_\beta}{\sigma_{-1}} \tag{6.5-15}$$

式中，$(\sigma_{-1})_\beta$ 为某种表面加工情况下试样的疲劳极限；σ_{-1} 为磨光试样的疲劳极限。

图 6.5-137 为钢试样弯曲或拉压循环载荷时的表面加工系数。对于扭转疲劳，在缺乏试验数据时，可取弯曲时的表面加工系数代之。

表 6.5-11 是表面加工系数的统计参数。

3.3.2　表面腐蚀状况

腐蚀环境对材料疲劳极限的影响，用腐蚀系数 β_2 表示，即

$$\beta_2 = \frac{(\sigma_{-1})_c}{\sigma_{-1}} \tag{6.5-16}$$

式中，$(\sigma_{-1})_c$ 为腐蚀环境中材料的疲劳极限；σ_{-1} 为空气中光滑试样的疲劳极限。

图 6.5-138 为腐蚀环境对钢试样的旋转弯曲疲劳极限影响的腐蚀系数。图 6.5-139 为铸铁在淡水中旋转弯曲条件下的腐蚀系数。表 6.5-12 为 1Cr13 钢在各种腐蚀环境中的腐蚀系数。

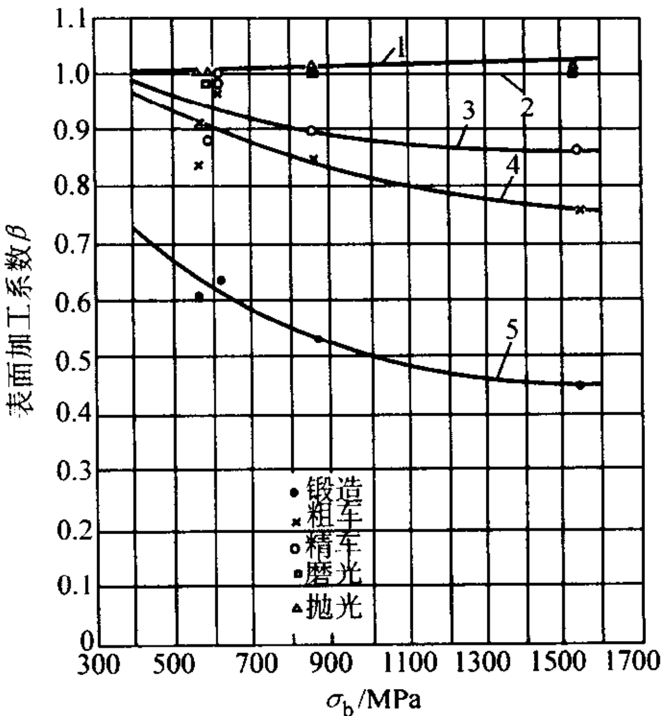


图 6.5-137　钢试样的表面加工系数 β

1—抛光；2—磨光；3—精车；4—粗车；5—锻造

表 6.5-11　表面加工系数的均值 $\bar\beta$ 及标准差 S_β

钢种	锻造		粗车 $R_a=12.5\,\mu m$		精车 $R_a=3.2\,\mu m$		磨削 $R_a=0.4\,\mu m$		抛光 $R_a=0.1\,\mu m$	
	$\bar\beta$	S_β	$\bar\beta$	S_β	$\bar\beta$	S_β	$\bar\beta$	S_β	$\bar\beta$	S_β
35	0.879 5	0.029 2	0.981 6	0.016 6	0.986 8	0.025 5	1	0.016 9	1.011 2	0.022 8
45	0.638 6	0.016 0	0.966 8	0.020 5	0.987 3	0.022 4	1	0.022 1	1.007 9	0.024 1
Q345	0.606 1	0.014 8	0.836 7	0.019 3	0.910 4	0.018 0	1	0.019 0	1.000 7	0.020 2
40Cr	0.535 3	0.020 9	0.847 9	0.050 5	0.901 1	0.035 5	1	0.049 9	1.021 0	0.040 1
60Si2Mn	0.456 0	0.017 3	0.762 2	0.030 7	0.866 1	0.034 6	1	0.035 3	1.014 3	0.033 1

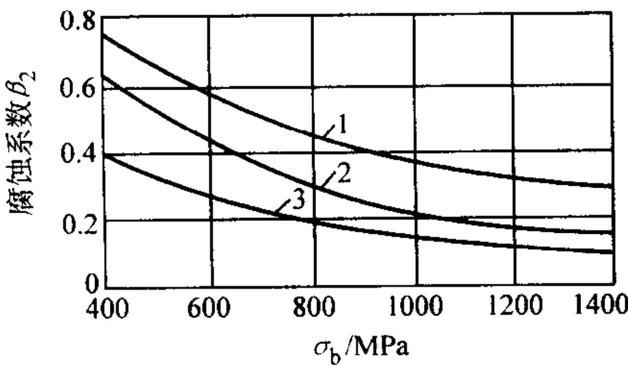


图 6.5-138　腐蚀环境对钢试样的旋转弯曲疲劳极限影响的腐蚀系数

1—淡水中无应力集中；

2—淡水中有应力集中，海水中无应力集中；

3—海水中有应力集中

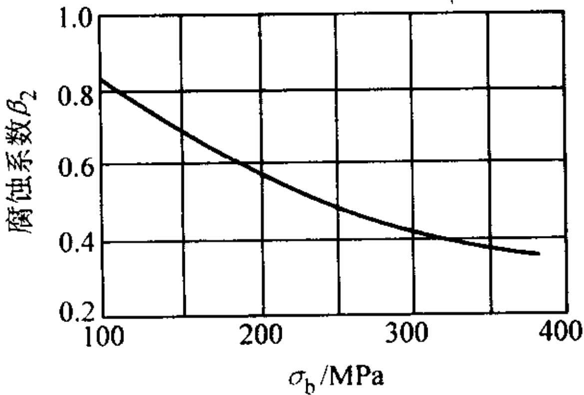


图 6.5-139　铸铁在淡水中旋转弯曲条件下的腐蚀系数

表 6.5-12　1Cr13 钢在各种腐蚀环境中的腐蚀系数

试验条件	试验温度/℃	腐蚀系数 β_2
在蒸汽气氛中	—	0.54
在蒸汽和空气的密封容器中	75	0.84
在 0.1 MPa 的蒸汽中	100	0.89
在压力为 4.4 MPa 的蒸汽中	150	0.90
在压力为 11.3 MPa 的蒸汽中	180	0.89
在压力为 16.2 MPa 的蒸汽中	370	0.89
在空气和湿蒸汽混合气体中	20	0.56

3.3.3　表面强化状况

由于机械零件的疲劳裂纹常开始于表层，所以强化表层是提高零件疲劳强度的有效方法。表面强化工艺可分为三类：①机械方法，如喷丸及辊压等；②化学方法，如渗碳及氮化等；③热处理，如高频、中频及工频电表层淬火，火焰淬火等。由此引入了表面强化系数 β_3，即

$$\beta_3 = \frac{(\sigma_{-1})_j}{\sigma_{-1}} \tag{6.5-17}$$

式中，$(\sigma_{-1})_j$ 为经强化工艺试样的疲劳极限；σ_{-1} 为未经强化工艺试样的疲劳极限。

各种强化工艺的强化系数 β_3，见表 6.5-13。

上述的表面加工系数 β_1，表面腐蚀系数 β_2 和表面强化系数 β_3，总称为表面系数。在疲劳强度计算中，应根据具

体情况选取相应的 β 值。例如，零件如只经过切削加工，则 $\beta=\beta_1$；如零件只经过强化，则 $\beta=\beta_3$；如零件在腐蚀介质中工作，则 $\beta=\beta_2$，不必将各 β 值相乘。

表6.5-14～表6.5-21给出感应加热淬火、渗氮、渗碳、辊压等强化处理后的疲劳试验结果。

表 6.5-13 表面强化系数 β_3 荐用值

强化方法	心部强度 σ_b/MPa	钢试样的表面强化系数 β_3		
		光滑试样	有应力集中的试样	
			$K_\sigma\leqslant1.5$	$K_\sigma\geqslant2.0$
高频淬火	600～800	1.3～1.5	1.4～1.5	1.8～2.2
	800～1 000	1.2～1.4	1.5～2.0	—
氮化	900～1 200	1.1～1.3	1.5～1.7	1.7～2.1
渗碳	400～600	1.8～2.0	3	—
	700～800	1.4～1.5	—	—
	1 000～1 200	1.2～1.3	2	—
辊压	600～1 500	1.1～1.4	1.4～1.6	1.6～2.0
喷丸	600～1 500	1.1～1.4	1.4～1.6	1.6～2.0
镀铬	—	0.5～0.7		
镀镍	—	0.5～0.9		
镀锌(热浸法)	—	0.6～0.95（电镀法取 $\beta_3=1.0$）		
镀铜	—	0.9		

表 6.5-14 感应淬火对圆柱钢试样对称弯曲疲劳极限的影响

表面硬度 HV	表面层厚度 t/mm	相对厚度 t/R	疲劳极限	
			MPa	%
250	—	—	380	100
650～800	1.5	0.187 5	590	155
650～800	1.9	0.237 5	598	157
650～850	3.3	0.412 5	664	175

注：材料为 $w_C=0.46$（%）碳钢，$d=2R=16$ mm，$\sigma_b=771$ MPa。

表 6.5-15 感应淬火对 $w_C=0.4\%$ 碳钢光滑和缺口试样旋转弯曲疲劳极限的影响（硬化层厚度1.2 mm）

处理型式	试样型式	疲劳极限 MPa	疲劳极限 %
正火	光滑	245	100
表面硬化	光滑	425	173
正火	缺口，0.4 mm深[①]	148	100
表面硬化	缺口，0.4 mm深[①]	422	282
正火	缺口，0.8 mm深[①]	143	100
表面硬化	缺口，0.8 mm深，硬化前加工	375	262
表面硬化	缺口，0.8 mm深，硬化后加工	382	269
正火	缺口，1.2 mm深	133	100
表面硬化	缺口，1.2 mm深，硬化前加工	285	214
表面硬化	缺口，1.2 mm深，硬化后加工	302	227
正火	孔，$d=3.6$ mm	145	100
表面硬化	孔，$d=3.6$ mm	245	169
正火	带压配轴套	142	100
表面硬化	带压配轴套	365	259

① 半径0.3 mm的U形缺口。

表 6.5-16 渗氮和渗碳的强化系数 β_3

表面处理	厚度 t/mm	硬度 HV	试样型式	试样直径/mm	β_3
渗氮	0.1～0.4	700～1 000	光滑	8～15	1.15～1.25
	0.1～0.4	700～1 000	光滑	30～40	1.10～1.15
	0.1～0.4	700～1 000	缺口	8～15	1.90～3.00
	0.1～0.4	700～1 000	缺口	30～40	1.30～2.00
渗碳	0.2～0.8	670～750	光滑	8～15	1.2～2.1
	0.2～0.8	670～750	光滑	30～40	1.1～1.5
	0.2～0.8	670～750	缺口	8～15	1.5～2.5
	0.2～0.8	670～750	缺口	30～40	1.2～2.5

表 6.5-17 氮化与未氮化的疲劳极限 MPa

材料	未氮化		氮化	
	光滑试样	缺口试样	光滑试样	缺口试样
普通铸铁	215	156	264	313
球墨铸铁	245	171	269	342
2Cr13	—	225	—	402

表 6.5-18 渗碳钢试样的旋转弯曲疲劳极限 MPa

材料	光滑试样		缺口试样[①]
	处理前	处理后	处理后
$w_C=0.20\%$钢	195	415	260
$w_C=0.35\%$钢	220	470	370
Cr-Ni钢	300	660	370

① 缺口半径 $R=0.75$ mm。

注：试样直径 $d=10$ mm，渗碳深度1.0～1.2 mm，渗碳温度1 050℃。

表 6.5-19　辊压对不同尺寸钢试样旋转弯曲疲劳极限的影响

钢材硬度 HV	硬化层相对厚度	直径 /mm	疲劳极限		硬度增量 HV
			MPa	%	
12ChN3A 315~325 ГОСТ	—	6.5	500	100	—
	—	35	430	100	—
	0.057	6.5	550	110	20
	0.057	35	480	112	65
	0.114	6.5	580	116	25
	0.114	35	500	116	95
	0.170	6.5	610	122	55
	0.170	35	530	123	95
37ChN3A 360~365 ГОСТ	—	6.5	640	100	—
	—	35	560	100	—
	0.057	6.5	700	109	20
	0.057	35	600	107	10
	0.114	6.5	720	112	50
	0.114	35	580	104	45
	0.170	6.5	720	112	35
	0.170	35	650	116	70
18ChNWA 325~370 ГОСТ	—	6.5	570	100	—
	—	35	480	100	—
	0.057	6.5	680	119	80
	0.057	35	550	114	105
	0.110	6.5	640	112	60
	0.170	6.5	670	117	75

表 6.5-20　42CrMo 钢辊压前后的疲劳极限

静强度/MPa		疲劳极限（$N=10^6$）/MPa	
σ_s	σ_b	未辊压	辊压
853	963	618	689
880	1 044	591	698
982	1 145	532	731

表 6.5-21　各种组织的铸铁的辊压效果

组织状态	辊压力/N	弯曲疲劳极限/MPa		提高率/%
		辊压前	辊压后	
珠光体+片状石墨	353	115	139	20
铁素体+片状石墨	490	61	131	114
珠光体+球状石墨	1 804	193	468	142
铁素体+球状石墨	1 451	123	360	193

3.4　载荷状况

载荷状况影响包括载荷类型、载荷频率、载荷变化情况及平均应力等。

3.4.1　载荷类型影响

机械零件承受载荷类型有拉、压、弯、扭及以上 4 种的组合作用。疲劳数据中多用旋转弯曲疲劳试验获得。在缺少其他加载方式的试验数据时，用载荷系数 C_L 来修正。一般取拉、压的载荷系数 $C_L=0.85$，扭转的载荷系数 $C_L=0.58$。

对于重要的零构件，应该用相同载荷类型下试验得到的数据来进行计算或设计。

3.4.2　载荷频率影响

对于高周疲劳，在空气中，室温下进行试验，频率对疲劳极限影响很小。在腐蚀环境或高温条件下试验时，频率对疲劳极限影响很大。图 6.5-140 是几种材料的频率和疲劳极限曲线。由图可见，当频率小于 1 000 Hz 时，疲劳极限随着频率的增加稍有增加，其后出现最大值。当频率再增加时，疲劳极限下降。

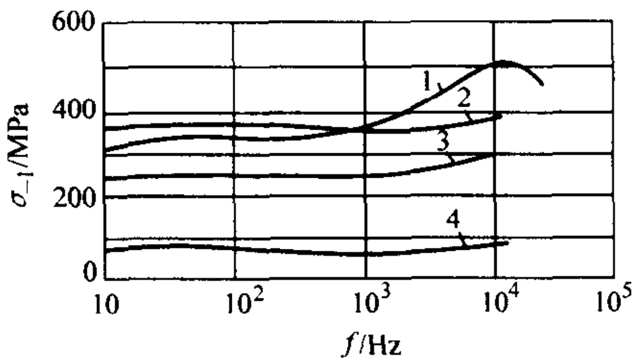


图 6.5-140　载荷频率对金属疲劳极限的影响
1—$w_C=0.86\%$碳素钢；2—$w_C=0.11\%$碳素钢；3—铜；4—铝

因此，在室温下工作的机械，一般不考虑频率的影响。在腐蚀及高温环境下工作的机械则必须考虑频率的影响。

图6.5-141 是 20Cr 钢的试验频率和腐蚀系数的关系曲线。

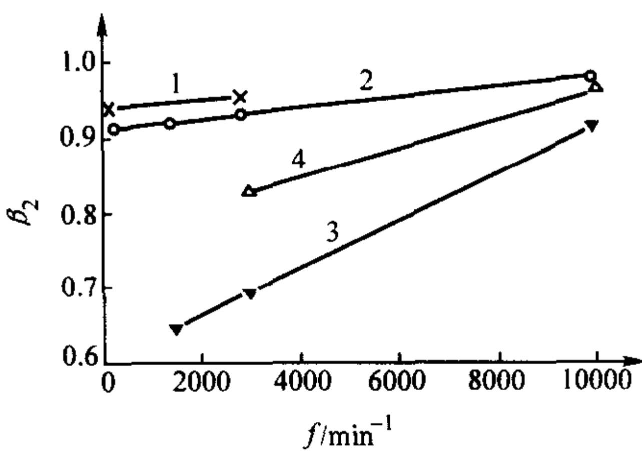


图 6.5-141　20Cr 钢试验频率对疲劳强度影响
1—在航空油中，试样磨光；
2—在航空油加 2% 油酸中，试样磨光；
3—在淡水加 2% 异戊醇中，试样磨光；
4—在淡水加 2% 异戊醇中，试样车削

图 6.5-142 是铸钢 ZG20SiMn 和 ZG0Cr13Ni4Mo 在淡水介质中腐蚀疲劳极限与试验频率的关系曲线。

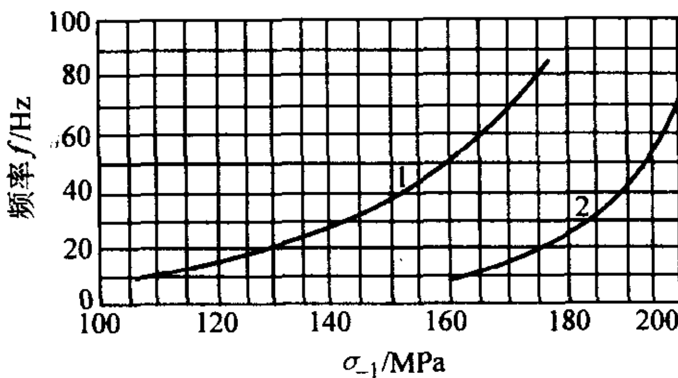


图 6.5-142　试验频率对铸钢在淡水介质的腐蚀疲劳强度的影响
1—铸造结构钢 ZG20SiMn；2—铸造不锈钢 ZG0Cr13Ni4Mo

3.4.3　载荷峰值影响

在随机载荷作用下的工作应力，有时偶然连续超过材料的疲劳极限，这种情况称为过载。少量次数的过载应力，可使材料应变强化和产生残余压应力，使疲劳极限提高，但循环次数超过一临界值后，因造成损伤而使疲劳极限明显降低。低于该临界值，则不造成疲劳损伤。

在低于材料疲劳极限某些应力水平运转一定循环次数后，可使疲劳极限明显提高，这种现象称为低载锻炼。低载锻炼的效果，决定于材料的力学性能、锻炼的应力水平和循环次数。应力越接近疲劳极限，锻炼的效果越明显。

在随机载荷中，每一应力循环的应力峰值和谷值是随机变化的，这时不仅存在过载和低载应力大小的影响，而且受到加载顺序的影响。

3.4.4　平均应力影响

不同的平均应力可用应力比反映出来。表6.5-22和表6.5-23是国产钢不同应力比下的拉－压和扭转的疲劳极限。

表6.5-22　7种国产钢不同应力比下的拉－压疲劳极限　　MPa

材料	α_σ	应力比 $r=-1$		应力比 $r=0$		应力比 $r=0.3$		应力比 $r=1$	
		均值	标准差	均值	标准差	均值	标准差	均值	标准差
Q345（热轧）	1	269	9.4	377	23.1	431	17.5	533	6.7
	2	169	5.7	327	7.6	421	11.4	734	15.3
	3	109	3.2	218	8.5	257	12.2	875	7.2
35（正火）	1	177	9.4	291	11.2	388	7.5	606	10.0
	2	131	6.6	243	10.6	313	16.3	730	7.8
	3	96	4.8	192	5.9	252	12.7	839	15.5
45（调质）	1	269	8.6	436	13.4	517	22.5	762	36.7
	2	173	7.1	334	12.3	418	19.7	922	32.8
	3	103	4.4	187	8.5	277	13.9	1 178	43.7
45（正火）	1	219	8.9	346	9.2	346	23.3	577	24.8
	2	165	5.7	313	12.2	399	18.6	782	14.8
	3	121	4.1	208	8.2	274	5.0	871	10.3
40Cr（调质）	1	345	17.3	629	44.7	671	25.3	855	21.4
	2	257	8.5	431	18.0	555	21.2	1 209	34.6
	3	163	1.6	257	6.0	337	8.3	1 358	38.3
40CrNiMo（调质）	1	499	4.5	805	18.7	856	31.0	1 001	74.6
	2	276	4.8	490	20.7	599	14.6	1 139	26.4
	3	188	5.9	322	14.3	439	17.2	1 383	18.9
60Si2Mn（淬火后中温回火）	1	487	26.3	749	33.8	1 118	29.0	1 442	31.4
	2	338	14.8	527	21.0	701	24.3	1 777	71.5
	3	215	10.4	356	20.7	468	33.0	2 041	70.5

表6.5-23　两种国产钢不同应力比下的扭转疲劳极限　　MPa

材料	α_σ	应力比 $r=-1$		应力比 $r=0$		应力比 $r=0.3$		应力比 $r=1$	
		均值	标准差	均值	标准差	均值	标准差	均值	标准差
45（正火）	1	233	5.6	450	18.7	—	—	317	7.1
	2	101	5.9	189	12.1	264	5.3	603	18.5
	3	119	5.1	177	6.9	239	6.7	556	10.7
40Cr（调质）	1	314	15.5	574	9.5	—	—	609	48.0
	2	141	3.4	235	6.9	319	10.0	794	42.0
	3	145	4.6	199	8.0	243	7.3	782	30.9

有平均应力时的载荷称不对称载荷，其相应的应力称不对称循环应力。在进行强度计算时，常将不对称循环应力折算成等效的对称循环应力。等效应力幅 $\sigma_A = \sigma_a + \psi_\sigma \sigma_m$，$\psi_\sigma$ 称不对称循环度系数或平均应力影响系数。

表 6.5-24 是 7 种国产材料的平均应力影响系数。图 6.5-143 和图 6.5-144 是国产 45 钢和 40Cr 在 3 种应力集中系数下的疲劳极限线图（或称等寿命曲线图）。图 6.5-145 是不同应力集中系数对平均应力影响系数的影响曲线。应用图 6.5-145 曲线可查出不同应力集中下的 ψ_σ 值。在缺少数据情况下，用光滑试样的 ψ_σ 值来代替有应力集中条件下的 ψ_σ，对于设计来说是偏于安全的。

其他加载情况和表面状态条件对 ψ_σ 值也有影响，见表 6.5-25 和表 6.5-26。

表 6.5-24 7 种国产钢的平均应力影响系数

材 料	热处理	α_σ	平均应力影响系数 ψ_σ		
			$r=0$	$r=0.3$	$r=1$
Q345	热轧	1	0.43	0.42	0.50
		2	0.04	0.08	0.23
		3	0.003	0.12	0.12
35	正火	1	0.22	0.17	0.29
		2	0.08	0.11	0.18
		3	0.014	0.048	0.11
45	正火	1	0.26	0.43	0.38
		2	0.06	0.10	0.21
		3	0.17	0.14	0.14
45	调质	1	0.23	0.26	0.35
		2	0.034	0.10	0.19
		3	0.10	0.034	0.09
40Cr	调质	1	0.10	0.25	0.40
		2	0.20	0.17	0.21
		3	0.27	0.21	0.12
40CrNiMo	调质	1	0.24	0.36	0.50
		2	0.12	0.17	0.24
		3	0.16	0.12	0.14
60Si2Mn	淬火后中温回火	1	0.23	0.13	0.34
		2	0.28	0.20	0.19
		3	0.21	0.17	0.11

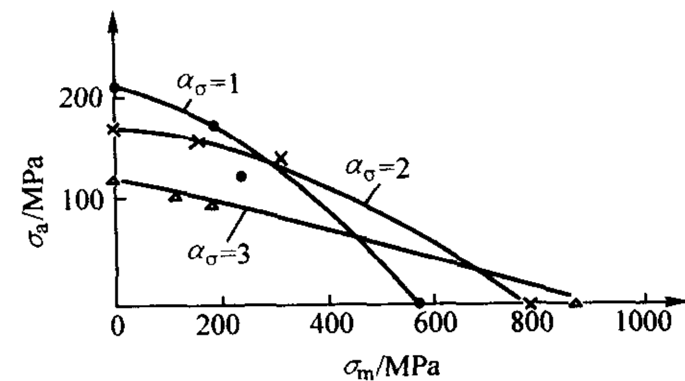


图 6.5-143 45 钢在不同应力集中系数下的疲劳极限线图
($N=10^7$)
45 钢经正火，其 $\sigma_b = 612$ MPa，$\sigma_s = 361$ MPa

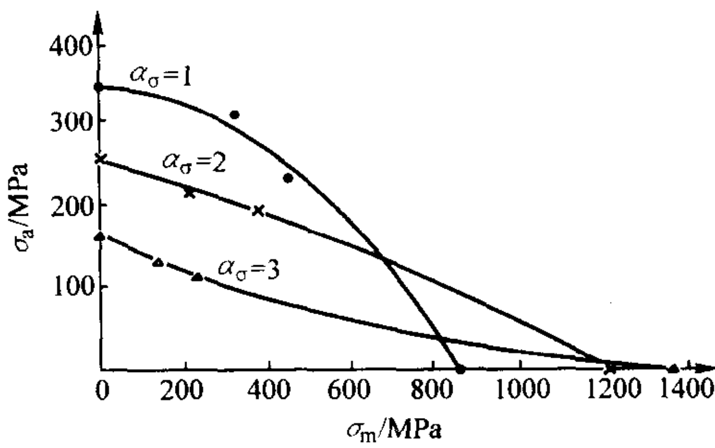


图 6.5-144 40Cr 在不同应力集中系数下的疲劳极限线图
($N=10^7$)
40Cr 经调质，其 $\sigma_b = 858$ MPa，$\sigma_s = 673$ MPa

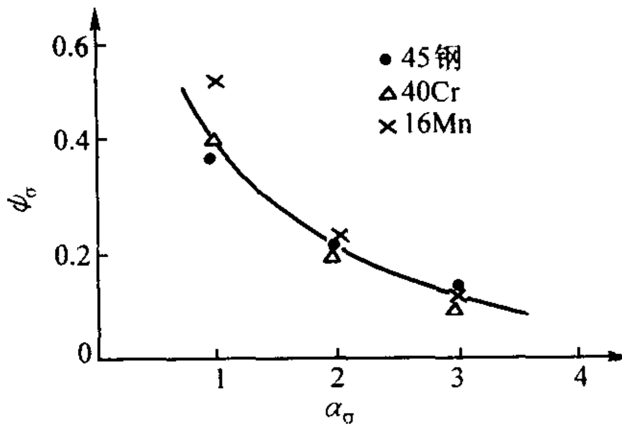


图 6.5-145 应力集中系数与平均应力影响系数 ψ_σ 的关系曲线

表 6.5-25 钢的系数 ψ_σ 和 ψ_τ

应力种类	系数	表面状态				
		抛光	磨削	车削	热轧	锻造
弯曲	ψ_σ	0.50	0.43	0.34	0.215	0.14
拉压	ψ_σ	0.41	0.36	0.30	0.18	0.10
扭转	ψ_τ	0.33	0.29	0.21	0.11	0.05

表 6.5-26 铸铁和铝合金的系数 ψ_σ 和 ψ_τ

材 料	ψ_σ		ψ_τ
	弯曲	拉压	扭转
铸 铁	0.49	0.41	0.48
铝合金	0.335	0.335	0.335

4 常规疲劳强度设计

试样和零件在高于 $10^4 \sim 10^5$ 次以上载荷循环而产生的疲劳，称为高周疲劳。对于高周疲劳通常采用常规疲劳强度设计方法。

常规疲劳强度设计是以名义应力为基本设计参数的抗疲劳设计方法，也称名义应力法。

它是假设零构件没有初始裂纹，应用标准试样试验得到的疲劳极限、$S-N$ 曲线及疲劳极限图等，再考虑零构件由于尺寸、表面状态及几何形状引起的应力集中等影响因素而进行的疲劳强度设计。把 $S-N$ 曲线用双对数坐标表示时，是由两根直线组成的折线。按水平线部分进行设计称无限寿命设计；按斜线部分进行设计称有限寿命设计。

无限寿命设计要求零构件在无限长的使用期间内不发生疲劳破坏。将零构件的工作应力限制在它的疲劳极限以下，得到的零构件的寿命在理论上是无限的。用这种准则进行设计常造成零构件结构尺寸大，过于笨重。但对于长时间运转的零构件，仍是一个较好的设计准则。

有限寿命设计，也称安全寿命设计。它保证机器在一定

使用期限内安全运行，所以它允许零构件的工作应力超过其疲劳极限。其基本依据是材料或零构件的 $S-N$ 曲线的斜线部分。计算的重点是零构件的裂纹形成寿命。这种设计准则能充分利用材料的承载能力，减小零构件的截面尺寸，减小质量。对于如飞机、汽车等要求减小质量、更新速度快的产品有重要意义。

对于有限寿命设计来说，疲劳损伤累积理论是其重要依据。而对于无限寿命设计则主要是计算其安全系数。

4.1　安全系数

一般的疲劳强度计算中，许用安全系数推荐用表6.5-27的数值。表6.5-28为初算时的安全系数荐用值。表6.5-29为各类机械零部件的许用安全系数。

表中所用符号：$n_{bp}=\sigma_b/\sigma_p$；$n_{sp}=\sigma_s/\sigma_p$；$n_{-1p}=\sigma_{-1}/\sigma_p$；$n_{0p}=\sigma_0/\sigma_p$（其中，$\sigma_b$ 为材料的抗拉强度；σ_s 为材料的屈服强度；σ_{-1} 为对称循环疲劳极限；σ_0 为脉动循环疲劳极限；下角p为“许用”）。校核零件的疲劳强度，必须使它同时满足静强度要求。

表6.5-27　许用安全系数

情　况	n_{-1p}
材料较均匀，载荷及应力较精确时	1.3
材料不够均匀，载荷及应力计算精度较差	1.5~1.8
材料均匀度很差，计算精度很差	1.8~2.5

表6.5-28　安全系数荐用值（初算用）

材　料		静载荷		冲击载荷		疲劳载荷			
						n_{bp}		n_{-1p}	
		n_{bp}	n_{sp}	n_{bp}	n_{sp}	一般零件	重要零件	一般零件	重要零件①
铸铁		3~4	—	10~15	—	8~10	12~15	—	—
高强度钢		2~3	—	—	—	—	—	—	—
结构钢	$\sigma_s/\sigma_b=0.45\sim0.6$，计算精确	2.4~2.6	1.2~1.5	2.0~2.8	1.5~2.2	5.0	7	1.3	1.5
	$\sigma_s/\sigma_b=0.6\sim0.8$，计算精度一般		1.4~1.8	2.5~4.0	2.0~2.8	5.5	8	1.5	1.8
	$\sigma_s/\sigma_b=0.8\sim0.9$，计算不精确		1.7~2.2	3.5~5.0	2.5~3.5	6.0	10	1.8	2.5

① 重要零件是指在整个使用期内不希望破坏的零件。

表6.5-29　各类机械零件的许用安全系数

机械种类	零部件名称	应力状态	材　料	安全系数	备　注
起重机械	主梁	弯	Q235A，Q345	$n_{sp}=1.4\sim1.6$ $n_{0p}=1.4\sim1.6$	运送液态金属的起重机用1.6
	端梁	弯	Q235A，Q345	$n_{sp}=2.4$	
	小车梁	弯	Q235A，Q345	$n_{sp}=3\sim4$	
	卷筒轴	弯曲疲劳	45	$n_{sp}=1.3\sim1.6$， $n_{-1p}=1.8$	手动，$n_{sp}=1.3$，吊钢水包 $n_{sp}=1.6$
	减速机低速轴	弯扭疲劳	45	$n_{sp}=1.6$， $n_{-1p}=1.8$	
	卷筒轴承侧法兰螺栓	拉伸疲劳	Q235A	$n_{sp}=3$，$n_{0p}=2.5$	
	吊钩钩体	拉、弯	20，36Mn2Si	$n_{sp}=1.6$	
	吊钩螺纹尾部	拉	20，36Mn2Si	$n_{sp}=5\sim7$	
	吊钩梁	弯	45	$n_{sp}=3$	
	拉板	拉、挤压	Q345	$n_{sp}=1.6$	
	吊钩滑轮轴	弯	45	$n_{sp}=1.6$	
	小车轮轴	弯扭疲劳	45	$n_{sp}=1.4$， $n_{-1p}=1.6$	
	大车轮轴	弯扭疲劳	45	$n_{sp}=1.4$， $n_{-1p}=1.6$	
矿山机械	矿井提升机卷筒	弯，压	Q235A，Q345	$n_{sp}=1.4\sim1.6$	
	矿井提升机主轴	弯扭疲劳	45	$n_{-1p}=1.2\sim1.5$	
	颚式破碎机机架	弯曲疲劳	ZG270-500	$n_{0p}=1.5$	
	颚式破碎机传动轴	弯扭疲劳	45	$n_{-1p}=1.5$	
	颚式破碎机主轴	弯扭疲劳	45	$n_{-1p}=1.4$	
	圆锥破碎机传动轴	弯扭疲劳	45	$n_{-1p}=1.4$	
	圆锥破碎机主轴	弯扭疲劳	24CrMoV	$n_{-1p}=2$	
	圆锥破碎机液压缸体	内压	ZG270-500	$n_{sp}=2\sim2.4$	
	球磨机筒体	弯	Q235A，20	$n_{sp}=3.5\sim4$	

续表 6.5-29

机械种类	零部件名称	应力状态	材料	安全系数	备注
冶金机械	轧钢机机架（初轧机）	弯、拉、拉伸疲劳	ZG270 - 500	$n_{bp}=6\sim8$	$n_{0p}=1.6$
	轧钢机机架（板热轧机）	弯、拉、拉伸疲劳	ZG270 - 500	$n_{bp}=7\sim10$	$n_{0p}=1.7$（厚板）
	轧钢机机架（板冷轧机）	弯、拉	ZG270 - 500	$n_{bp}=8\sim12$	考虑刚度
	轧钢机轧辊（初轧机辊身）	弯扭疲劳	60CrMnMo，60CrMo，55CrMo	$n_{bp}=6\sim8$	$n_{-1p}=1.8$
	轧钢机轧辊（热轧板工作辊）	弯扭疲劳	HT250，球铁	$n_{bp}=6.5$	$n_{-1p}=1.5\sim2.5$
	冷轧薄板工作辊	弯扭疲劳	9Cr2	$n_{-1p}=1.1$	
	热轧板支承辊	弯曲疲劳	37SiMn2MoV，8CrMoV，40Mn2MoV	$n_{bp}=6$	$n_{-1p}=1.2\sim2$
	冷轧板支承辊	弯曲疲劳	9Cr2，9Cr2Mo	$n_{-1p}=1.2$	
	轧钢机的机架辊	弯扭疲劳	45	$n_{bp}=6$	$n_{-1p}=1.8$
	轧钢机万向接轴	弯扭疲劳	45CrV	$n_{sp}=3$	$n_{-1p}=2.0$
	轧钢机万向接轴叉头	弯扭疲劳	45CrV	$n_{sp}=2.6$	$n_{-1p}=1.8$
	六连杆式热剪机的上剪股	弯曲疲劳	ZG35CrMo，32SiMn2MoV	$n_{sp}=2$	$n_{0p}=1.5$
	六连杆式热剪机的下剪股	弯曲疲劳	ZG35CrMo，32SiMn2MoV	$n_{sp}=3$	$n_{0p}=1.6$
	六连杆式热剪机的偏心轴	弯扭疲劳	40	$n_{sp}=3$	$n_{-1p}=2.0$
	六连杆式热剪机的连杆	拉压弯	40	$n_{sp}=3$	
	六连杆式热剪机的传动轴	弯扭疲劳	35CrMo，35SiMn2MoV	$n_{sp}=3$	$n_{-1p}=2.5$
	摆式飞剪机曲轴	弯扭疲劳	35SiMn2MoV	$n_{-1p}=2$	
	辊式矫直机的工作辊辊身	弯扭疲劳	9Cr2，60CrMoV	$n_{sp}=4\sim12$	考虑刚度
	辊式矫直机的支承辊辊身	弯扭疲劳	9Cr2	$n_{sp}=3\sim6$	考虑刚度
	辊式矫直机的支承辊辊颈	扭	9Cr2	$n_{sp}=1.7$	
	辊式矫直机的支承辊小轴	弯	42MnMoV	$n_{sp}=2$	
	辊式矫直机的机架（铸铁）	弯	HT250	$n_{bp}=6$	
	辊式矫直机的机架（钢）	弯	Q235A	$n_{sp}=3$	
	辊式矫直机的机架盖	弯	Q235A	$n_{bp}=5$	
	辊式矫直机万向接轴	弯扭疲劳	35SiMn	$n_{sp}=4\sim5$	$n_{-1p}=1.6$
	辊式矫直机压下螺杆	扭、压	45，35SiMn	$n_{sp}=2.7$	
	辊式矫直机拉杆	拉	35SiMn	$n_{sp}=3$	
	高炉大钟拉杆	拉	20	$n_{sp}=5$	考虑温度
	转炉托圈	弯	—	$n_{sp}=8$	考虑温度
	转炉耳轴	弯	40Cr，38SiMnV	$n_{sp}=3$	
	盛钢桶桶体	内压	Q235A	$n_{sp}=2.5$	$n_{-1p}=2$
	盛钢桶耳轴	弯	ZG270 - 500	$n_{sp}=7$	
	铁水车减速机轴	弯扭疲劳	—	$n_{-1p}=2.3$	
锻压机械	水压机立柱（光滑部分）	拉、弯	40，45，20MnV，20SiMnMo	$n_{sp}=1.7\sim2$	
	水压机立柱（螺纹部分）	拉、弯	40，45，20MnV，20SiMnMo	$n_{sp}=4\sim5$	$n_{-1p}=1.5$
	水压机上横梁	弯	ZG270 - 500，Q235A	$n_{bp}=6\sim8$	$n_{-1p}=1.4\sim1.6$
	水压机活动横梁	弯	Q235A	$n_{bp}=5\sim6$	
	水压机下横梁	弯	Q235A	$n_{bp}=8\sim12$	

续表 6.5-29

机械种类	零部件名称	应力状态	材料	安全系数	备注
锻压机械	水压机液压缸缸体	内压	35，45，20MnV，Q345（12MnV）	$n_{sp}=2\sim3$	
	水压机液压缸法兰	弯、压	35，45，ZG35，22MnMo	$n_{sp}=2.2$	有冲击时 $n_{sp}=3\sim4$
	水压机液压缸柱塞	内压	45	$n_{sp}=2.2$	
	水压机高压水罐	内压	20，14CrMnMoV	$n_{sp}=2$	
	水压机充水罐	内压	Q235A	$n_{sp}=3$	
	挤压机柱子（光滑部分）	拉、弯	18MnMoNb	$n_{sp}=2$	
	挤压机柱子（螺纹部分）	拉、弯	18MnMoNb	$n_{sp}=4$	$n_{-1p}=1.9$
	挤压机机架	弯	Q235A，ZG270－500	$n_{sp}=3\sim6$	
	挤压机主缸缸体	内压	18MnMoNb	$n_{sp}=2.5\sim3$	$n_{bp}=4\sim5$
	挤压机动梁回程缸缸体	内压	18MnMoNb	$n_{sp}=3\sim4.5$	
	挤压机穿孔缸缸体	内压	35SiMn，18MnMoNb	$n_{sp}=3\sim4.5$	
	挤压机穿孔回程缸缸体	内压	45	$n_{sp}=2.5$	
	挤压机剪刀缸缸体	内压	45	$n_{sp}=2.5$	
	挤压机移动缸缸体	内压	45	$n_{sp}=2.0$	
	精压机传动轴	弯扭疲劳	35SiMn2MoV	$n_{-1p}=2$	
	锻锤机架	弯、拉伸疲劳	ZG270－500	$n_{bp}=5$	$n_{0p}=1.6$
	锻锤拉杆	拉	40Cr，35CrMnV	$n_{sp}=2.5$	
	热模锻曲轴	弯扭疲劳	40CrNi，35SiMn2MoV	$n_{-1p}=1.6\sim2$	
橡胶塑料机械	橡胶塑料辊机辊筒	弯扭疲劳	HT200	$n_{-1p}=2.5\sim3$	冷硬铸铁
	橡胶塑料辊机机架	弯	HT250，HT300	$n_{bp}=12$	$n_{0p}=5$
	橡胶塑料辊机机架盖	弯	HT250	$n_{bp}=10$	$n_{0p}=4.5$
	橡胶塑料挤出机螺杆	扭	38CrMoAl	$n_{sp}=3$	
	橡胶塑料挤出机机筒		38CrMoAl	$n_{sp}=3$	考虑结构要求
内燃机	内燃机曲轴主轴颈	扭转疲劳	QT600－3，45，40MnB，40Cr，40，45Mn2，30MnMoW，30Mn2MoTiB，40Mn2SiV，15SiMn3MoWVA，37SiMnMoWV	$n^{\tau}_{-1p}=3\sim4$（扭应力）	汽车发动机
	内燃机曲轴主轴颈	扭转疲劳		$n^{\tau}_{-1p}=4\sim5$（扭应力）	拖拉机发动机
	内燃机曲轴主轴颈	扭转疲劳		$n^{\tau}_{-1p}=2\sim3$（扭应力）	高增压柴油机
	内燃机曲轴曲柄销	弯扭疲劳		$n_{-1p}=1.3\sim1.5$	汽车发动机
	内燃机曲轴曲柄销	弯扭疲劳		$n_{-1p}=1.5\sim2$	拖拉机发动机
	内燃机曲轴曲柄销	弯扭疲劳		$n_{-1p}=1.2\sim1.4$	高增压柴油机
	内燃机曲轴曲柄臂	弯扭疲劳		$n_{-1p}=2\sim3$	汽车发动机
	内燃机曲轴曲柄臂	弯扭疲劳		$n_{-1p}=3\sim3.5$	拖拉机发动机
	内燃机曲轴曲柄臂	弯扭疲劳		$n_{-1p}=1.3\sim2$	高增压柴油机
内燃机	内燃机活塞销	弯、剪	20，20Cr，20Mn2，18CrMnTi，20SiMnVB	$n_{sp}=2\sim2.2$	渗碳
	内燃机连杆小头	弯压疲劳	45，40Cr，35CrMo，40MnVB	$n_{-1p}=2.5\sim5$	汽车发动机
	内燃机连杆杆身	弯压疲劳		$n_{-1p}=2\sim2.5$	拖拉机发动机
	内燃机连杆杆身	弯压疲劳		$n_{-1p}=2.5\sim3$	船用中、高速柴油机
	内燃机连杆杆身	弯压疲劳		$n_{-1p}=2\sim3$	高速强载柴油机
	内燃机连杆大头	弯压疲劳		$n_{-1p}=2.0$	汽车、拖拉机发动机
	内燃机连杆大头	弯压疲劳		$n_{-1p}=1.5$	高速强载柴油机
	内燃机连杆螺栓	拉伸疲劳	45，40Cr，35CrMo，40MnVB	$n_{-1p}=1.5\sim2$	
	汽缸体紧螺栓	拉伸疲劳	40Cr，40MnB，35CrMo，40CrMo	$n_{-1p}=1.3\sim2$	

续表 6.5-29

机械种类	零部件名称	应力状态	材　料	安全系数	备　注
气体压缩机	气体压缩机曲轴	弯扭疲劳	45	$n_{-1p}=2\sim2.5$	$n_{sp}=3\sim6$
	气体压缩机曲柄臂	弯扭疲劳	45	$n_{-1p}=1.5$	
	气体压缩机连杆	弯扭疲劳	30	$n_{sp}=3$	
	气体压缩机活塞杆	弯扭疲劳	45	$n_{bp}=10$	
	气体压缩机高压缸阀腔	内压	40	$n_{-1p}=1.4\sim2$	
汽车拖拉机	汽车变速箱轴	弯扭疲劳	40Cr，40MnB，18CrMnTi	$n_{-1p}=1.3$	曲轴、连杆的安全系数见内燃机
	汽车后桥半轴	弯扭疲劳	40MnB，35CrMnSiA	$n_{-1p}=2$	
	拖拉机变速箱轴	弯扭疲劳	40，18CrMnTi	$n_{-1p}=2$	
	拖拉机传动轴	弯扭疲劳	40	$n_{-1p}=1.3$	
	拖拉机履带驱动轮轴	弯扭疲劳	40Cr	$n_{-1p}=1.1$	
水轮机	水轮机转轮叶片	拉、弯	ZG20SiMn，ZG0Cr13Ni4Mo	$n_{sp}=2.5$ $n_{-1p}=2$	
	水轮机主轴轴身	拉、弯、扭	45，20SiMn	$n_{sp}=2.5\sim3$	
	水轮机主轴法兰	弯、压	45，20SiMn	$n_{sp}=1.8\sim2.3$	
	水轮机导叶体	弯、扭	ZG270－500，ZG20SiMn	$n_{sp}=2$	
	水轮机导叶体轴颈	弯、扭	ZG270－500，ZG20SiMn	$n_{sp}=1.8$	
	水轮机导叶臂	弯、扭	ZG270－500	$n_{sp}=1.8$	
	水轮机导叶套筒	弯	HT200	$n_{bp}=10$	
	水轮机接力器缸体	内压	HT200	$n_{bp}=10$	
	水轮机接力器油缸法兰	弯、压	HT200	$n_{bp}=5$	
水轮机	水轮机涡壳	内压	Q235A，Q345（16Mn），Q390（15MnV，15MnTi）	$n_{sp}=1.8\sim2$	混流式水轮机 n_{-1p} 的数值随使用年限而定，对于使用年限较短时，可取 $n_{-1p}=1.5\sim1.8$
	水轮机顶盖和支持盖	弯	HT200，HT300	$n_{bp}=8.5\sim10$	
	水轮机顶盖和支持盖	弯	ZG270－500	$n_{sp}=2$	
	水轮机导水机构盖板	弯、拉	Q235A	$n_{sp}=2$	
	水轮机联结板	弯、拉	Q235A	$n_{sp}=2$	
	水轮机旋管，导管体	拉	Q235A	$n_{sp}=2.5$	
	水轮机耳柄	拉	35，40Cr	$n_{sp}=2.5$	
	水轮机转臂	弯、挤压	35	$n_{sp}=2.5$	
	水轮机连杆	拉、压	ZG270－500	$n_{sp}=2$	
	水轮机活塞销，连杆销	弯	35	$n_{sp}=2$	
	水轮机叶销	剪切	45	$n_{sp}=2$	
	水轮机联轴螺栓	弯、拉	35，40Cr	$n_{sp}=2.5$	
	水轮机叶片螺栓	弯、拉	35，40Cr	$n_{sp}=2$	
	水轮机分半键，导向键	剪切	Q235A，35	$n_{sp}=2$	
	水轮机叶片键，卡环	剪切	45	$n_{sp}=2$	

4.2　疲劳损伤累积

当材料或零件承受高于疲劳极限的应力时，每一循环都使材料产生一定量损伤，这种损伤能够累积，当损伤累积到临界值时将发生破坏，这就是疲劳损伤累积理论。

所谓损伤就是材料或零构件中细微“结构”的变化，在循环应力作用下，形成微裂纹，并成长和合并，导致材料的变质和恶化。损伤累积的结果，往往产生宏观裂纹，导致最终破断。

疲劳累积损伤理论，归纳起来可分三大类。

1）线性疲劳累积损伤理论，材料在各个应力下的疲劳损伤是独立进行的，并且总损伤可以线性地累加起来。其中最有代表性的是帕姆格伦－迈因纳（Palmgren－Miner）理论。简称迈因纳理论。

2）非线性疲劳累积损伤理论　基于假定载荷历程和损伤之间存在着相互的作用，即在各个应力下产生的损伤与前面应力作用的量值和次数有关。这一理论的代表是科尔顿和多兰（Corten&Dolan）理论。

3）其他的累积损伤理论　大都是从实验、观测和分析推导出来的损伤公式，多属于经验和半经验公式。如莱维（Levy）、科津（Kozin）等理论即是。

线性累积损伤理论，特别是迈因纳理论，形式简单、使用方便，在工程中得到了广泛的应用。

4.2.1　线性疲劳累积损伤理论

线性累积损伤理论认为，材料在各个应力下的疲劳损伤是独立进行的，并且总损伤可以线性地累加起来。

图6.5-146为疲劳损伤线性累积示意图。图6.5-146a为变化的应力，图6.5-146b为 $S-N$ 曲线。

应力 σ_1 作用 n_1 次，在该应力水平下材料达到破坏的总循环次数为 N_1。设 D 为最终断裂时的损伤临界值，根据线性疲劳累积损伤理论，应力 σ_1 每作用一次对材料的损伤为 D/N_1，经 n_1 次循环作用后，σ_1 对材料的总损伤为 n_1D/N_1。同样可找出仅有 σ_2 作用时，材料发生破坏的应力循环数 N_2，应力 σ_2 每循环一次对材料的损伤为 D/N_2，经 n_2 次循环后，σ_2 对材料的总损伤应力 n_2D/N_2。如此类推，应力 σ_3，循环作用 n_3 次，对材料造成的总损伤为 n_3D/N_3。应力 σ_4 小于材料疲劳极限 σ_{-1}，它可以作用无限次循环而不引起材料疲劳损伤，计算中可以不予考虑。

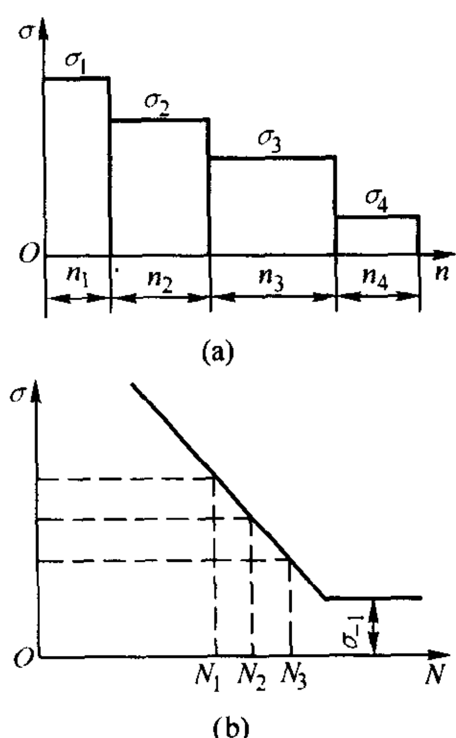


图6.5-146　疲劳损伤线性累积示意图

当各级应力对材料的损伤总和达到临界值 D 时，材料即发生破坏。用公式表示为

$$\frac{n_1D}{N_1}+\frac{n_2D}{N_2}+\frac{n_3D}{N_3}=D$$

或写成

$$\frac{n_1}{N_1}+\frac{n_2}{N_2}+\frac{n_3}{N_3}=1$$

上面的关系式推广到更普遍的情况时，即有

$$\frac{n_1}{N_1}+\frac{n_2}{N_2}+\cdots+\frac{n_n}{N_n}=1$$

或写成

$$\sum_{i=1}^{m}\frac{n_i}{N_i}=1 \tag{6.5-18}$$

式（6.5-18）称为线性疲劳累积损伤方程式。迈因纳理论与试验结果并不完全相符合。这是因为疲劳损伤的累积不但决定于当前的应力状况，而且还和过去作用的应力历史有关。即材料以前作用的应力历史，对以后应力作用下损伤有干涉效应。另外，加载顺序对损伤有明显影响。先作用高应力还是先作用低应力，所得结果不一样。因而使得式（6.5-18）的右边不等于1，而等于某一数值 a。数值 a 是在某一区间内变化。

a 值取0.7时，其寿命估算结果比迈因纳公式计算更安全，从总体上看其寿命估算精度也有提高。

4.2.2　相对迈因纳（Miner）法则

由于上述的迈因纳法则没有考虑载荷次序和残余应力的复杂非线性相互影响。因而分散性很大。其 a 值在0.3~10之间变化。而相对迈因纳法则一方面保留了迈因纳法则中第一个假设，即线性累积假设，另一方面又避开了累积损伤 $a=1$的第二个假设。

相对迈因纳法则的数学表达式为

$$N_A=N_B\frac{\left(\sum\frac{n_i}{N_i}\right)_B}{\left(\sum\frac{n_i}{N_i}\right)_A} \tag{6.5-19}$$

式中，N_A 为载荷谱A作用下估算的疲劳裂纹形成寿命；N_B 为载荷谱B作用下估算的疲劳裂纹形成寿命；$\left(\sum\frac{n_i}{N_i}\right)_A$ 为载荷谱A的计算累积损伤；$\left(\sum\frac{n_i}{N_i}\right)_B$ 为载荷谱B的计算累积损伤。

式（6.5-19）表明，只要两个谱的载荷历程相似，则两个谱的寿命之比等于它们的累积损伤之比的倒数。

使用相对迈因纳法则的关键是确定相似谱B。其中有两点假设：①相似谱B的主要峰谷顺序应和计算谱A相近或相同（保证相似谱能模拟计算谱的载荷次序特征）；②相似谱B的主要峰谷大小和计算谱A成比例或近似成比例。比例因子最好接近1，以便保证相似谱能够模拟计算谱在缺口根部造成的塑性变形。

用相对迈因纳法则计算和试验结果比较可见，能大幅度消除迈因纳法则计算数值引起的误差。提高其计算精度。

4.3　无限寿命设计

4.3.1　单向应力时无限寿命设计

零部件受单向循环应力，是指只承受单向正应力或单向切应力。例如，只承受单向拉压循环应力、弯曲循环应力或扭转循环应力。在单向循环应力下工作的零部件很多，如高炉上料机的钢丝绳受单向波动拉伸应力，曲柄压力机的连杆受单向脉动应力。只承受弯曲力矩的心轴，转动时表面上各点的应力状态是对称循环弯曲应力等。

（1）计算公式

表6.5-30中列出了不同受载情况下单向应力时安全系数的计算公式。公式的符号涵义为：n_σ、n_τ 为计算的安全系数；σ_{-1}、τ_{-1} 为材料在对称循环下的疲劳极限，弯曲时为 σ_{-1}，拉压时为 σ_{-1l}，扭转时为 τ_{-1}；K_σ、K_τ 为弯曲和扭转时的有效应力集中系数；ε 为尺寸系数；β 为表面系数；ψ_σ、ψ_τ 为不对称循环系数，一般计算式为 $\psi_\sigma=(2\sigma_{-1}-\sigma_0)/\sigma_0$，$\psi_\tau=(2\tau_{-1}-\tau_0)/\tau_0$；$\sigma_0$、$\tau_0$ 为弯曲和扭转时的脉动循环疲劳极限；σ_i、τ_i 为作用试样上的第 i 个应力水平；n_i 为第 i 个应力水平 σ_i 或 τ_i 作用时的循环数；σ_{max}、τ_{max}为载荷谱中的最大应力；N_0 为无限寿命的最小循环数，即循环基数；N 为总寿命，即整个工作循环数；m 为材料常数，即 $S-N$ 曲线在对数坐标中的倾斜率的负值，$m=-\lg N_i/\lg\sigma_i$；N_i 为在应力水平 σ_i 作用下，材料达到疲劳破坏的循环数；σ_{di}、τ_{di}为第 i 个当量应力，计算式为

$$\sigma_{di}=\left[\frac{K_\sigma}{\varepsilon\beta}(\sigma_a)_d\right]_i$$

$$\tau_{di}=\left[\frac{K_\tau}{\varepsilon\beta}(\tau_a)_d\right]_i$$

$$(\sigma_a)_d=\sigma_a+\psi_\sigma\sigma_m$$

$$(\tau_a)_d=\tau_a+\psi_\tau\tau_m$$

表 6.5-30 单向应力时安全系数计算式

受载情况	弯曲或拉压时的安全系数	扭转时的安全系数
恒幅对称循环	$n_\sigma=\dfrac{\sigma_{-1}}{\dfrac{K_\sigma}{\varepsilon\beta}\sigma_a}$	$n_\tau=\dfrac{\tau_{-1}}{\dfrac{K_\tau}{\varepsilon\beta}\tau_a}$
恒幅不对称循环	$n_\sigma=\dfrac{\sigma_{-1}}{\dfrac{K_\sigma}{\varepsilon\beta}\sigma_a+\psi_\sigma\sigma_m}$	$n_\tau=\dfrac{\tau_{-1}}{\dfrac{K_\tau}{\varepsilon\beta}\tau_a+\psi_\tau\tau_m}$
变幅对称循环	$n_\sigma=\dfrac{\sigma_{-1}}{\dfrac{K_\sigma}{\varepsilon\beta}\sqrt[m]{\dfrac{N}{N_0}\sum\limits_i\left(\dfrac{\sigma_i}{\sigma_{max}}\right)^m\dfrac{n_i}{N}}\sigma_{max}}$	$n_\tau=\dfrac{\tau_{-1}}{\dfrac{K_\tau}{\varepsilon\beta}\sqrt[m]{\dfrac{N}{N_0}\sum\limits_i\left(\dfrac{\tau_i}{\tau_{max}}\right)^m\dfrac{n_i}{N}}\tau_{max}}$
变幅不对称循环	$n_\sigma=\dfrac{\sigma_{-1}}{\sqrt[m]{\dfrac{N}{N_0}\sum\limits_i\left(\dfrac{\sigma_{di}}{\sigma_{dmax}}\right)^m\dfrac{n_i}{N}}\sigma_{dmax}}$	$n_\tau=\dfrac{\tau_{-1}}{\sqrt[m]{\dfrac{N}{N_0}\sum\limits_i\left(\dfrac{\tau_{di}}{\tau_{dmax}}\right)^m\dfrac{n_i}{N}}\tau_{dmax}}$

（2）算例

图 6.5-147 所示的轴。载荷 F 为对称循环载荷，$F=50\,000$ MPa，轴材料为 45 钢，调质。表面加工方法为精车，校核 $A-A$ 截面的疲劳强度。

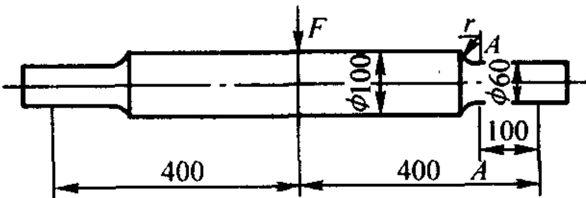


图 6.5-147 轴

解 1）计算公式 因载荷是等幅对称循环，故用公式

$$n_\sigma=\frac{\sigma_{-1}}{\dfrac{K_\sigma}{\varepsilon\beta}\sigma_a}$$

2）求 σ_a 该轴为简支梁，故 $A-A$ 截面的应力为

$$\sigma_a=\frac{M}{W}=\frac{16Fl}{\pi d^3}=\frac{16\times50\,000\times100}{\pi\times60^3}=117.9\ \text{MPa}$$

3）求 σ_{-1} 查表 6.5-2，材料 45 钢调质状态时，$\sigma_{-1}=388$ MPa，$\sigma_b=710$ MPa

4）求 K_σ 用式（6.5-13），$K_\sigma=1+q(\alpha_\sigma-1)$

查图 6.5-109，当 $\sigma_b=710$ MPa，$r=5$ mm 时，$q=0.8$。

查图 6.5-76，当 $\rho/d=5/60=0.083$，$D/d=100/60=1.67$ 时，$\alpha_\sigma=1.78$，得 $K_\sigma=1.624$。

或从图 6.5-113 直接查得 $K_\sigma\approx1.70$，两者差异不大，取 $K_\sigma=1.7$。

5）求 ε 查图 6.5-136，当 $d=60$ mm，45 钢时 $\varepsilon=0.825$。

6）求 β 表面加工方法为精车，故 $\beta=\beta_1$。

查图 6.5-137，当 $\sigma_b=710$ MPa 及精车时 $\beta_1=0.92$。

7）求 n_σ

$$n_\sigma=\frac{\sigma_{-1}}{\dfrac{K_\sigma}{\varepsilon\beta}\sigma_a}=\frac{388}{\dfrac{1.7}{0.825\times0.92}\times117.9}=1.47>n_p=1.3$$

故该轴 $A-A$ 截面的疲劳强度符合要求。

4.3.2 多向应力时无限寿命设计

在多向应力情况下，把多向应力转化成单向应力，然后利用上述的单向应力设计方法进行设计。变形能强度理论及最大切应力理论是将多向应力状态与单向应力状态联系起来，比较符合实际的理论。这里根据变形能强度理论，把多向应力转化成单向当量应力，其计算公式为

当量应力幅

$$\sigma_{da}=\frac{[(\sigma_{a1}-\sigma_{a2})^2+(\sigma_{a2}-\sigma_{a3})^2+(\sigma_{a3}-\sigma_{a1})^2]^{1/2}}{\sqrt2}\tag{6.5-20}$$

当量平均应力

$$\sigma_{dm}=\frac{[(\sigma_{m1}-\sigma_{m2})^2+(\sigma_{m2}-\sigma_{m3})^2+(\sigma_{m3}-\sigma_{m1})^2]^{1/2}}{\sqrt2}\tag{6.5-21}$$

式中，σ_{a1}、σ_{a2}、σ_{a3} 为主应力幅；σ_{m1}、σ_{m2}、σ_{m3} 为主应力幅方向的平均应力。对于二向应力状态，公式可简化为

$$\left.\begin{aligned}\sigma_{da}&=(\sigma_{a1}^2-\sigma_{a1}\sigma_{a2}+\sigma_{a2}^2)^{1/2}\\ \sigma_{dm}&=(\sigma_{m1}^2-\sigma_{m1}\sigma_{m2}+\sigma_{m2}^2)^{1/2}\end{aligned}\right\}\tag{6.5-22}$$

有了这两个当量应力后，可以运用单向应力计算公式进行设计。

在二向应力状态时，最常见的承受弯曲和扭转复合循环应力作用的传动轴和曲轴等的设计中，常采用下面公式计算其安全系数，即

$$n=\frac{1}{\sqrt{\left(\dfrac{1}{n_\sigma}\right)^2+\left(\dfrac{1}{n_\tau}\right)^2}}\tag{6.5-23}$$

这里的 n_σ 和 n_τ，就是上述的单向弯曲和单向扭转状态下的安全系数（参见表 6.5-30）。

4.4 有限寿命设计

4.4.1 安全系数计算公式

在有限寿命设计中，多向应力状态的处理方法与无限寿命设计的方法是一样的，将它转化为单向当量应力。

安全系数计算公式与无限寿命设计中的公式一样，只是其中有些系数取值不一样。推荐的系数取值列于表 6.5-31 中。

表 6.5-31　系数取值

系　数	取　值
有效应力集中系数 $K_{\sigma x}$	$N \leqslant 10^3$　$K_{\sigma x}=1.0$ $10^3 < N < 10^6$ $K_{\sigma x}=1.0+(K_\sigma-1.0)(x-3)/3$ $N \geqslant 10^6$　$K_{\sigma x}=K_\sigma$ x 为循环数的对数，即 $x=\lg N_x$ K_σ 为无限寿命时的有效应力集中系数
尺寸系数 ε 表面加工系数 β 不对称循环系数 ψ_σ、ψ_τ	与无限寿命设计中相同

4.4.2　寿命估算

在进行有限寿命设计时，不但要计算零构件的工作安全系数，还要计算零构件的疲劳寿命。常用的疲劳寿命计算公式列于表 6.5-32 中。

表 6.5-32　寿命估算方法

应力状态	方　法	内　容
恒幅	简单估算法	根据计算确定的零件危险点处应力幅 σ_a，在零件的 $S-N$ 曲线上确定对应的循环数，就是所要求的寿命
变幅	线性累积损伤理论的方法	$N=\dfrac{1}{\sum_i \dfrac{1}{N_i}\times\dfrac{n_i}{N}}$ n_i/N 可从载荷谱中求得，N_i 是对应于 σ_i 的循环次数，可以从 $S-N$ 曲线求得

4.4.3　随机疲劳寿命估算

承受随机载荷的零构件，在进行疲劳强度和疲劳寿命计算之前，必须先确定其载荷谱。由于载荷谱编制时既可以用程序谱给出，也可用概率密度函数给出，故其寿命计算公式不相同。

1）程序谱的疲劳寿命计算　设程序谱的一个周期内含有 k 级应力水平 σ_1、σ_2、…、σ_k，各级应力水平下的循环数分别为 n_1、n_2、…、n_k。由考虑各种影响因素的 $S-N$ 曲线，查得各级应力水平单独作用时的破坏循环数 N_1、N_2、…、N_k。用 T 表示周期总数，当零件破坏时（出现宏观裂纹），则零件寿命为

$$T=\frac{1}{\sum_{i=1}^{k}\frac{n_i}{N_i}} \tag{6.5-24}$$

2）概率密度函数给出的连续谱寿命计算　设载荷幅值变化的概率密度函数为 $P(\sigma)$，T 为零件的寿命，则

$$T=\frac{1}{\int_{\sigma_r}^{\sigma_{max}}\frac{P(\sigma)}{C}\sigma^m \mathrm{d}\sigma} \tag{6.5-25}$$

式中，σ_r 为零件材料的疲劳极限；σ_{max} 为载荷时间历程中最大的应力幅；$P(\sigma)$ 为应力幅值频率密度函数；m、C 为材料有关的系数。m 可由试样的疲劳试验确定。一般情况下，$m=3\sim10$。当 m 已知时，由 $\sigma_i^m N_i=C$ 即可求得 C 值，则由式（6.5-25）可算得寿命 T。若式（6.5-25）积分困难时，可用数值计算方法解决。

4.4.4　算例

计算一起重机吊钩上端螺纹的疲劳寿命。已知螺纹为 M64 的标准螺纹，螺纹材料是 20 钢锻造，其力学性能为：$\sigma_b=412$ MPa，$\sigma_s=245.3$ MPa。

解 1）确定载荷　由于吊钩螺纹为松螺纹联接，没有预紧力，所以吊钩挂的重力就是螺纹所受之力。用统计的方法根据吊钩每天的吊重情况，可确定螺纹上承受的名义应力及每一名义应力作用的次数，见表 6.5-33 中的第一列及第二列。由统计表中可知，吊钩每天工作的总循环数 $N=144$ 次，每一应力水平的循环数 n_i 由表中第一列可知，则 n_i/N，即各应力水平所占总循环数的百分数见表中第二列。

2）确定各系数　根据 20 钢锻造的 $\sigma_b=412$ MPa，由表 6.5-7 得有效应力集中系数 $K_\sigma=3.0$（估值）

查图 6.5-136，得 $\varepsilon=0.85$

查图 6.5-137，得 $\beta=0.88$（螺纹为粗车表面）

由此得

$$\frac{K_\sigma}{\varepsilon\beta}=\frac{3.0}{0.85\times0.88}=4.0$$

螺杆的应力状态是脉动循环变幅应力，将名义应力乘以 $K_\sigma/\varepsilon\beta=4.0$，得表 6.5-33 中第四列的数据。

3）确定疲劳极限　20 钢的疲劳极限由本章 2.2.1 中的经验公式求得，即

对于对称拉压

$$\sigma_{-1l}=0.23(\sigma_b+\sigma_s)=0.23\times(412+245.3)=151.2\ \text{MPa}$$

对于脉动拉压

$$\sigma_{0l}=1.42\times151.2=214.5\ \text{MPa}$$

将表 6.5-33 中第四列数据与疲劳极限比较可知，表中大部分数值超过疲劳极限。因此，这个螺杆的应力变化情况属于有限寿命设计。

表 6.5-33　计算数据

每天工作的循环数	循环数占的百分数/%	名义应力/MPa	当 $K_\sigma/\varepsilon\beta=4.0$	
			σ_i/MPa	N_i
1	0.695	80.4	323.7	4×10^3
3	2.08	78.5	313.9	6×10^3
5	3.47	73.6	294.3	2.5×10^4
7	4.86	69.7	279.6	4×10^4
9	6.24	63.8	255.1	1×10^5
11	7.64	59.8	240.3	1.7×10^5
13	9.02	55.9	225.6	3.5×10^5
15	10.4	51.0	206.0	1.4×10^6
17	11.8	46.1	186.4	8×10^6
19	13.2	41.2	166.8	$>10^7$
21	14.6	34.3	137.3	$>10^7$
23	16.0	14.2	56.9	$>10^7$

4）确定 $S-N$ 曲线　因没有 20 钢的 $S-N$ 曲线，所以用近似法作 $S-N$ 曲线。在双对数坐标纸上作两条：一点是 $N=10^3$，$\sigma=0.9\sigma_b=0.9\times412=370.8$ MPa；一点是 $N=10^7$，$\sigma=0.45\sigma_b=185.4$ MPa。连接该两点得一斜线，即为所求的 $S-N$ 曲线，如图 6.5-148 所示。

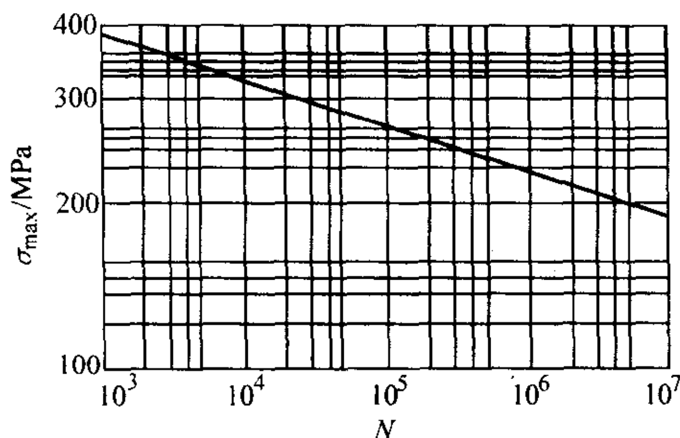


图 6.5-148　20 钢的 $S-N$ 曲线

由图 6.5-148 的 $S-N$ 曲线，查出在应力水平 σ_i 下到达破坏时的循环数 N_i，列于表 6.5-33 中的第五列。由该列的数值可看到，当 $\sigma_i < 186.4$ MPa 以后，$N_i > 10^7$。但由经验公式求得的 $\sigma_{0l} = 214.5$ MPa，大于 186.4 MPa，说明两种假设的近似法之间有误差。本题按表 6.5-33 中数据计算偏于安全。

假设 $N_i \geqslant 10^7$ 时，不产生疲劳损伤，则总寿命为

$$N = 1\bigg/\left[\frac{0.006\,95}{4\times10^3}+\frac{0.020\,8}{6\times10^3}+\frac{0.034\,7}{2.5\times10^4}+\frac{0.048\,6}{4\times10^4}+\frac{0.062\,4}{10^5}+\frac{0.076\,4}{1.7\times10^5}+\frac{0.090\,2}{3.5\times10^5}+\frac{0.104}{1.4\times10^6}+\frac{0.118}{8\times10^6}\right] = 1.082\times10^5$$

因每天工作循环数为 144，则工作天数为

$$\frac{1.082\times10^5}{144} = 752\ \text{d}$$

如起重机每年工作 360 d，则工作年数为

$$\frac{752}{360} = 2.09\ \text{a}$$

即该起重机吊钩的螺杆部分的寿命为 2.09 a，如这部分为吊钩的薄弱环节，为保证安全工作，每工作 2 a 后，需要更新。

5　现代疲劳强度设计

常规疲劳设计是以名义应力为基本设计参数，根据名义应力进行抗疲劳设计。而实际上决定零构件疲劳强度和寿命的是应变集中（或应力集中）处的最大局部应力和应变。因此，在低周疲劳研究和应变分析研究成果基础上，不同于常规疲劳设计的新的疲劳寿命估算方法——局部应力应变法，被称为现代疲劳设计方法。

它的设计思路是，零构件的疲劳破坏都是从应变集中部位的最大应变处起始，并且在裂纹萌生以前都要产生一定的局部塑性变形，局部塑性变形是疲劳裂纹萌生和扩展的先决条件。因此，决定零构件疲劳强度和寿命的是应变集中处的最大局部应力应变，只要最大局部应力应变相同，疲劳寿命就相同。因而有应力集中的零构件的疲劳寿命，可以使用局部应力应变相同的光滑试样的应变－寿命曲线进行计算，也可使用局部应力应变相同的光滑试样进行疲劳试验来模拟。

该方法有以下优点。

1）应变是可以测量的，而且已被证明是一个与低周疲劳相关的极好参数，根据应变分析的方法，可将高、低周疲劳寿命的估算方法统一起来。

2）使用这种方法时，只需知道应变集中部位的局部应力应变和基本的材料疲劳性能数据，就可以估算零件的裂纹形成寿命，避免了大量的结构疲劳试验。

3）这种方法可以考虑载荷顺序对应力应变的影响，特别适用于随机载荷下的寿命估算。

4）这种方法易于与计数法结合起来，可以利用计算机进行复杂的计算。

尽管局部应力应变法有许多优点，但并不能取代名义应力法。因为

1）这种方法只能用于有限寿命下的寿命估算，而不能用于无限寿命，当然也无法代替常规的无限寿命设计法；

2）这种方法目前还不够完善，未考虑尺寸因素和表面情况的影响，对高周疲劳有较大误差；

3）这种方法目前主要限于对单个零件进行分析，对于复杂的联接件，难于进行精确的应力应变分析，难于使用。

还应指出，用名义应力有限寿命设计法估算出的是零件总寿命，而局部应力应变法估算出的是裂纹形成寿命。这种方法常与断裂力学方法联合使用，用这种方法估算出裂纹形成寿命以后，再用断裂力学方法估算出裂纹扩展寿命，两阶段寿命之和即为零件的总寿命。

5.1　$\varepsilon-N$ 曲线

低周疲劳的应力水平较高，其峰值应力常高于材料的弹性极限，有明显宏观塑性变形，故低周疲劳又称为应变疲劳或塑性疲劳。

低周疲劳中的 $S-N$ 曲线，常以 $\varepsilon-N$ 曲线形式给出。在 $\varepsilon-N$ 曲线中，N 可以是循环数，也可以是“反向”数，在恒幅载荷中，反向数为循环数的两倍，所以有些资料中的横坐标用“$2N$”作为计量单位。

图 6.5-149～图 6.5-165 是机械和航空行业中几种常用材料的应变－寿命曲线图。

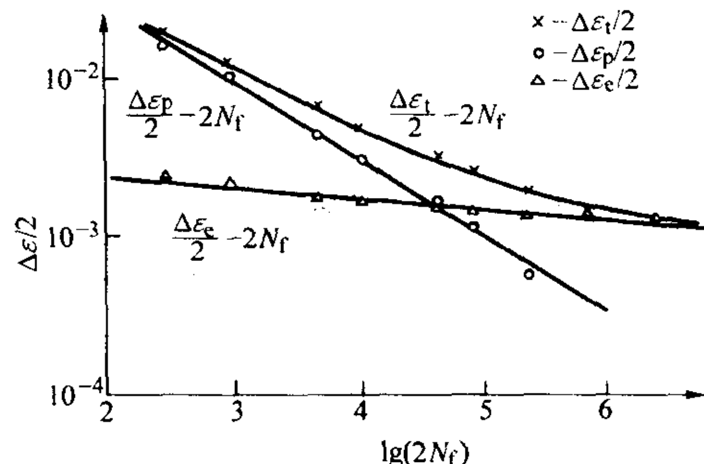


图 6.5-149　Q235A 钢的应变－寿命曲线

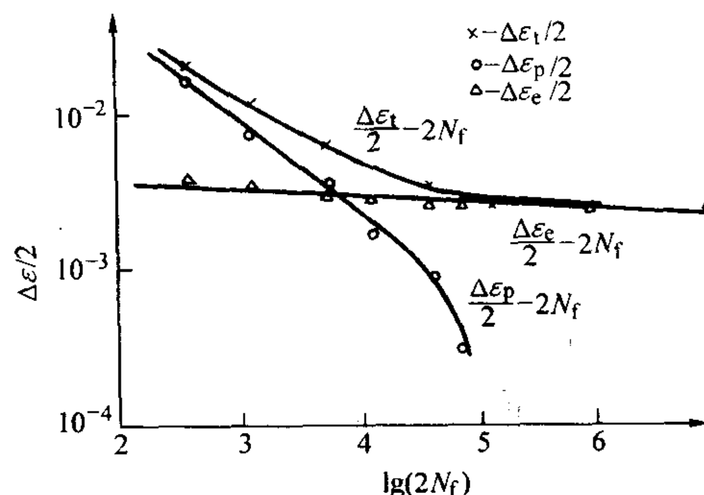


图 6.5-150　45 钢的应变－寿命曲线

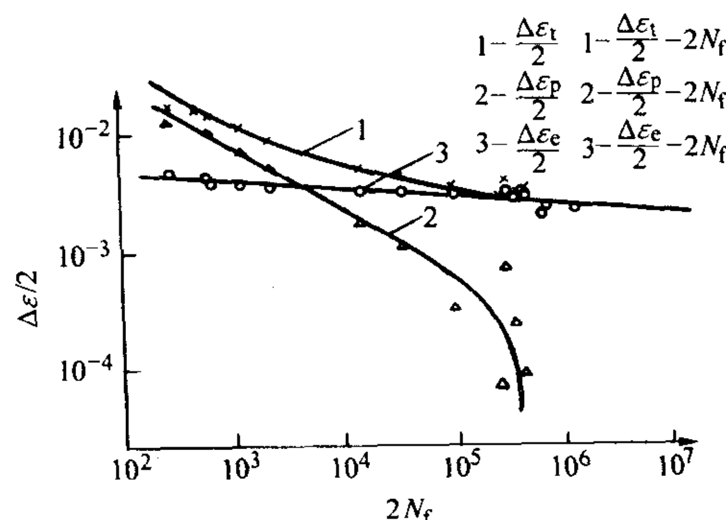


图 6.5-151　40Cr 钢的应变－寿命曲线

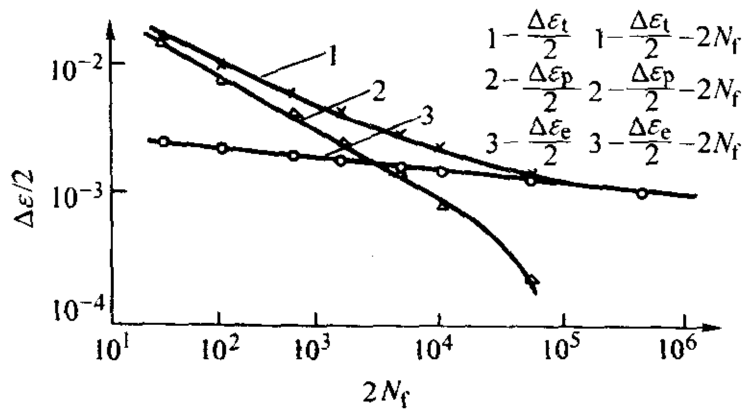


图 6.5-152　Q345 钢应变-寿命曲线

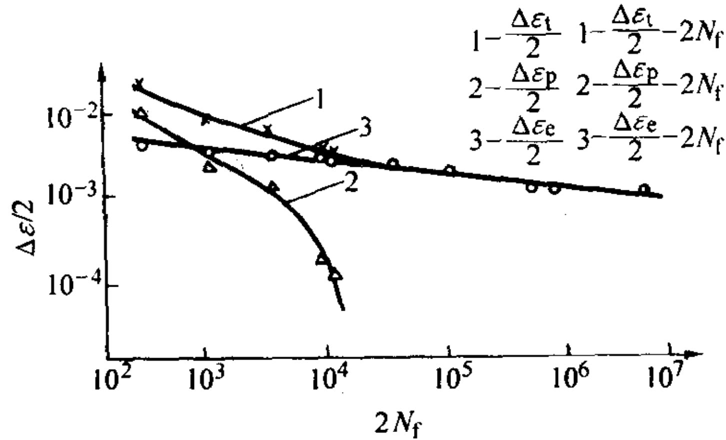


图 6.5-153　60Si2Mn 钢的应变-寿命曲线

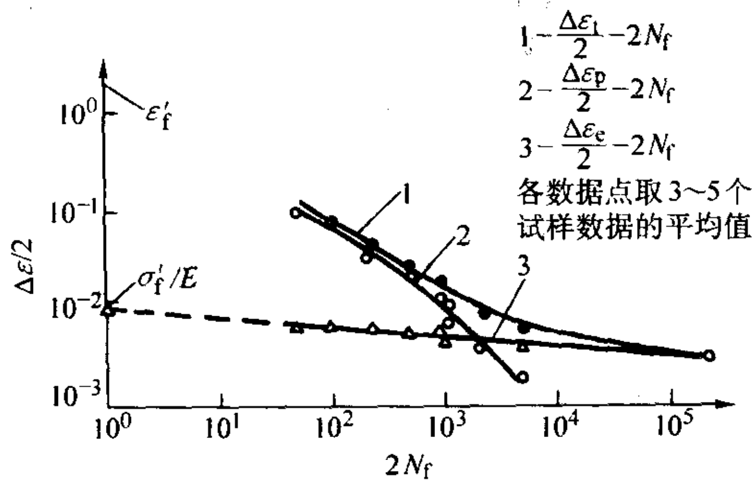


图 6.5-154　30CrMnSiA 钢的应变-寿命曲线

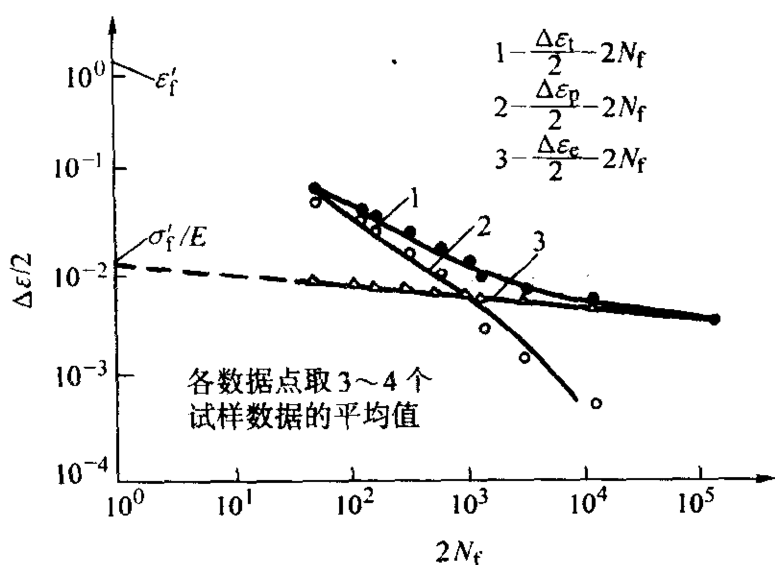


图 6.5-155　30CrMnSiNi2A 钢的应变-寿命曲线

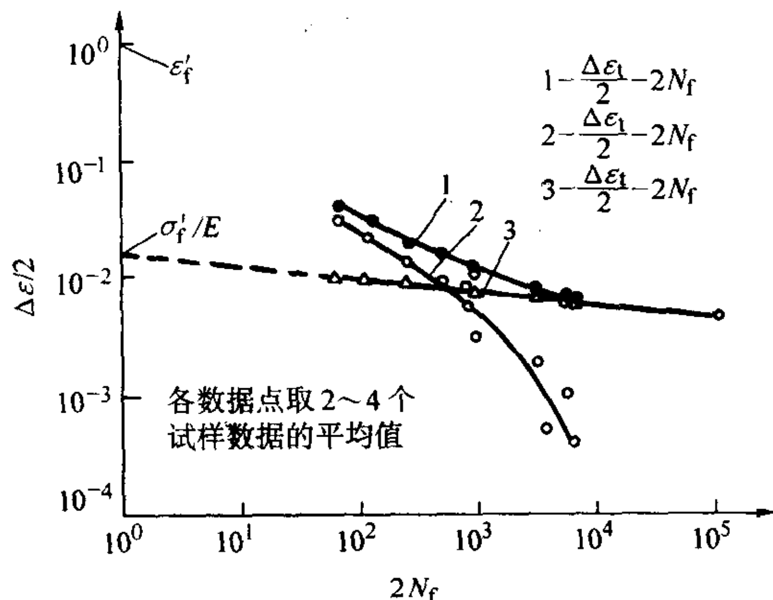


图 6.5-156　40CrMnSiMoVA 钢的应变-寿命曲线

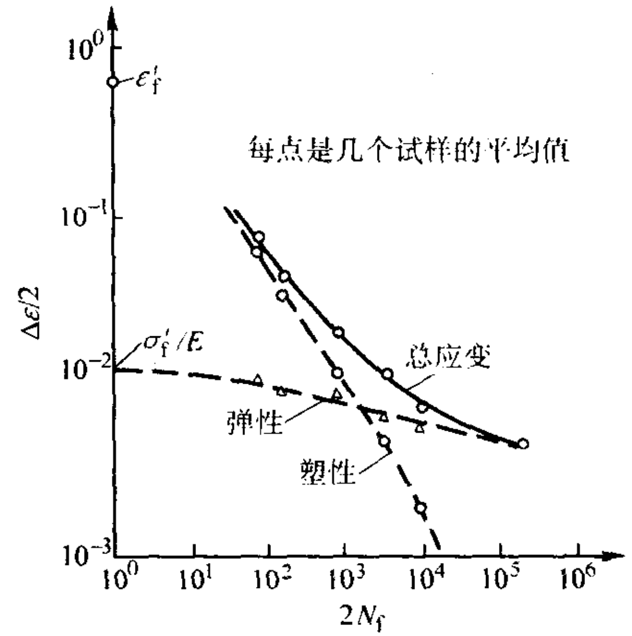


图 6.5-157　Ti-8Al-1Mo-1V 钛合金的应变-寿命曲线

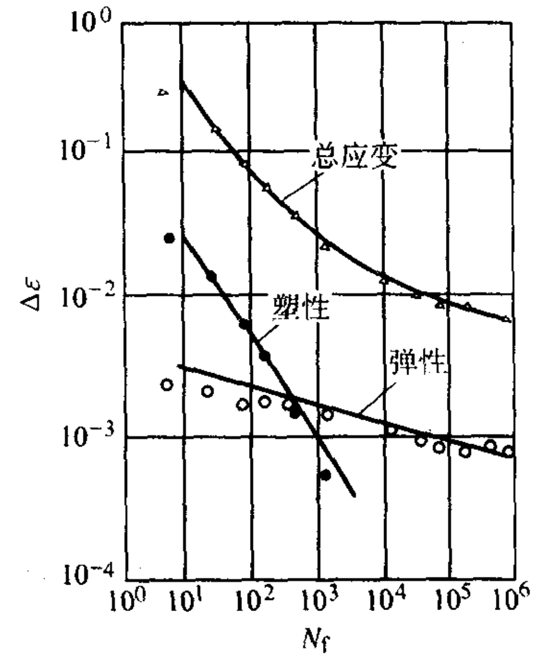


图 6.5-158　Ti-6Al-4V 钛合金的应变-寿命曲线

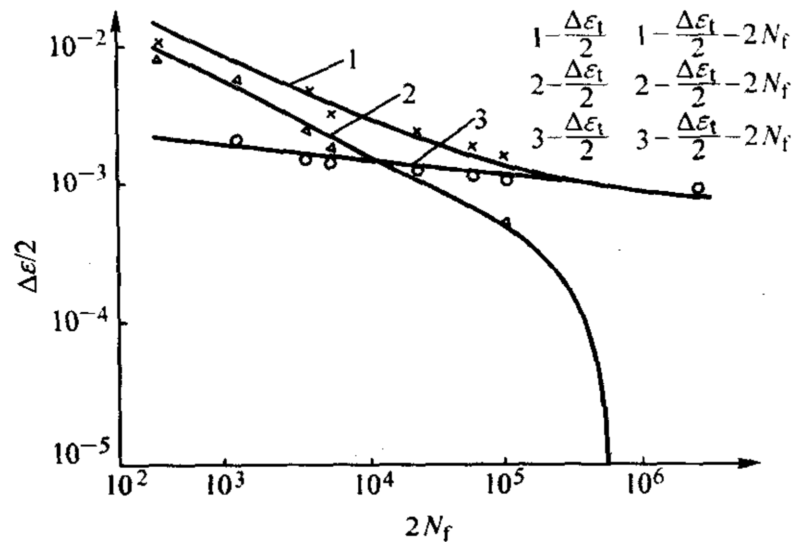


图 6.5-159　ZG270-500 铸钢的应变-寿命曲线

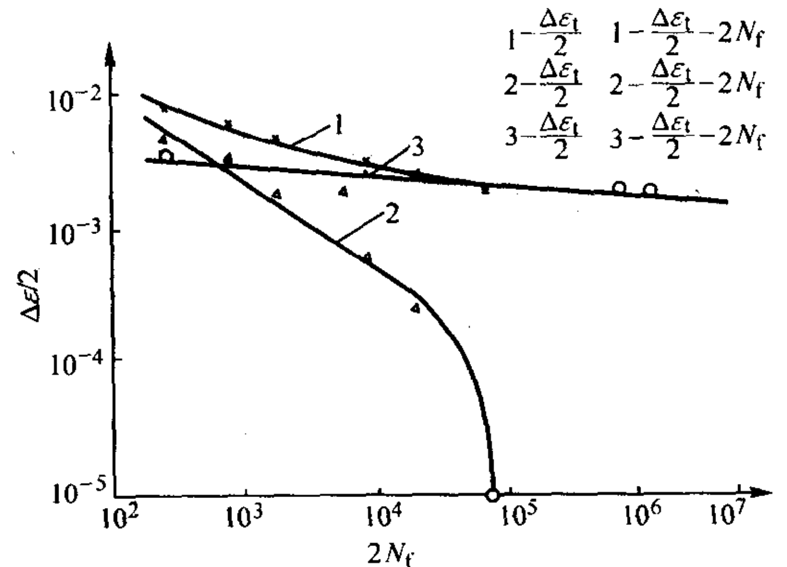


图 6.5-160　QT600-3 球铁（铸件为 Y 型试块）的应变-寿命曲线

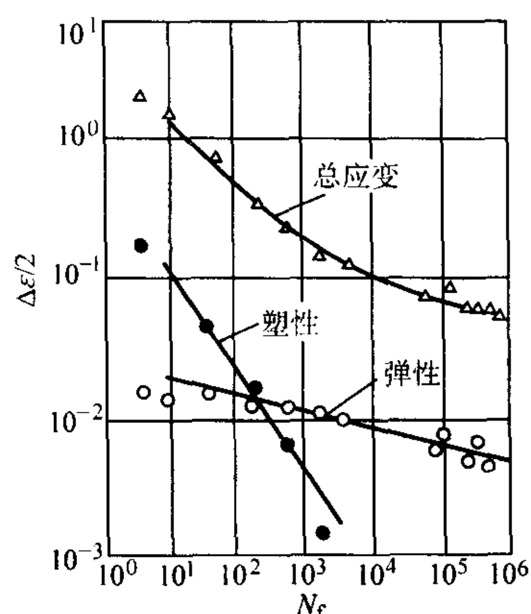


图 6.5-161 2014-T6 铝合金的应变-寿命曲线

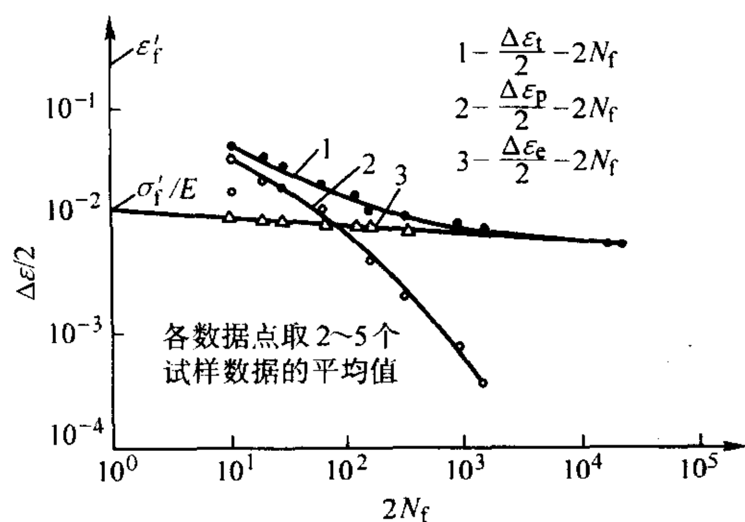


图 6.5-162 7A04-T6 铝合金的应变-寿命曲线

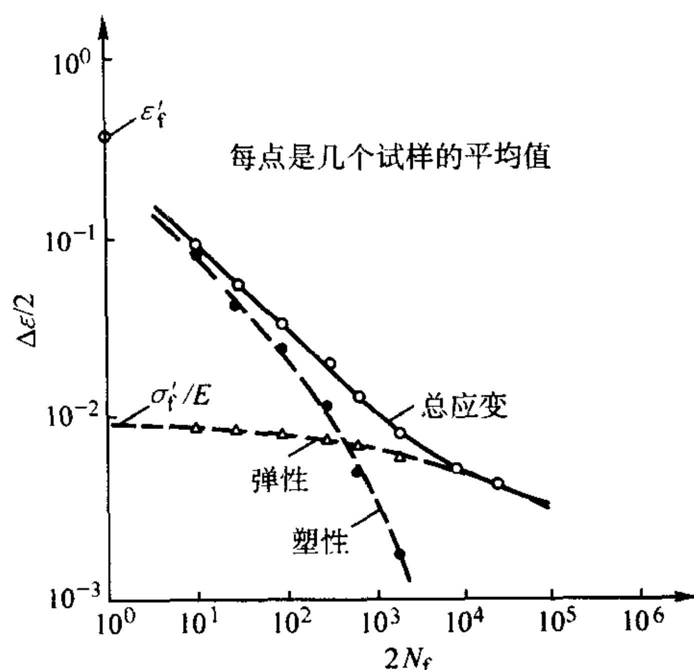


图 6.5-163 2024-T4 铝合金的应力-寿命曲线

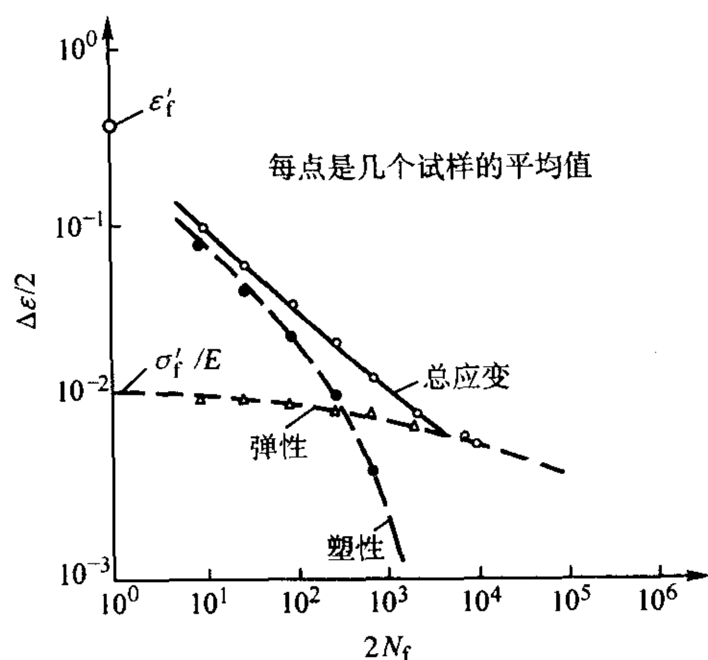


图 6.5-164 7075-T6 铝合金的应变-寿命曲线

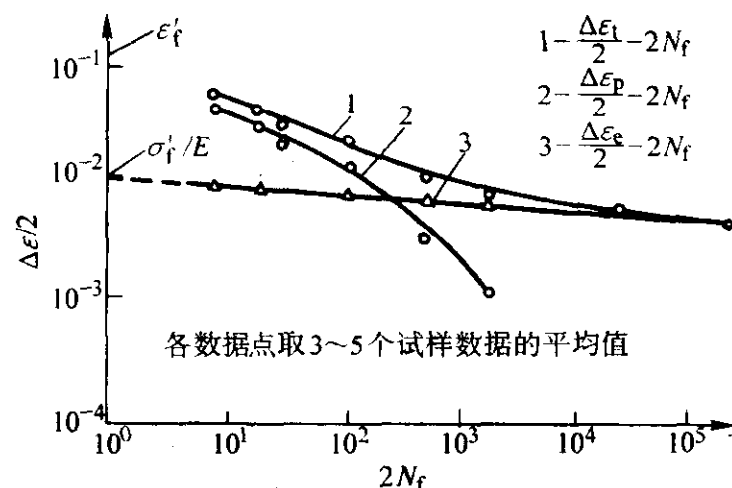


图 6.5-165 2A12-T4 铝合金（棒材）的应变-寿命曲线

5.2 循环 $\sigma-\varepsilon$ 曲线

5.2.1 滞回线

试样一次拉伸试验的应力-应变曲线为 OA，如图 6.5-166a 所示。若用相同的试样作压缩试验，则应力-应变曲线为 OB。曲线 BOA 表示材料一次加载的应力-应变关系，称为单调应力-应变（$\sigma-\varepsilon$）曲线。一般仅考虑 OA 段曲线。

将试样先拉伸，应力-应变曲线由 O 点到 A 点；然后进行压缩，应力-应变曲线由 A 点到 B 点；再进行拉伸，应力-应变曲线由 B 点回到 A 点，完成一个应力循环，如图 6.5-166b 所示。这种应力-应变循环曲线称为滞回线。滞回线不仅表示了应力的循环变化，还能反映每个循环中塑性应变的大小。

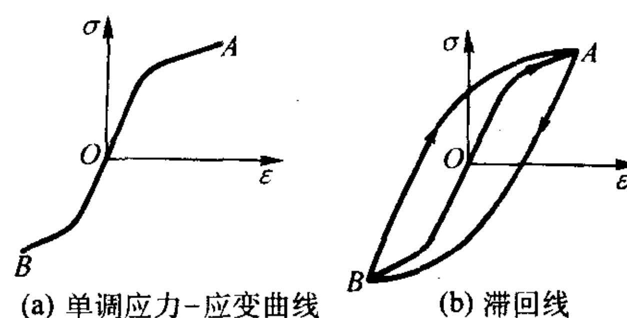


图 6.5-166 应力-应变曲线

5.2.2 循环硬化与循环软化

对于循环硬化的材料，其应变抗力随着循环数的增加而增大。因此，在恒应变幅度下，材料在每一循环中所需的应力将随循环数的增加而逐渐增大；或在恒应力幅度下，材料在每一循环中的应变量随循环数的增加而变小。

对于循环软化的材料，其应变抗力随着循环数的增加而变小。因此，在恒应变幅度下，材料在每一循环中所需的应力将随循环数的增加而逐渐变小；或在恒应力幅度下，材料在每一循环中的应变量随循环数的增加而变大。

材料是循环硬化还是循环软化，由材料的屈强比 σ_s/σ_b 而定。屈强比小于 0.7 时，材料产生循环硬化；屈强比大于 0.8 时，材料产生循环软化。所以，一般的退火材料产生循环硬化，冷加工的材料产生循环软化。

无论是循环硬化材料或循环软化材料，虽然在试验开始阶段所得的应力-应变滞回线并不闭合，但经过一定次数的循环后，滞回线接近于封闭环，即可得到稳定的滞回线。把应变幅控制在不同的水平上，可以得到一系列大小不同的稳定的滞回线，将这些滞回线的顶点连接起来，便得到如图 6.5-167 所示的曲线 OC，这曲线称为该金属材料的循环应力-应变（$\sigma-\varepsilon$）曲线。

5.2.3 循环应力-应变曲线

根据图 6.5-167 循环应力-应变曲线的作图法可知，曲线上的任一点实际上是一个滞回线的顶点，其坐标为该滞回

线的应力幅 σ_a 和应变幅 ε_a。因此，循环应力-应变曲线可以用下式拟合，即

$$\varepsilon_a=\varepsilon_e+\varepsilon_p=\frac{\sigma_a}{E}+\left(\frac{\sigma_a}{K'}\right)^{\frac{1}{n'}} \tag{6.5-26}$$

或写成幅度的形式（应力幅度 $\Delta\sigma=2\sigma_a$），应变幅度 $\Delta\varepsilon=2\varepsilon_a$），即

$$\frac{\Delta\varepsilon}{2}=\frac{\Delta\sigma}{2E}+\left(\frac{\Delta\sigma}{2K'}\right)^{\frac{1}{n'}} \tag{6.5-27}$$

式中，ε_e 为应变幅的弹性分量；ε_p 为应变幅的塑性分量；ε_a 为总应变幅；K' 为循环强度系数；n' 为循环应变硬化指数。

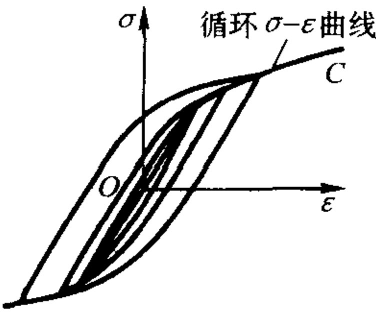


图 6.5-167　循环应力-应变曲线

图 6.5-168～图 6.5-184 给出机械和航空行业中几种常用材料的循环稳定和单调拉伸的应力-应变曲线。

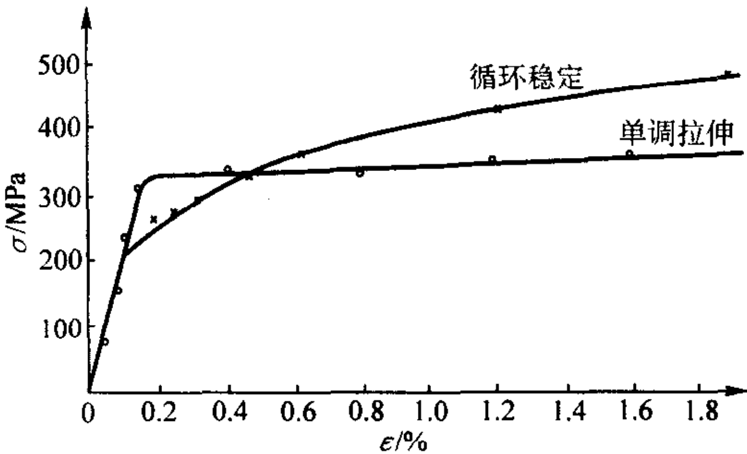


图 6.5-168　Q235A 钢的循环稳定与单调拉伸应力-应变曲线

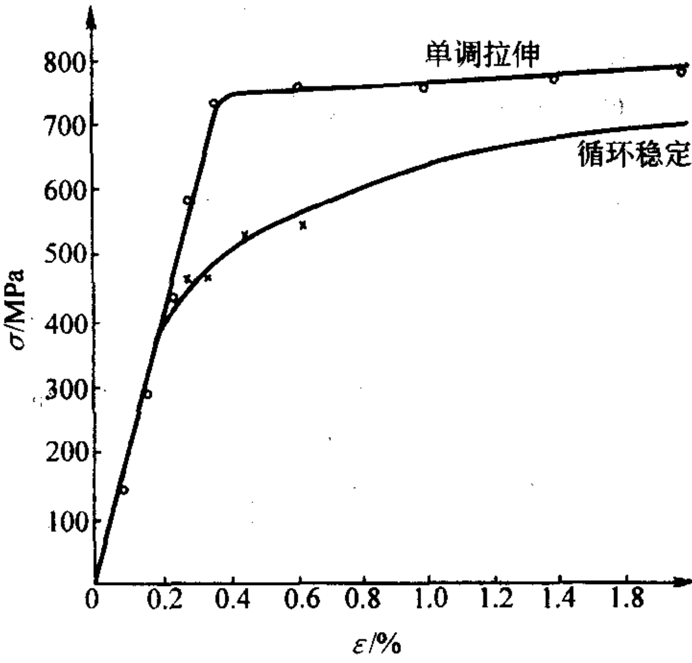


图 6.5-169　45 钢的循环稳定与单调拉伸应力-应变曲线

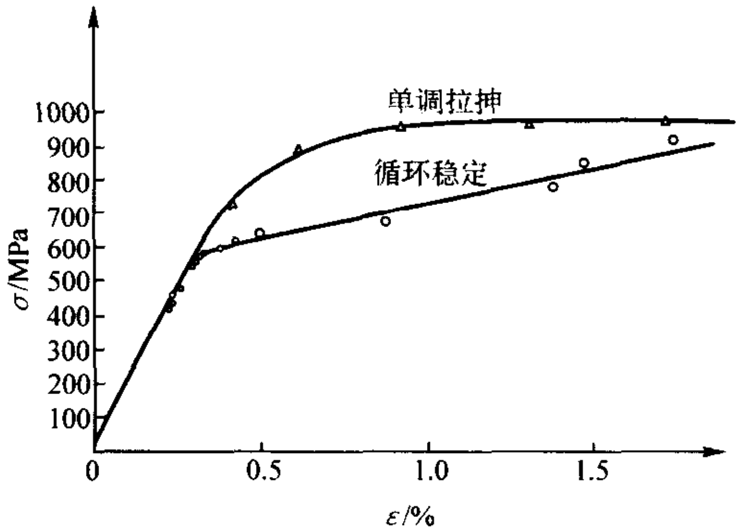


图 6.5-170　40Cr 钢循环稳定与单调拉伸应力-应变曲线

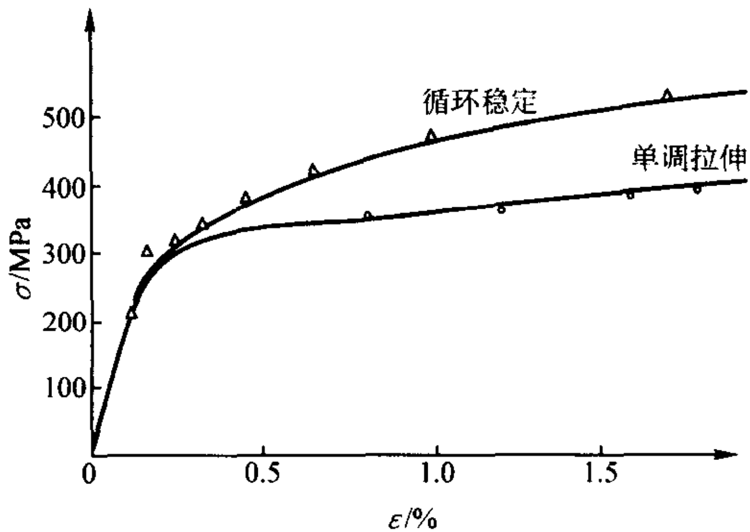


图 6.5-171　16Mn 钢的循环稳定与单调拉伸应力-应变曲线

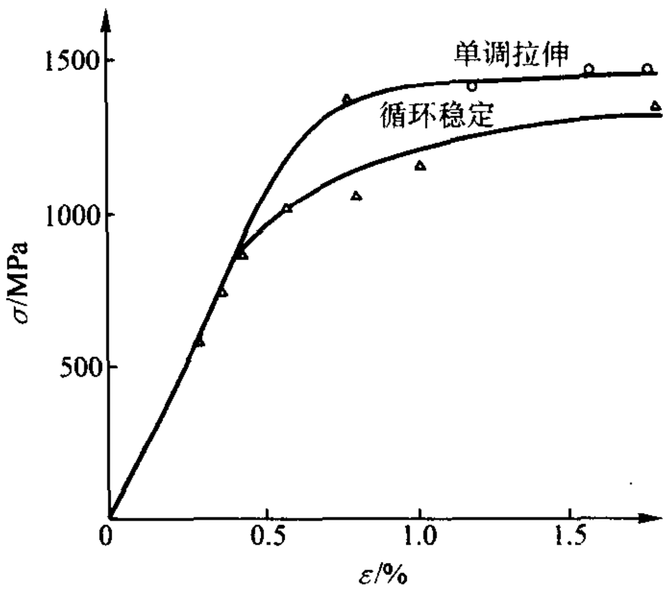


图 6.5-172　60Si2Mn 钢的循环稳定与单调拉伸应力-应变曲线

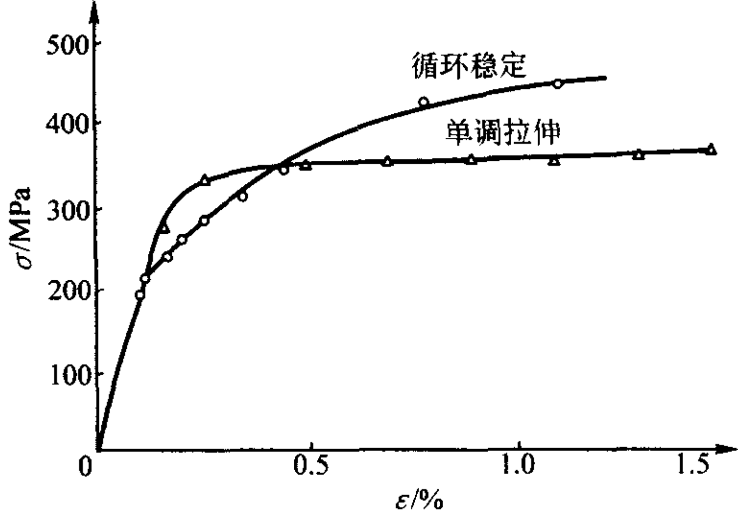


图 6.5-173　ZG270-500 铸钢的循环稳定与单调拉伸应力-应变曲线

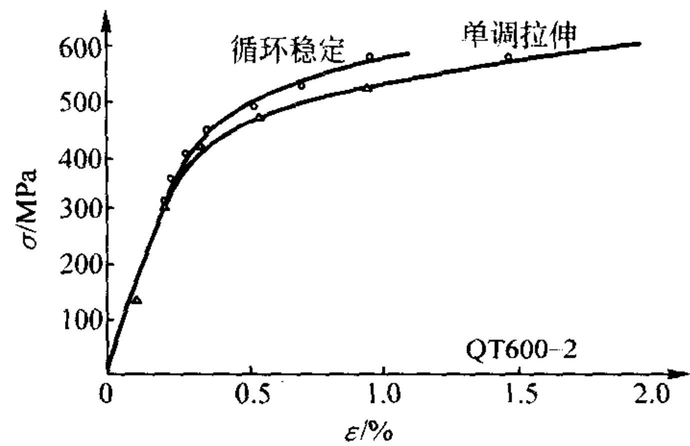


图 6.5-174 QT600-3 球铁（铸件为 Y 型试块）的循环稳定与单调拉伸应力-应变曲线

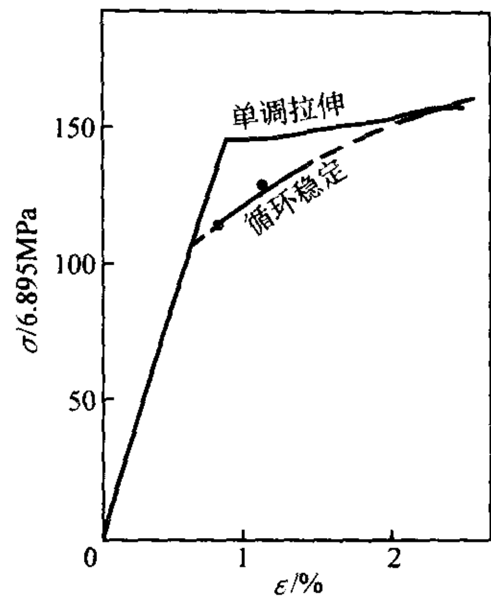


图 6.5-178 Ti-8Al-1Mo-1V 钛合金的循环稳定与单调拉伸应力-应变曲线

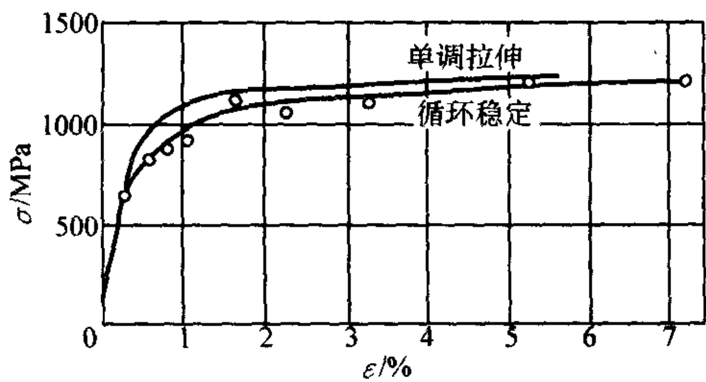


图 6.5-175 30CrMnSiA 钢的循环稳定与单调拉伸应力-应变曲线

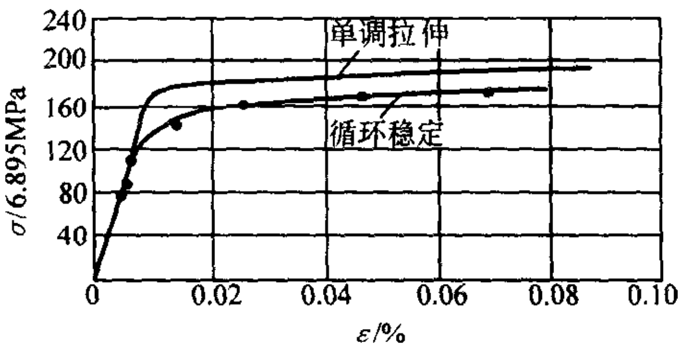


图 6.5-179 Ti-6Al-4V 钛合金的循环稳定与单调拉伸应力-应变曲线

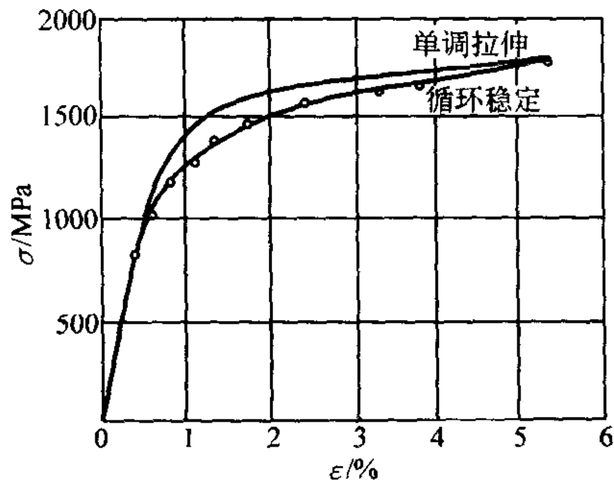


图 6.5-176 30CrMnSiNi2A 钢的循环稳定与单调拉伸应力-应变曲线

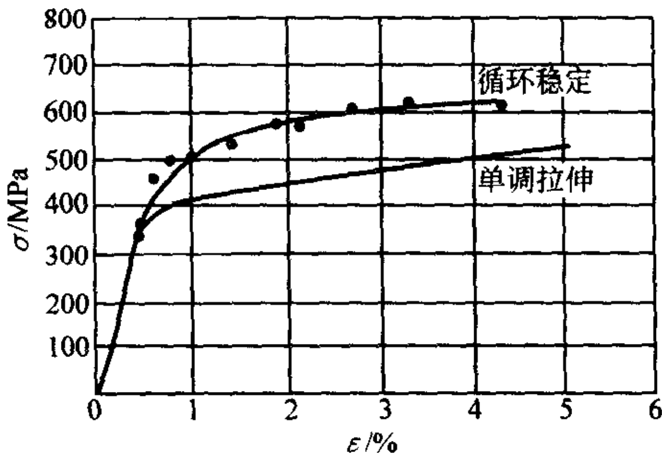


图 6.5-180 2A12-T4 铝合金（棒材）的循环稳定与单调拉伸应力-应变曲线

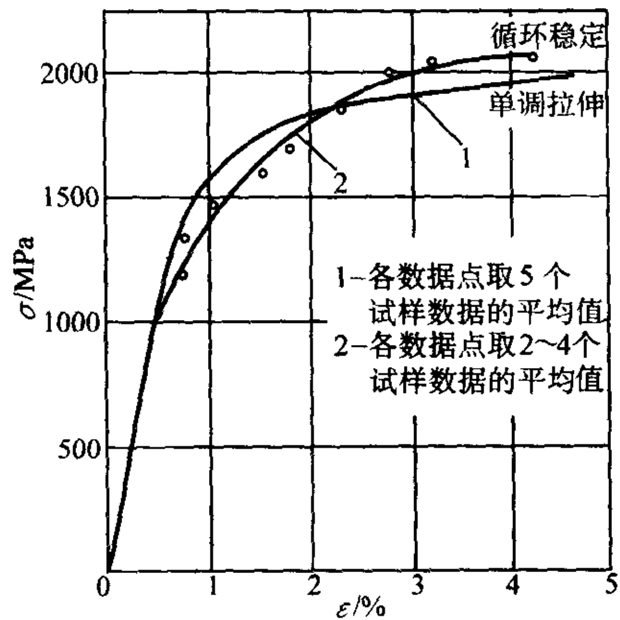


图 6.5-177 40CrMnSiMoVA 钢的循环稳定与单调拉伸应力-应变曲线

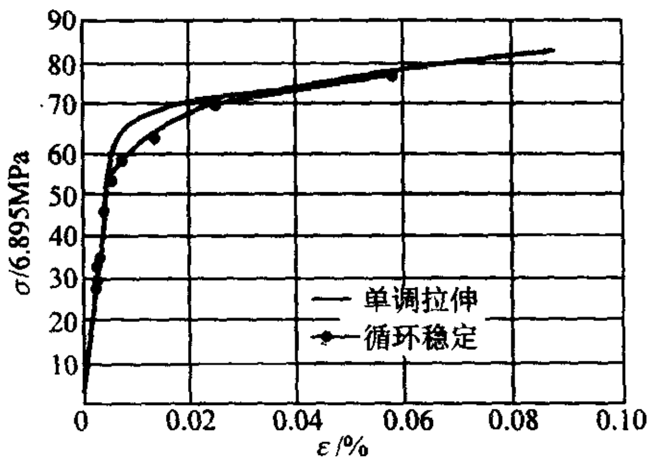


图 6.5-181 2014-T6 铝合金的循环稳定与单调拉伸应力-应变曲线

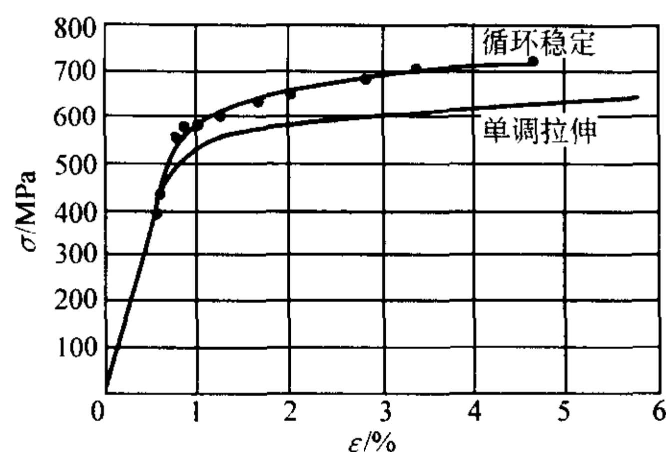


图 6.5-182 7A04-T6 铝合金的循环稳定与单调拉伸应力-应变曲线

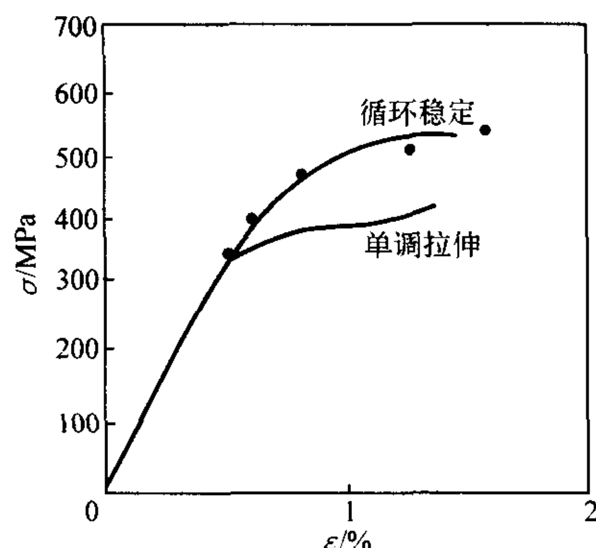


图 6.5-183 2024-T4 铝合金的循环稳定与单调拉伸应力-应变曲线

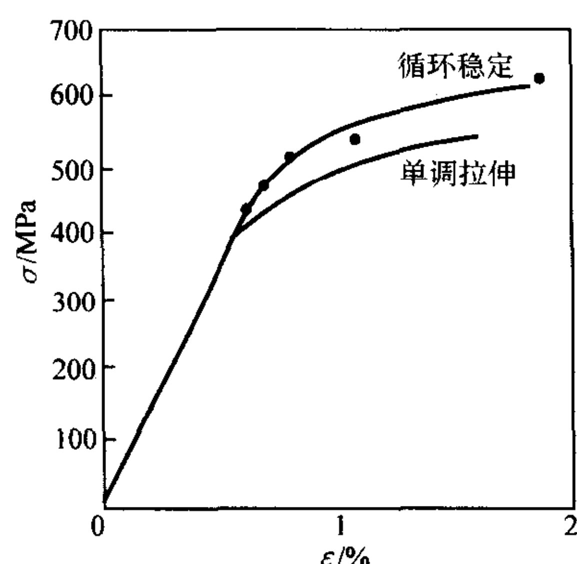


图 6.5-184 7075-T6 铝合金的循环稳定与单调拉伸应力-应变曲线

5.3 应变-寿命曲线的获得与低周疲劳寿命估算

5.3.1 曼森-科芬方程

准备一组材料和尺寸完全相同的试样，对每个试样施加不同的载荷，试样将产生不同的应变，疲劳循环次数由计数器自动记录，这样就得到一组应变和破坏循环数的记录数据。由于试验时控制总应变幅常常比较方便，所以得到的数据，一般是总应变幅与破坏循环数。图 6.5-149～图 6.5-165 就是对不同材料得出的总应变幅 ε_a ($\Delta\varepsilon/2$) 与破坏循环数 N 的曲线，即 $\varepsilon-N$ 曲线。

每一个总应变值可分为弹性应变分量和塑性应变分量(图 6.5-185)。假设在总应变幅为 0.6%时的疲劳寿命为 10^4 次循环。根据实测可知，总应变幅中 1/3 为塑性应变幅，其余 2/3，即 0.4%为弹性应变幅。反之，对同一种材料，只要循环弹性应变幅等于 0.4%，其寿命将是 10^4 次循环。同样，只要知道塑性应变幅为 0.2%，也可以推断它的寿命为 10^4 次循环。

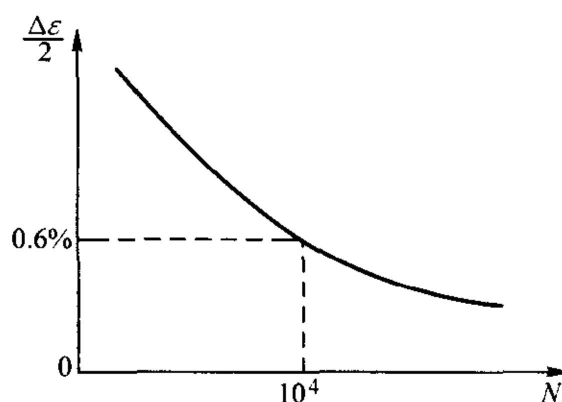


图 6.5-185 总应变幅-寿命曲线

指定一个弹性应变幅或塑性应变幅，就可以得到破坏循环数 N。因此，在同一张总应变幅-寿命曲线图上，可以画出弹性应变-寿命曲线和塑性应变-寿命曲线。在双对数坐标图上，弹性应变-寿命曲线和塑性应变-寿命曲线都是一条近似直线，如图 6.5-186 所示。这两直线的交点 P，称为过渡寿命点；P 点在横轴上的坐标 N_T，称为过渡寿命，它是一试验常数。交点 P 表示低周疲劳与高周疲劳的分界点：在 P 点的右侧，弹性应变起主导作用，在 P 点的左侧，塑性应变起主导作用。或者说，P 点的右侧为高周疲劳区，P 点的左侧为低周疲劳区。当提高材料强度时，P 点左移，提高材料韧性时，P 点右移。

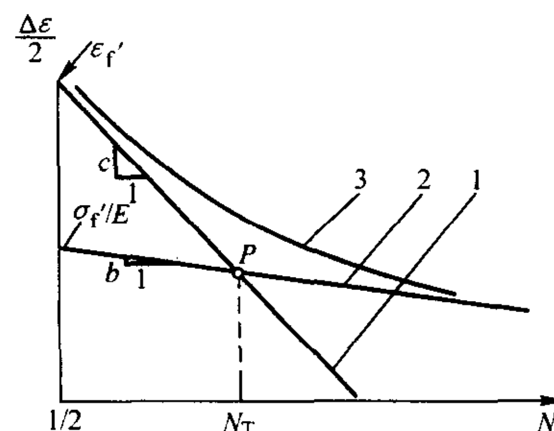


图 6.5-186 通用斜率法的应变-寿命曲线
（双对数坐标）
1—塑性应变-寿命曲线；2—弹性应变-寿命曲线；
3—总应变-寿命曲线

图 6.5-186 中塑性应变-寿命曲线 1 的方程，可以用幂指数函数形式表示为

$$\Delta\varepsilon_p N^{\beta} = C_1 \quad (6.5\text{-}28)$$

弹性应变幅度 $\Delta\varepsilon_e$ 和塑性应变幅度 $\Delta\varepsilon_p$ 还可以写成一般常用的形式，即

$$\frac{\Delta\varepsilon_e}{2} = \frac{\sigma_f'}{E}(2N)^b$$

$$\frac{\Delta\varepsilon_p}{2} = \varepsilon_f'(2N)^c$$

式中，σ_f'为疲劳强度系数；σ_f'/E 为循环数 $N=\frac{1}{2}$处直线 2 的纵坐标截距；b 为疲劳强度指数，直线的斜率；E 为材料的弹性模量；ε_f'为疲劳塑性系数，$N=\frac{1}{2}$处直线 1 的纵坐标截距；c 为疲劳塑性指数。

总应变幅-寿命曲线 3 的数学表达式为

$$\frac{\Delta\varepsilon}{2} = \frac{\Delta\varepsilon_e}{2} + \frac{\Delta\varepsilon_p}{2} = \frac{\sigma'_f}{E}(2N)^b + \varepsilon_f'(2N)^c \quad (6.5\text{-}29)$$

这里 N 为反向次数，“$2N$”在恒幅循环载荷中为循环次数。式 (6.5-29) 称为曼森-科芬方程。

式 (6.5-28) 和式 (6.5-29) 中的 6 个参数：b、c、ε'_f 和 σ'_f，是表征低周疲劳特性的主要参数。对于机械设计中几种常用钢材的 6 个参数见表 6.5-34。

表 6.5-34　低周疲劳性能参数

材料	K′/MPa	n′	b	c	ε′f	σ′f/MPa
45 钢（正火）	1 153	0.179	-0.123	-0.526	0.465	1 115
40Cr	1 592	0.173	-0.120	-0.559	0.388	1 306
40CrNiMoA	1 439	0.152	-0.061	-0.643	0.463	898
Q345	1 045	0.151				
45 钢（调质）	1 324	0.160				
20 钢（热轧）	772	0.18	-0.12	-0.51	0.41	896

5.3.2　四点法求应变 - 寿命曲线

曼森指出，确定 $\Delta\varepsilon_e-N$ 和 $\Delta\varepsilon_p-N$ 两条直线只要四个点。这四个点可以由单调拉伸试验数据获得，而不用去做疲劳试验。如图 6.5-187，该四个点分别为：

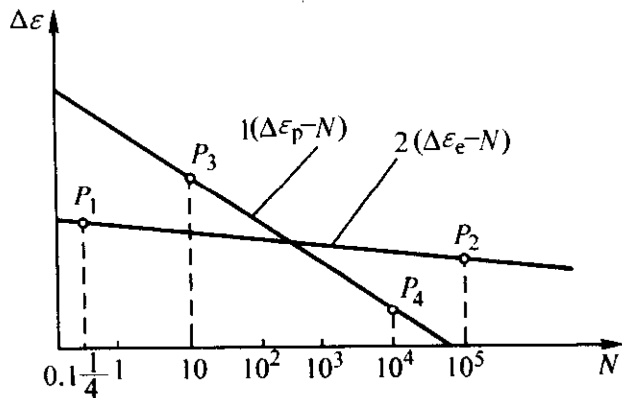


图 6.5-187　四点法求应变 - 寿命曲线

P_1 为对应于$\frac{1}{4}$次循环（即一次拉伸至破坏）的应变幅度的弹性分量

$$\Delta\varepsilon_e = 2.5\ (\sigma_f/E) \qquad (6.5\text{-}30)$$

P_2 为对应于 10^5 次循环的应变幅度的弹性分量

$$\Delta\varepsilon_e = 0.90\ (\sigma_b/E) \qquad (6.5\text{-}31)$$

连接 P_1 和 P_2 点，得图 6.5-187 中的曲线 2，即 $\Delta\varepsilon_e$-N 曲线。这里 $\Delta\varepsilon_e$ 为弹性应变幅度；N 为破断循环数；σ_f 为单调拉断时的真实应力；σ_b 为强度极限。

P_3 为对应于 10 次循环的应变幅度的塑性分量为

$$\Delta\varepsilon_p = \frac{1}{4}\varepsilon_f^{3/4} \qquad (6.5\text{-}32)$$

P_4 为对应于 10^4 次循环的应变幅度的塑性分量为

$$\Delta\varepsilon_p = \frac{0.013\,2 - \Delta\varepsilon_e^*}{1.91} \qquad (6.5\text{-}33)$$

连接 P_3 和 P_4 点，得 $\Delta\varepsilon_p-N$ 曲线 1。这里 $\Delta\varepsilon_e^*$ 为曲线 2 上 $N=10^4$ 所对应的弹性应变幅度；ε_f 为单调拉断时的真实应变，用截面收缩率 ψ（以%计）近似求得

$$\varepsilon_f = \ln\frac{100}{100-\psi} \qquad (6.5\text{-}34)$$

用四点法求材料的应变 - 寿命曲线，适合于碳钢、合金钢、铝、钛等金属材料。

曼森对 29 种材料的疲劳试验结果进行了整理归纳，在双对数坐标平面上得出（参见图 6.5-187）塑性应变 - 寿命直线 1 的斜率为 -0.6，弹性应变 - 寿命直线 2 的斜率为 -0.12，从而得到下面的关系式，即

$$\Delta\varepsilon = 3.5\frac{\sigma_b}{E}N^{-0.12} + \varepsilon_f^{0.6}N^{-0.6} \qquad (6.5\text{-}35)$$

由于斜率是根据 29 种材料归纳出来的，即这个斜率对多种材料通用，故本法称为通用斜率法。

5.3.3　低周疲劳寿命估算

估算低周疲劳寿命常用两种方法：①类似常规疲劳设计方法，即用 $\varepsilon-N$ 曲线直接推算出寿命；②用局部应力 - 应变法估算裂纹形成寿命（见本章第 5.4 小节）。

用应变 - 寿命（$\varepsilon-N$）曲线直接推算出寿命时，关键是获得材料的 $\varepsilon-N$ 曲线。这可以通过疲劳试验获得，如图 6.5-149～图 6.5-165；或用四点法求得弹性应变幅度 - 寿命（$\Delta\varepsilon_e-N$）曲线和塑性应变幅度 - 寿命（$\Delta\varepsilon_p-N$）曲线（见图 6.5-187），然后将弹性应变幅度与塑性应变幅度相加得总应变幅度，得出总应变幅度 - 寿命曲线。当应变比 $r=-1$ 时，得 ε_a-N 曲线。

在实际计算中，一般可按弹性理论求应力幅 σ_a，然后假设以 σ_a 为理论弹性应力幅，近似用公式 $\sigma_a=\frac{1}{2}E\varepsilon_a$ 计算 ε_a，最后用 ε_a-N 曲线直接推算出疲劳寿命。

当给出材料低周疲劳的应力 - 寿命（σ_a-N）曲线时，也可以用弹性理论求得的 σ_a，直接从 σ_a-N 曲线推算出疲劳寿命。

上述的寿命估算方法，是用材料力学或弹性理论的方法来计算零件和构件危险点的名义应力为出发点的，故称这种方法为名义应力法。而低周疲劳的应力 - 寿命曲线中的 σ_a，是真实应力幅，应变 - 寿命曲线中的 ε_a 是真实应变幅，所以名义应力法在低周疲劳寿命估算中，误差很大，只能用于粗略的寿命估算。对于较重要设备的寿命估算，建议用本章第 5.4 节的局部应力 - 应变法估算疲劳寿命。

表 6.5-35 和表 6.5-36 是国产常用的机械材料和航空材料的单调与循环应变特性数据，供寿命估算中应用。

表 6.5-35　某些国产机械材料的单调与循环应变特性

材料	热处理	σb/MPa	σs/σb	K/K′ /（MPa/MPa）	n/n′	εf/ε′f	σf/σ′f /（MPa/MPa）	b	c	E/MPa	循环硬化（软化）特性
Q235A	轧态	470.4	0.69	928.2/969.6	0.259 0/0.182 4	1.021 7/0.274 7	976.4/658.8	-0.070 9	-0.490 7	198 753.4	循环硬化
Q345	轧态	572.5	0.63	856.1/1 164.8	0.181 3/0.187 1	1.072 9/0.464 4	1 118.3/947.1	-0.094 3	-0.539 5	200 741	循环硬化
45	调质	897.7	0.91	928.7/1 112.5	0.036 9/0.115 8	0.839 3/1.504 8	1 511.7/1 041.4	-0.070 4	-0.733 8	193 500	循环软化
40Cr	调质	1 084.9	0.94	1 285.1/1 228.9	0.051 2/0.090 3	0.731 9/0.380 9	1 264.7/1 385.1	-0.078 9	-0.576 5	202 860	循环软化
60Si2Mn	淬火后中温回火	1 504.8	0.91	1 721.2/1 925.0	0.035 0/0.090 6	0.455 7/0.320 3	2 172.4/2 690.6	-0.113 0	-0.582 6	203 395	循环软化

续表 6.5-35

材料	热处理	σ_b/MPa	σ_s/σ_b	K/K' /（MPa/MPa）	n/n'	$\varepsilon_f/\varepsilon'_f$	σ_f/σ'_f /（MPa/MPa）	b	c	E/MPa	循环硬化（软化）特性
ZG270－500	正火	572.3	0.64	1 218.1/1 267.5	0.285 0/0.222 0	0.238 3/0.181 3	809.4/781.5	－0.098 8	－0.506 3	204 555.4	循环硬化
QT450－10①	铸态	498.1	0.79	－/1 127.9	－/0.140 5	－/0.146 1	－/856.9	－0.102 7	－723 7	166 108.5	循环硬化
QT600－3②	正火	748.4	0.61	1 439.9/1 039.8	0.199 6/0.116 5	0.076 0/0.372 5	856.5/885.2	－0.077 7	－0.710 4	154 000	循环硬化
QT600－3①	正火	677.0	0.77	1 621.5/979.3	0.183 4/0.087 6	0.037 7/0.027 1	888.8/1 109.8	－0.105 6	－0.339 3	150 376.5	循环硬化
QT800－2②	正火	913.0	0.64	1 777.3/1 437.7	0.203 4/0.147 0	0.045 5/0.168 4	946.8/1 067.4	－0.083 0	－0.579 2	160 500	循环硬化

① ϕ30 棒料。
② Y 型试块。

表 6.5-36　某些国产航空材料的单调与循环应变特性

材料	热处理	σ_b/MPa	$\sigma_{0.2}$/MPa	K/K' /(MPa/MPa)	n/n'	$\varepsilon_f/\varepsilon'_f$ /（%/%）	σ_f/σ'_f /(MPa/MPa)	b	c	E/MPa	是否 Masing 材料
30CrMnSiA	调质	1 177.0	1 104.5	1 475.76/1 771.93	0.063/0.127	77.27/161.15	1 795.07/1 755.94	－0.085 9	－0.771 2	203 004.9	是
30CrMnSiNi2A	等温淬火后回火	1 655.4	1 308.3	2 355.35/2 647.69	0.091/0.13	74/120.71	2 600.52/2 773.22	－0.102 6	－0.781 6	200 062.8	否
40CrMnSiMoVA	等温淬火后回火	1 875.3	1 513.2	3 150.20/3 411.36	0.146 8/0.14	63.32/96.86	3 511.55/3 254.35	－0.105 4	－0.785 0	200 455.1	否
2A12－T4（棒材）	T4	545.1	399.5	870.47/849.78	0.097/0.158	13.67/18	723.76/643.44	－0.062 7	－0.653 9	73 160.2	否
2A12－T4（板材）	T4	475.6	331.5	545.17/645.79	0.088 9/0.066 9	30.19/16.50	618.04/670.21	－0.102 7	－0.511 4	71 022.3	—
7A04－T6	T6	613.9	570.8	775.05/949.61	0.063/0.08	18.00/24.52	710.62/884.69	－0.072 7	－0.776 1	72 571.8	是
7A09－T74	T74	560.2	518.2	724.64/905.87	0.071/0.101	28.34/77.08	748.47/807.80	－0.074 3	－0.935 1	72 179.5	—

5.4　局部应力应变分析

5.4.1　真实应力与真实应变

工程上常用材料的应力－应变曲线（图 6.5-188a），是由拉伸试验确定的，其名义应力 s 等于载荷 F 除以原始截面面积 A_0，其名义应变 e 为伸长量 ΔL 除以原始长度 L_0（标距长度）（图 6.5-188b）。即

$$s=\frac{F}{A_0} \tag{6.5-36}$$

$$e=\frac{L-L_0}{L_0}=\frac{\Delta L}{L_0}$$

由于在拉伸过程中，试样的截面面积是变化的，直到拉断，则真实应力 σ 为

$$\sigma=F/A \tag{6.5-37}$$

式中，A 为颈缩处的横截面面积。

当试样拉伸至 L 长时，假设试样长度有一微小增量 dL，则此时的应变增量为

$$\mathrm{d}\varepsilon=\frac{\mathrm{d}L}{L}$$

上式由 L_0 至 L 积分，得真实应变为

$$\varepsilon=\ln\frac{L}{L_0} \tag{6.5-38}$$

真实应力、应变与名义应力、应变的关系为

$$\sigma=s\ (1+e) \tag{6.5-39}$$

$$\varepsilon=\ln\ (1+e) \tag{6.5-40}$$

真实应变反映了物体变形的实际情况，也称为自然应变或对数应变；名义应变也称为工程应变。在大应变问题中，只有用真实应变才能得出合理的结果。

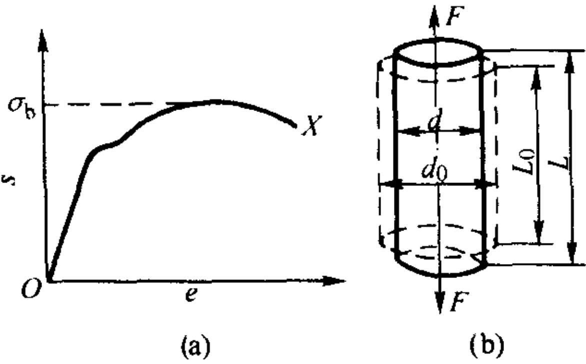


图 6.5-188　应力－应变曲线

5.4.2　玛辛特性

改变应力水平，可以得到不同应力水平下的滞回线（参见图 6.5-167）。图 6.5-189a 为不同应力水平下的滞回线 ADA、BEB、CFC，将坐标轴平移，使原点与各滞回线的最低点相重合，若滞回线的最高点的连线与其上行段迹线相吻合（见图 6.5-189b），则该材料具有玛辛特性，称为玛辛材料。

将材料的循环 $\sigma-\varepsilon$ 曲线画于图 6.5-189b 上，可以看

出，滞回线上行段迹线的纵坐标，为循环 $\sigma-\varepsilon$ 曲线的纵坐标的两倍。

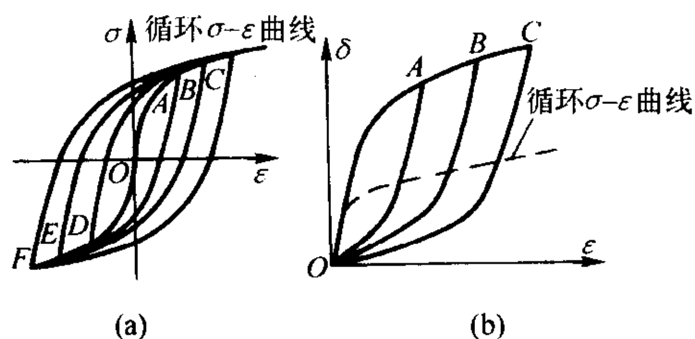


图 6.5-189 坐标平移后的滞回线

5.4.3 材料的记忆特性

图 6.5-190a 表示载荷 - 时间历程，图 6.5-190b 表示材料在该载荷 - 时间历程中的应力 - 应变响应。加载时由 1 到 2，相应的应力 - 应变响应由 A 到 B；由 2 到 3 加反向载荷时，应力 - 应变曲线由 B 到 C；再由 3 到 2′加载时，应力 - 应变曲线由 C 到 B'，B' 和 B 重合。此后继续加载，则应力 - 应变曲线并不沿 CB' 曲线的延长线（图中虚线所示），而且急剧转弯沿原先 AB 曲线的延长线，似乎材料"记忆"了原先的路径，这就是材料的记忆特性。

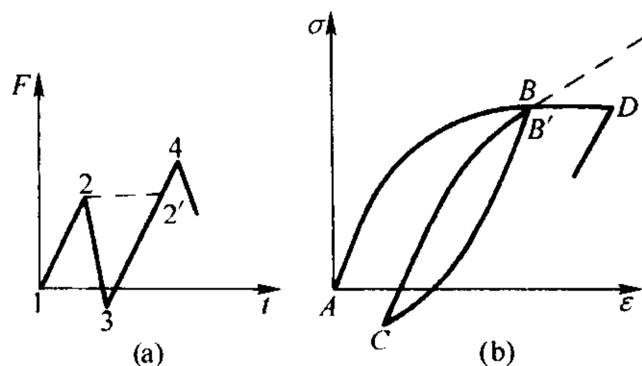


图 6.5-190 材料的记忆特性

5.4.4 载荷顺序效应

缺口零件在拉伸载荷作用下，缺口根部应力集中处材料发生屈服。卸载后因处于弹性状态的材料要恢复原来的状态，而已塑性变形的材料阻止这种恢复行为，故两者相互挤压，使缺口根部产生残余压应力。如大载荷环后面紧接着出现小载荷环，则该小载荷环引起的应力将叠加在这个残余应力之上，因此该小载荷环造成的损伤受到前面大载荷环的影响，而且这种影响往往是很大的。图 6.5-191 所示的两种载荷 - 时间历程，除第一载荷环以外，二者都相同，只是第一个大载荷环的过载方向不同。图 6.5-191a 所示的大载荷环以压缩载荷结束，应力集中处产生残余拉应力（$+\sigma_m$）。图 6.5-191b 所示的大载荷环以拉伸载荷结束，应力集中处产生残余压应力（$-\sigma_m$）。由于两种载荷 - 时间历程所产生的残余应力不同，所以滞回线的形状不同，即载荷顺序对局部应力 - 应变是有影响的。

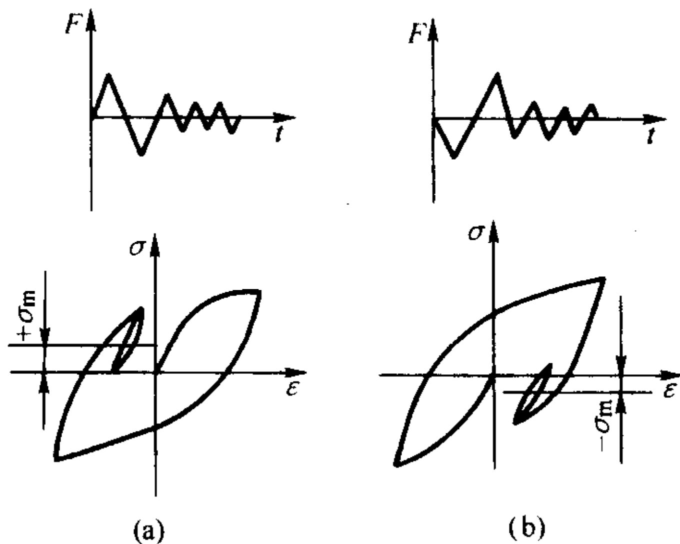


图 6.5-191 载荷顺序对滞回线影响

5.4.5 滞回线方程

局部应力 - 应变法认为，在疲劳强度问题中，材料的本构关系应由循环应力 - 应变曲线确定。材料的滞回线形状是通过循环应力 - 应变曲线来描述的。因此，循环 $\sigma-\varepsilon$ 曲线在局部应力 - 应变法中具有特殊重要的位置。由式（6.5-27）给出循环应力 - 应变曲线用幅度表达的方程式

$$\frac{\Delta\varepsilon}{2}=\frac{\Delta\sigma}{2E}+\left(\frac{\Delta\sigma}{2K'}\right)^{\frac{1}{n'}}$$

对于具有玛辛特性的材料，若使坐标原点与各应力水平下的滞回线最低点相重合，则滞回线的最高点的连线，与其上行段迹线相吻合（图 6.5-189b）。许多试验表明，多数金属材料的滞回线，可以用放大一倍后的循环 $\sigma-\varepsilon$ 曲线来近似描述。这样，就可得出下面的滞回线方程式，即

加载时

$$\frac{\varepsilon-\varepsilon_r}{2}=\frac{\sigma-\sigma_r}{2E}+\left(\frac{\sigma-\sigma_r}{2K'}\right)^{\frac{1}{n'}} \tag{6.5-41}$$

卸载时

$$\frac{\varepsilon_r-\varepsilon}{2}=\frac{\sigma_r-\sigma}{2E}+\left(\frac{\sigma_r-\sigma}{2K'}\right)^{\frac{1}{n'}} \tag{6.5-42}$$

式中，ε_r、σ_r 为滞回线顶点的坐标。

5.4.6 诺伯法

确定局部应力 - 应变的方法有：电阻应变计测定法、光弹性法、脆性涂层法和云纹法等实验方法，以及用有限元法求数值解。弹塑性有限元法是计算局部应力 - 应变的较精确的方法，但由于计算工作量大，工程上倾向于采用简单的近似方法。例如诺伯法、线性应变法、修正的斯托威尔法和莫尔斯基等效能量法。其中，应用最多的是诺伯法。

诺伯提出的在弹塑性状态下的通用公式

$$\alpha_\sigma^2=K'_\sigma K'_\varepsilon \tag{6.5-43}$$

式中，α_σ 为理论应力集中系数；K'_σ 为真实应力集中系数，$K'_\sigma=\sigma/s$；K'_ε 为真实应变集中系数，$K'_\varepsilon=\varepsilon/e$；$s$ 为缺口件的名义应力；e 为缺口件的名义应变；σ 为缺口件的真实应力；ε 为缺口件的真实应变。

通过式（6.5-43），就可以简单地把局部应力 - 应变与名义应力 - 应变联系起来。式（6.5-43）可写成下面形式，即

$$\sigma\varepsilon=\alpha_\sigma^2 se$$

一般情况下，名义应力和名义应变均在弹性范围内，即有 $s=Ee$。故有

$$\sigma\varepsilon=\frac{(\alpha_\sigma s)^2}{E} \tag{6.5-44}$$

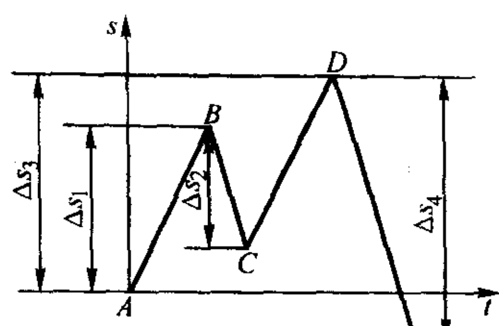
当名义应力确定以后，$\sigma\varepsilon=(\alpha_\sigma s)^2/E$ 是个常数，称为诺伯常数。于是式（6.5-44）可以写成 $\sigma\varepsilon=C$。这是一个双曲线方程，也称为诺伯双曲线。

如果已知 α_σ、s 和 E，再结合材料的 $\sigma-\varepsilon$ 曲线，就可以算出相应的局部应力和应变，如图 6.5-192 所示。将式（6.5-44）改写成幅度形式

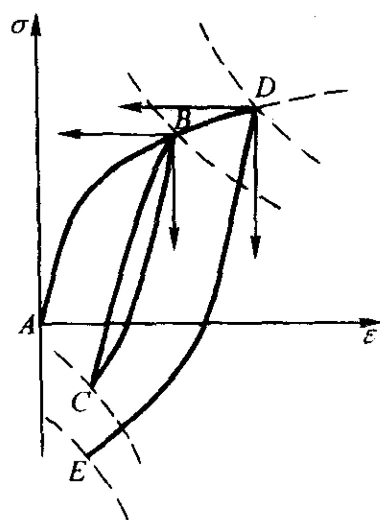
$$\Delta\sigma\Delta\varepsilon=\frac{\alpha_\sigma^2\ (\Delta s)^2}{E} \tag{6.5-45}$$

根据所给的载荷谱，名义应力幅度 Δs 是知道的，联立解式（6.5-27）和式（6.5-45），就可以求出 $\Delta\sigma$ 和 $\Delta\varepsilon$，加上坐标原点的应力和应变值，就是该点的局部真实应力和真实应变值。

例如，图 6.5-192a 是用名义应力表示的加载历程，图 6.5-192b 表示用诺伯法得到的零件危险点的局部应力 - 应变的情况。具体确定方法如下。



(a) 名义应力历程



(b) 局部应力-应变的确定

图 6.5-192　诺伯法确定局部应力 - 应变

1) B 点的确定　以 A 点作为坐标原点，画出循环 $\sigma-\varepsilon$ 曲线，并用 AB 间的名义应力幅度 Δs_1 画出 $\Delta\sigma\Delta\varepsilon=(\alpha_\sigma\Delta s_1)^2/E$ 双曲线，这两条曲线的交点 B 的纵坐标和横坐标，就是加载到 B 点时的局部应力和局部应变值。

2) C 点的确定　以 B 点作为坐标原点，向下画出滞回线（两倍于循环 $\sigma-\varepsilon$ 曲线），并用 BC 间的名义应力幅度 Δs_2 画出 $\Delta\sigma\Delta\varepsilon=(\alpha_\sigma\Delta s_2)^2/E$ 双曲线，这两条曲线的交点 C 的纵坐标和横坐标，即为从 B 点到 C 点的局部应力和应变幅度，在卸载时为负。加上 B 点的局部应力和应变值后，就得到加载到 C 点时的局部应力和应变值。

3) D 点的确定　从 C 点加载超过 B 点时要考虑"记忆特性"，即从 C 点到 D 点可以看作从 A 点直接加载到 D 点，故要以 A 点为坐标原点画出循环 $\sigma-\varepsilon$ 曲线，并画出 $\Delta\sigma\Delta\varepsilon=(\alpha_\sigma\Delta s_3)^2/E$ 双曲线，两条曲线的交点 D 的纵坐标和横坐标，即为加载到 D 点时的局部应力和应变值。

4) E 点的确定　以 D 点作为坐标原点，向下画出滞回线，并画出 $\Delta\sigma\Delta\varepsilon=(\alpha_\sigma\Delta s_4)^2/E$ 双曲线，由这两条曲线的交点 E 的纵坐标和横坐标，得到从 D 点到 E 点的局部应力和应变幅度，在卸载时为负。加上 D 点的局部应力和应变值后，就得到加载到 E 点时的局部应力和应变值。

按这个步骤对名义应力谱编制程序，在计算机上进行计算。

诺伯公式高估了局部应力和应变。因此，把公式中的理论应力集中系数 α_σ 改为有效应力集中系数 K_σ，得诺伯修正公式

$$\Delta\sigma\Delta\varepsilon=\frac{K_\sigma^2\ (\Delta s)^2}{E} \tag{6.5-46}$$

5.5　裂纹形成寿命的估算

5.5.1　损伤计算

局部应力 - 应变法计算损伤的出发点是应变 - 寿命关系式（6.5-29），即

$$\frac{\Delta\varepsilon}{2}=\frac{\Delta\varepsilon_e}{2}+\frac{\Delta\varepsilon_p}{2}=\frac{\sigma'_f}{E}\ (2N)^b+\varepsilon'_f\ (2N)^c$$

或分开写成

$$\frac{\Delta\varepsilon_e}{2}=\frac{\sigma'_f}{E}\ (2N)^b \tag{6.5-47}$$

$$\frac{\Delta\varepsilon_{p^*}}{2}=\varepsilon'_f\ (2N)^c \tag{6.5-48}$$

$\varepsilon-N$ 曲线是在对称循环条件下得出的。对于复杂载荷 - 时间历程作用下的疲劳问题，平均应力的存在是不可避免的，需要对上式进行修正。

当材料处于弹性范围时，平均应力对疲劳寿命的影响很大。而当材料出现塑性变形后，由于平均应力的松弛效应，其影响就大大减弱了。所以通常只对 $\varepsilon-N$ 曲线的弹性部分，即式（6.5-47）予以修正，一般应用的修正公式为

$$\sigma_r=\sigma_a\frac{\sigma'_f}{\sigma'_f-\sigma_m} \tag{6.5-49}$$

式中，σ_a 为应力幅；σ_m 为平均应力；σ_r 为等效应力幅。

修正后的应变 - 寿命关系为

$$\frac{\Delta\varepsilon_e}{2}=\frac{\sigma'_f-\sigma_m}{E}\ (2N)^b \tag{6.5-50}$$

$$\frac{\Delta\varepsilon_p}{2}=\varepsilon'_f\ (2N)^c \tag{6.5-51}$$

根据上述的寿命关系式，即式（6.5-47）、式（6.5-48）和式（6.5-50），采用不同的损伤参量，可以得到不同的损伤公式。局部应力 - 应变法中常用的损伤公式有以下几种。

1) 兰德格拉夫损伤公式　R.W. 兰德格拉夫认为，损伤由 $\Delta\varepsilon_p$ 与 $\Delta\varepsilon_e$ 的比值来控制。由式（6.5-47）和式（6.5-48）可推导出每个局部应变为 $\Delta\varepsilon$（$=\varepsilon_p+\varepsilon_e$）的应变循环造成的损伤为

$$\frac{1}{N}=2\left(\frac{\sigma'_f}{E\varepsilon'_f}\times\frac{\Delta\varepsilon_p}{\Delta\varepsilon_e}\right)^{\frac{1}{(b-c)}} \tag{6.5-52}$$

计入平均应力的影响，修正后的损伤公式为

$$\frac{1}{N}=2\left(\frac{\sigma'_f}{E\varepsilon'_f}\times\frac{\Delta\varepsilon_p}{\Delta\varepsilon_e}\times\frac{\sigma'_f}{\sigma'_f-\sigma_m}\right)^{1/(b-c)} \tag{6.5-53}$$

2) 道林损伤公式　N.E. 道林等人认为，以过渡疲劳寿命 N_T 为界，当 $\varepsilon_p>\varepsilon_e$ 时，应该以塑性应变分量为损伤参量，此时损伤公式为

$$\frac{1}{N}=2\left(\frac{\varepsilon'_f}{\varepsilon_p}\right)^{\frac{1}{c}} \tag{6.5-54}$$

当 $\varepsilon_p<\varepsilon_e$ 时，应该以弹性应变分量为损伤参量，损伤公式为

$$\frac{1}{N}=2\left(\frac{\sigma'_f}{E\varepsilon_e}\right)^{\frac{1}{b}} \tag{6.5-55}$$

若考虑平均应力的影响进行修正，则有

$$\frac{1}{N}=2\left(\frac{\sigma'_f-\sigma_m}{E\varepsilon_e}\right)^{\frac{1}{b}} \tag{6.5-56}$$

3) 史密斯损伤公式　K.N. 史密斯等人为反映平均应力的影响，对试验结果进行分析，提出用 $\sigma_{max}\Delta\varepsilon$ 来计算损伤，并推导出损伤公式

$$\sigma_{max}\Delta\varepsilon=\frac{2\sigma'_f}{E}\ (2N)^{2b}+2\sigma'_f\varepsilon'_f\ (2N)^{b+c} \tag{6.5-57}$$

该方程要用数值方法求解。

根据不同的 $\Delta\varepsilon_p/\Delta\varepsilon_e$ 比值，选用相应的损伤计算式。

5.5.2　估算裂纹形成寿命步骤

用局部应力 - 应变法估算裂纹形成寿命的步骤如下。

1) 把载荷谱、材料性能常数和应力集中系数作为输入计算机的信息。

2) 对载荷 - 时间历程进行循环计数。

3) 根据载荷 - 时间历程确定名义应力和应变 - 时间历程。

4) 根据选定的损伤公式，按循环计数的结果计算每一个载荷循环造成的损伤。

5) 对损伤进行累积计算，即根据累积损伤公式算出裂

纹形成寿命。

5.5.3 算例

在本例中，采用雨流法计数，用诺伯公式进行局部 $\sigma-\varepsilon$ 分析，用道林公式计算损伤。具体步骤如下。

首先将载荷－时间历程化为计算点上的名义应力－时间历程（见图 6.5-193a），并进行雨流计数，得到 1－4－7、2－3－2′和 5－6－5′三个循环。然后根据材料的 $\sigma-\varepsilon$ 曲线（滞回线）和零件的有效应力集中系数 K_σ，用诺伯法确定局部应力－应变响应。

循环 $\sigma-\varepsilon$ 曲线的方程为

$$\frac{\Delta\varepsilon}{2}=\frac{\Delta\sigma}{2E}+\left(\frac{\Delta\sigma}{2K'}\right)^{\frac{1}{n'}} \tag{a}$$

根据倍增原理，上升段的滞回线方程为

$$\frac{\varepsilon-\varepsilon_r}{2}=\frac{\sigma-\sigma_r}{2E}+\left(\frac{\sigma-\sigma_r}{2K'}\right)^{\frac{1}{n'}} \tag{b}$$

下降段的滞回线方程为

$$\frac{\varepsilon_r-\varepsilon}{2}=\frac{\sigma_r-\sigma}{2E}+\left(\frac{\sigma_r-\sigma}{2K'}\right)^{\frac{1}{n'}} \tag{c}$$

式中，σ、ε 为局部应力、应变的流动值；σ_r、ε_r 为前一峰值点的局部应力、应变值。

本例的材料是汽车用热轧低碳钢，其化学成分为：$w_C=0.23\%$；$w_{Mn}=1.57\%$；$w_P=0.016\%$；$w_S=0.022\%$；$w_{Si}=0.01\%$；$w_{Cu}=0.22\%$。其强度极限 $\sigma_b=540\sim565$ MPa，屈服强度 $\sigma_s=315\sim325$ MPa，截面缩减率 $\psi=64\%\sim69\%$，弹性模量 $E=192$ GPa，$n'=0.193$，$K'=1\ 125.9$ MPa。

应用诺伯公式（6.5-46）：

$$\Delta\sigma\Delta\varepsilon=\frac{K_\sigma^2\ (\Delta s)^2}{E} \tag{d}$$

根据所计算的危险点处的几何形状和材料，查应力集中系数图得 $K_\sigma=2.60$。

根据图 6.5-193a 所示的名义应力－时间历程，即可逐个反复地进行局部应力－应变分析。

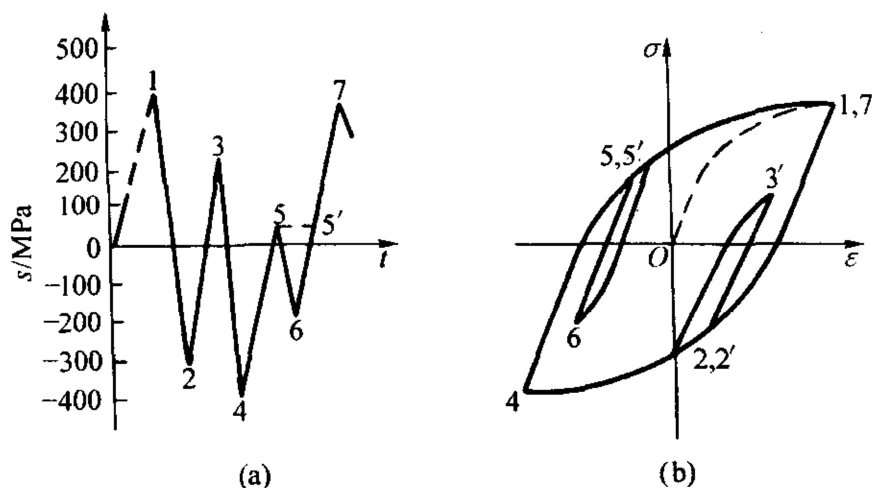


图 6.5-193　名义应力－时间历程及应力－应变响应

1）从 0－1 加载时，由于是从零开始，循环 $\sigma-\varepsilon$ 方程用式（6.5-26）

$$\varepsilon_a=\varepsilon_e+\varepsilon_p=\frac{\sigma_a}{E}+\left(\frac{\sigma_a}{K'}\right)^{\frac{1}{n'}} \tag{e}$$

再与诺伯公式（d）联立求解。将 $E=192$ GPa、$K'=1\ 125.9$ MPa、$n'=0.193$、$K_\sigma=2.60$ 代入，有

$$\left.\begin{aligned}\Delta\varepsilon&=\frac{\Delta\sigma}{192\ 000}+\left(\frac{\Delta\sigma}{1\ 125.9}\right)^{\frac{1}{0.193}}\\ \Delta\sigma\Delta\varepsilon&=\frac{2.6^2\times\Delta s_{01}^2}{192\ 000}\end{aligned}\right\}$$

此时，$\Delta s_{01}=395.5$ MPa，于是 $\Delta\sigma\Delta\varepsilon=5.5$。解联立方程得

$$\Delta\sigma=458.3\ \text{MPa},\quad \Delta\varepsilon=0.012$$

即 1 点的局部应力和应变为

$$\sigma=458.3\ \text{MPa},\quad \varepsilon=0.012$$

2）从 1－2 卸载时，根据卸载滞回线计算。将有关数据代入式（a）和式（d），有

$$\left.\begin{aligned}\frac{\Delta\varepsilon}{2}&=\frac{\Delta\sigma}{2\times192\ 000}+\left(\frac{\Delta\sigma}{2\times1\ 125.9}\right)^{\frac{1}{0.193}}\\ \Delta\sigma\Delta\varepsilon&=\frac{2.6^2\times\Delta s_{12}^2}{192\ 000}\end{aligned}\right\}$$

此时，$\Delta s_{12}=699.0$ MPa，于是 $\Delta\sigma\cdot\Delta\varepsilon=17.2$。解联立方程得

$$\Delta\sigma=870\ \text{MPa},\quad \Delta\varepsilon=0.0198$$

2 点的局部应力和应变为

$$\sigma=458.3-870=-411.7\ \text{MPa},$$
$$\varepsilon=0.012-0.019\ 8=-0.007\ 8$$

3）从 2－3 加载时，根据加载滞回线计算

$$\left.\begin{aligned}\frac{\Delta\varepsilon}{2}&=\frac{\Delta\sigma}{2\times192\ 000}+\left(\frac{\Delta\sigma}{2\times1\ 125.9}\right)^{\frac{1}{0.193}}\\ \Delta\sigma\Delta\varepsilon&=\frac{2.6^2\times\Delta s_{23}^2}{192\ 000}\end{aligned}\right\}$$

此时，$\Delta s_{23}=521.1$ MPa，于是 $\Delta\sigma\Delta\varepsilon=9.56$。解联立方程得

$$\Delta\sigma=780\ \text{MPa},\quad \Delta\varepsilon=0.012\ 2$$

3 点的局部应力和应变为

$$\sigma=-411.7+780=368.3\ \text{MPa},$$
$$\varepsilon=-0.007\ 8+0.012\ 2=0.004\ 4$$

4）在 3－4 的卸载过程中，由于从 3 卸载到 2′时，形成了一个封闭的应力－应变滞回线，所以根据材料的记忆特性，计算 4 点的应力和应变时，应根据从 1 点出发的滞回线，并取应力幅度 Δs_{14} 进行计算

$$\left.\begin{aligned}\frac{\Delta\varepsilon}{2}&=\frac{\Delta\sigma}{2\times192\ 000}+\left(\frac{\Delta\sigma}{2\times1\ 125.9}\right)^{\frac{1}{0.193}}\\ \Delta\sigma\Delta\varepsilon&=\frac{2.6^2\times\Delta s_{14}^2}{192\ 000}\end{aligned}\right\}$$

此时，$\Delta s_{14}=790.7$ MPa，于是 $\Delta\sigma\Delta\varepsilon=22.0$。解联立方程得

$$\Delta\sigma=910\ \text{MPa},\quad \Delta\varepsilon=0.024$$

4 点的局部应力和应变为

$$\sigma=458.3-910=-451.7\ \text{MPa}$$
$$\varepsilon=0.012-0.024=-0.012$$

5）从 4－5 加载时，根据加载滞回线计算

$$\left.\begin{aligned}\frac{\Delta\varepsilon}{2}&=\frac{\Delta\sigma}{2\times192\ 000}+\left(\frac{\Delta\sigma}{2\times1\ 125.9}\right)^{\frac{1}{0.193}}\\ \Delta\sigma\Delta\varepsilon&=\frac{2.6^2\times\Delta s_{45}^2}{192\ 000}\end{aligned}\right\}$$

此时，$\Delta s_{45}=434.1$ MPa，于是 $\Delta\sigma\Delta\varepsilon=6.6$。解联立方程得

$$\Delta\sigma=721\ \text{MPa},\quad \Delta\varepsilon=0.009\ 2$$

5 点的局部应力和应变为

$$\sigma=-451.7+721=269.3\ \text{MPa}$$
$$\varepsilon=-0.012+0.009\ 2=-0.002\ 8$$

6）从 5－6 加载时，根据卸载滞回线计算

$$\left.\begin{aligned}\frac{\Delta\varepsilon}{2}&=\frac{\Delta\sigma}{2\times192\ 000}+\left(\frac{\Delta\sigma}{2\times1\ 125.9}\right)^{\frac{1}{0.193}}\\ \Delta\sigma\Delta\varepsilon&=\frac{2.6^2\times\Delta s_{56}^2}{192\ 000}\end{aligned}\right\}$$

此时，$\Delta s_{56}=239.9$ MPa，得 $\Delta\sigma\Delta\varepsilon=2.0$。解联立方程得

$$\Delta\sigma=531\ \text{MPa},\quad \Delta\varepsilon=0.003\ 8$$

6 点的局部应力和应变为

$$\sigma=269.3-531=-261.7\ \text{MPa}$$
$$\varepsilon=-0.002\ 8-0.003\ 8=-0.006\ 6$$

7）从 6－7 加载时，根据图 6.5-193b 所示，7 点的应力和应变值与 1 点相同。得局部应力和应变为

$\sigma = 458.3\ \text{MPa},\ \varepsilon = 0.012$

有了局部应力－应变响应，就可以进行损伤计算。损伤是根据每一应力－应变循环的幅值和均值，用道林公式计算的。现将上面分析得到的三个应力－应变循环2－3－2′、5－6－5′和1－4－7中的应力幅值σ_a、应变幅值ε_a、平均应力σ_m、平均应变ε_m及弹性应变分量ε_e、塑性应变分量ε_p列入表6.5-37。

表6.5-37　三个应力－应变循环的应力和应变值

应力循环	σ_a/MPa	ε_a	σ_m/MPa	ε_m	ε_e	ε_p
2－3－2′	390.0	0.006 1	－21.7	－0.001 7	0.002 0	0.004 1
5－6－5′	265.5	0.002 0	3.8	－0.004 4	0.001 4	0.000 6
1－4－7	455.0	0.012 0	3.3	0	0.002 4	0.009 6

下面进行损伤计算。

对于2－3－2′循环，由于$\varepsilon_p > \varepsilon_e$，故用$\varepsilon_p$计算损伤。由式（6.5-54）有

$$D_1 = \frac{1}{N} = 2\left(\frac{\varepsilon'_f}{\varepsilon_p}\right)^{\frac{1}{c}}$$

本例中，$\varepsilon'_f = 0.26$，$c = -0.47$，所以

$$D_1 = 2\left(\frac{0.26}{0.004\ 1}\right)^{-\frac{1}{0.47}} = 2.93\times10^{-4}$$

对于5－6－5′循环，由于$\varepsilon_e > \varepsilon_p$，故用$\varepsilon_e$计算损伤。式（6.5-56）中的$E\varepsilon_e$以总应力幅$\sigma_a$代替，有

$$D_2 = \frac{1}{N} = 2\left(\frac{\sigma'_f - \sigma_m}{\sigma'_a}\right)^{\frac{1}{b}}$$

本例中，$\sigma'_f = 935.9$ MPa，$b = -0.095$，$\sigma_m = 3.8$ MPa，$\sigma_a = 265.5$ MPa，$\varepsilon_e = 0.001\ 4$，$E = 192$ GPa。于是

$$D_2 = 2\left(\frac{935.9-3.8}{265.5}\right)^{-\frac{1}{0.095}} = 3.63\times10^{-6}$$

对于1－4－7循环，由于$\varepsilon_p > \varepsilon_e$，故用$\varepsilon_p$计算损伤

$$D_3 = 2\left(\frac{\varepsilon'_f}{\varepsilon_p}\right)^{\frac{1}{c}} = 2\left(\frac{0.26}{0.009\ 6}\right)^{-\frac{1}{0.47}} = 1.79\times10^{-3}$$

根据迈因纳定律求疲劳累积损伤，得

$$\begin{aligned} D &= \sum_i D_i = D_1 + D_2 + D_3 \\ &= 2.93\times10^{-4} + 3.63\times10^{-6} + 1.79\times10^{-3} \\ &= 2.087\times10^{-3} \end{aligned}$$

所以疲劳破坏时载荷循环块数（即载荷－时间历程1－7的反向次数）B为

$$B = \frac{1}{\sum_i D_i} = \frac{1}{2.087\times10^{-3}} = 479.2$$

若每个载荷块经历的时间为h_0，则零件的疲劳寿命为

$$h = Bh_0$$

上述计算均可由计算机完成。

局部应力－应变法是在应变分析和低周疲劳基础上发展起来的一种疲劳寿命估算方法。因此它特别适用于低周疲劳。将其应用于高周疲劳时，由于它没有考虑高周疲劳中表面状态和尺寸的影响因素，因此计算结果误差大，需采用修正的方法以减小误差。

6　接触疲劳强度设计

零件在循环接触应力作用下，产生局部永久性累积损伤，经过一定的循环次数后，接触表面发生麻点、浅层或深层剥落的过程，称为接触疲劳。齿轮、滚动轴承和凸轮是典型的接触疲劳失效零件。

6.1　接触疲劳失效机理

以圆柱体滚子在弹性体平面上滚动为例，弹性体内切应力τ_{yz}的变化如图6.5-194所示。图6.5-194a表示滚子在位置1，图6.5-194b表示滚子在位置2，图6.5-194c表示滚子在位置3。有剖面线的滚子表示滚子的所在位置，没有剖面线的滚子表示在该瞬间滚子不在这位置上。

图6.5-194a表示当滚子在位置1时，在位置1的正下方$\tau_{yz}=0$；但在左边位置2的下方（例如，位置2在位置1左边$0.85b$处），有切应力$\tau_{yz}=\tau_{zy}$的最大值，等于$0.256\sigma_{zmax}$（图6.5-195），并以τ_0表示。此时，位置3下方的切应力很小，可以不考虑。

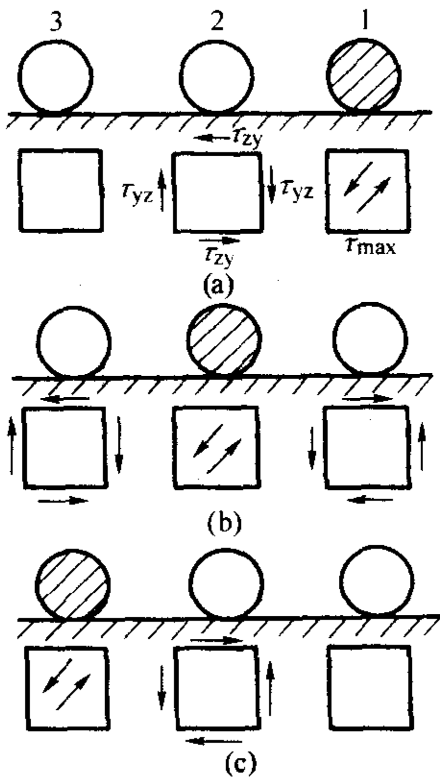


图6.5-194　圆柱体滚子在平面上滚动时的切应力变化

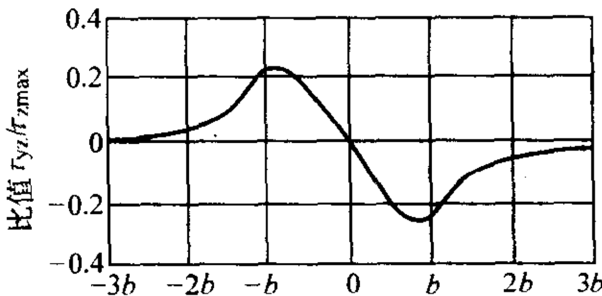


图6.5-195　在表面下0.5b处切应力τ_{yz}的变化曲线

图6.5-194b表示滚子在位置2时弹性体内切应力τ_{yz}的情况。这时，在位置2的正下方，$\tau_{yz}=0$；在左边位置3及右边位置1的下方，$\tau_{yz}=\tau_{zy}=\tau_0$，但切应力的方向相反。

图6.5-194c表示滚子在位置3时弹性体内切应力τ_{yz}的情况。这时，在位置3的正下方，$\tau_{yz}=0$；在右边位置2的下方，$\tau_{yz}=\tau_{zy}=\tau_0$；在位置1的下方，切应力很小，可以不考虑。

因此，弹性体内最危险的切应力是离表面$0.5b$处的τ_{yz}，其最大值为$0.256\sigma_{zmax}=\tau_0$，即应力幅为$0.256\sigma_{zmax}$的对称循环切应力。

在切应力τ_0的循环作用下，接触物体表面下形成平行于表面的裂纹，裂纹在滚动方向平行于表面扩展，再延伸到表面使之剥落。这种破坏的裂纹扩展速度较慢，断口光滑。

6.2　接触应力

图6.5-196表示两物体相接触时采用的坐标系，未加载

时于 O 点接触。假设：①两物体为完全弹性体，并且各向同性；②作用于物体上的载荷仅产生弹性变形并遵循虎克定律；③两物体的接触区面积比物体的总面积小很多；④压力垂直于接触表面，即接触区中的摩擦力略去不计；⑤表面光滑，无承载油膜。

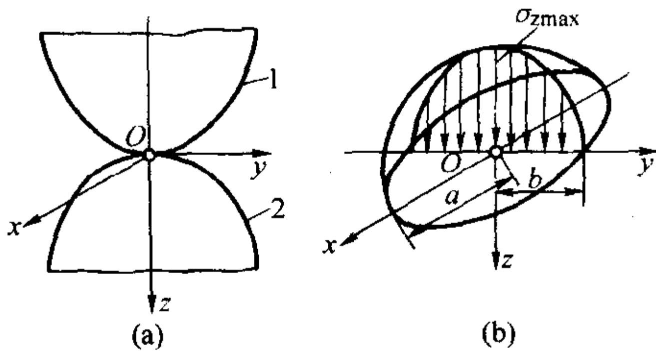


图 6.5-196　两物体的接触

在弹性体接触问题中，原为点接触的两物体受压力后，接触面的一般形状为椭圆，其长半轴为 a，短半轴为 b。取椭圆的中心为原点 O，压力分布为半椭球形，在 O 点的最大名义接触应力以 σ_{zmax} 表示（图 6.5-196b）。令半椭球体的体积等于总压力 F，即得

$$\sigma_{zmax}=\frac{3F}{2\pi ab} \tag{6.5-58}$$

椭圆面积上的平均应力 σ_{zm} 为

$$\sigma_{zm}=\frac{F}{\pi ab}$$

由此可以看出，σ_{zmax} 为 σ_{zm} 的 1.5 倍。由弹性力学可得椭圆的长半轴 a 和短半轴 b 分别为

$$\left.\begin{aligned}a&=m\left[\frac{3\pi F\ (k_1+k_2)}{4\ (A+B)}\right]^{\frac{1}{3}}\\ b&=n\left[\frac{3\pi F\ (k_1+k_2)}{4\ (A+B)}\right]^{\frac{1}{3}}\end{aligned}\right\} \tag{6.5-59}$$

式中的常数 A 和 B 都是正值。

设上边物体 1 的表面在接触点处的主曲率半径为 R_1 及 R'_1，下边物体 2 的表面在接触点处的主曲率半径为 R_2 及 R'_2，R_1 及 R_2 两曲率半径所在平面的夹角为 ψ，则 A 和 B 决定于以下两个方程，即

$$\left.\begin{aligned}A+B&=\frac{1}{2}\left(\frac{1}{R_1}+\frac{1}{R'_1}+\frac{1}{R_2}+\frac{1}{R'_2}\right)\\ B-A&=\frac{1}{2}\left[\left(\frac{1}{R_1}-\frac{1}{R'_1}\right)^2+\left(\frac{1}{R_2}-\frac{1}{R'_2}\right)^2+2\left(\frac{1}{R_1}-\frac{1}{R'_1}\right)\right.\\ &\quad\left.\times\left(\frac{1}{R_2}-\frac{1}{R'_2}\right)\cos2\psi\right]^{1/2}\end{aligned}\right\} \tag{6.5-60}$$

m 及 n 是与比值（$B-A$）/（$A+B$）有关的系数。引用符号

$$\cos\theta=\frac{B-A}{A+B}$$

则 m 及 n 与 θ 的关系可列于表 6.5-38。

表 6.5-38　m 及 n 的数值

θ	30°	40°	50°	60°	70°	80°	90°
m	2.73	2.14	1.75	1.49	1.28	1.13	1.00
n	0.49	0.57	0.64	0.72	0.80	0.89	1.00

$$k_1=\frac{1-\nu_1^2}{\pi E_1},\quad k_2=\frac{1-\nu_2^2}{\pi E_2} \tag{6.5-61}$$

式中，ν_1 和 ν_2 为物体 1 和物体 2 的泊松比；E_1 和 E_2 为物体 1 和物体 2 的弹性模量。

比值 a/b 越大，接触面的椭圆越是长而窄。当 a/b 趋于无限大时，就得到两个轴线平行的圆柱体相接触的情况。这时，接触面是宽度为 $2b$ 的狭矩形，而名义接触应力沿接触面宽度方向按半椭圆分布。令接触面的单位长度上的接触力为 F'，则

$$F'=\frac{\pi b\sigma_{zmax}}{2}$$

从而得到最大名义接触应力

$$\sigma_{zmax}=\frac{2F'}{\pi b} \tag{6.5-62}$$

它等于平均应力 $F'/(2b)$ 的 $4/\pi$ 倍。对局部应变进行分析，可得

$$b=\sqrt{\frac{4F'\ (k_1+k_2)\ R_1R_2}{R_1+R_2}} \tag{6.5-63}$$

及

$$\sigma_{zmax}=\sqrt{\frac{F'\ (R_1+R_2)}{\pi^2\ (k_1+k_2)\ R_1R_2}} \tag{6.5-64}$$

当两物体的材料相同，即 $E_1=E_2=E$ 及 $\nu_1=\nu_2=\nu=0.3$ 时，得工程上常用的公式，即

$$\left.\begin{aligned}b&=1.52\sqrt{\frac{F'R_1R_2}{E\ (R_1+R_2)}}\\ \sigma_{zmax}&=0.418\sqrt{\frac{F'E\ (R_1+R_2)}{R_1R_2}}\end{aligned}\right\} \tag{6.5-65}$$

对于圆柱体与平面相接触的情况，只需在上面的公式中，令 $R_1\to\infty$，对于圆柱体与圆柱座相接触的情况，只须在以上的公式中取 R_1 为负值。

用解析法可以推导出接触物体的切应力 τ_{yz} 的表达式。当轴线平行的两圆柱体相接触时，在接触区表面下 $0.5b$ 处的平面上，切应力 τ_{yz} 的变化如图 6.5-195 所示。图中横坐标由 $-b$ 到 b 的区域为接触面的宽度，O 点为接触区的中点，该点的压应力为最大应力 σ_{zmax}。切应力的最大值在表面以下 $0.5b$ 处，离中心点 O 的距离为 $0.85b$，且 $+0.85b$ 处的 τ_{yz} 与 $-0.85b$ 处的 τ_{yz} 方向相反，故此切应力为对称循环切应力。

6.3　接触疲劳强度计算

接触疲劳强度计算，也是以 $S-N$ 曲线为依据的。但接触疲劳的 $S-N$ 曲线与拉伸和弯曲疲劳的 $S-N$ 曲线不同。对材料接触疲劳试验的基本要求是，应尽可能地将被试验的材料做成滚子形零件（试样），并接近实际使用条件。这些条件包括：试样的加载形式，润滑油的选择与供油方法，材料的化学成分及组织状态，试样的形状和试样的表面加工特性等。

接触疲劳的 $S-N$ 曲线的纵坐标是最大名义接触应力 σ_{zmax}。σ_{zmax} 的计算不考虑应力集中和局部塑性变形后应力重新分配等因素，而按弹性理论的公式进行。在每个应力水平下，对成组试验法的数据进行统计，得到 σ_{zmax} 的均值，根据各个应力水平下的 σ_{zmax} 的均值画出 $S-N$ 曲线。

对于在每个应力水平下，用一个试样的接触疲劳常规试验法，试样的数目不得少于 12 个，其中在接触疲劳极限水平区段进行试验的试样数不得少于 3 个。

图 6.5-197 为 $w_C=0.34\%$ 碳钢、$w_C=0.10\%$ 碳钢和硬铝用机油润滑进行接触疲劳试验得到的 $S-N$ 曲线。图 6.5-198 为 14CrMnSiNi2Mo 钢经碳氮共渗和渗碳后淬火试验得的 $S-N$ 曲线。表 6.5-39 为某些材料的接触疲劳极限。

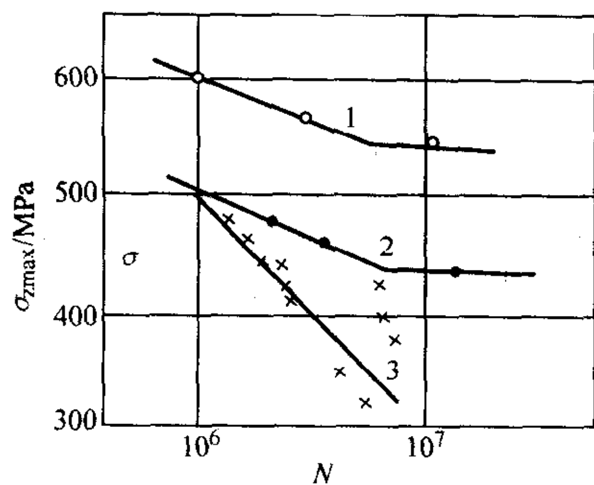


图 6.5-197 接触疲劳的 $S-N$ 曲线
1—$w_C=0.34\%$碳钢；2—$w_C=0.10\%$碳钢；
3—硬铝

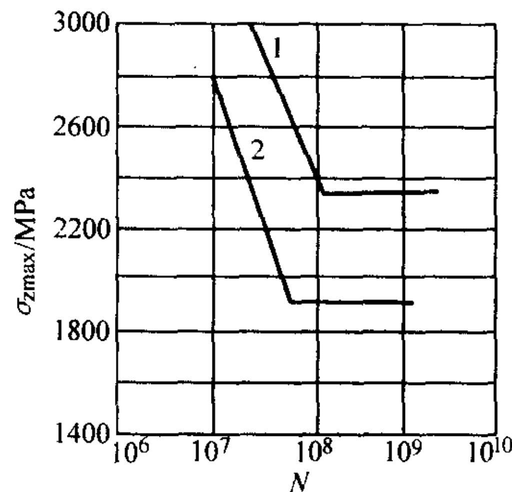


图 6.5-198 14CrMnSiNi2Mo 钢的接触疲劳 $S-N$ 曲线
1—碳氮共渗试样，渗层厚度为 0.66 mm；
2—渗碳试样，渗层厚度为 0.76 mm

接触疲劳极限的循环基数 N_0，以不产生大量扩展性点蚀为依据。对于低碳钢，$N_0=(2\sim4)\times10^6$；调质钢为 $(10\sim20)\times10^6$；铸铁为 $(2\sim6)\times10^6$；青铜与铜合金为 $(3\sim12)\times10^6$ 循环。当应力低于接触疲劳极限时，经过相当多循环后，也可能产生一些非扩展性的点蚀。

当由试验得出材料在某种具体接触情况下的 $S-N$ 曲线后，有关接触疲劳的强度计算和寿命计算，与高周疲劳计算相同。

表 6.5-39 某些材料的接触疲劳极限

材料及热处理	σ_b/MPa	σ_s/MPa	δ/%	硬度 HB		接触疲劳极限 σ_{zmax}/MPa
				试验前	试验后	
St42①	485	338	32	140	169	422
St50②	568	358	28	159	191	481
St60	650	342	25	187	218	530
St70	806	364	18	230	271	647
Si-Mn 钢，调质	797	521	23	229	266	706
Si-Mn 钢，调质	867	555	21	255	292	770
Cr-Mn 钢，调质	—	—	—	345	—	1 040
Cr-Mn 钢，调质	1 149	1 065	15	347	363	1 040
轴承钢，不淬火	989	928	18	310	322	559
氮化钢，淬火	1 903	—	—	555	594	1 687
轴承钢，淬火	1 981	—	—	573	592	1 726
轴承钢，淬火	2 010	—	—	629	642	1 912
灰铸铁	—	—	—	156	169	275
灰铸铁	—	—	—	234	251	422
特种铸铁	400	—	—	295	310	608
Stg38③	431	256	37	130	142	422
Stg52④	534	295	29	142	164	461
Stg60⑤	637	333	20	185	211	618
铝青铜	490	177	15	130	150	412
铝青铜	686～736	226	30	200	227	549

① 相当于甲类普通碳素钢 Q235A。
② 相当于 Q275。
③ 相近于 ZG230-450。
④ 相近于 ZG270-500。
⑤ 相近于 ZG310-570。

编写：王　雷（东北大学）
何雪浤（东北大学）

第6章　材料强度概率设计

1　强度概率设计原理

1.1　概率设计的特点与步骤

在进行材料强度设计时，有两种基本原理——根据概率理论进行设计，或根据确定性假设进行设计。由于对设计对象越来越多地提出了可靠性要求，因此要重新评价传统的基本设计原则，要应用概率论与数理统计知识。

简单地讲，传统的强度设计准则是要求强度大于应力，而概率设计准则是要求保证强度大于应力的概率不小于预期指标。这样看似简单的设计准则的改变，对设计过程本身有着深远的影响。

概率设计允许有一个小的失效概率，把强度与应力作为多值现象处理，并根据统计数字建立变量的模型和根据统计代数建立函数。设计变量的函数可作为随机变量之间的函数关系处理。每个随机变量均有一个明确的概率密度函数，函数模型是一个多元概率密度函数。

概率设计方法的优点包括：理论上的正确性，可靠性，经济性。

1.1.1　强度概率设计的基本特点

1）把应力和强度作为随机变量。

2）应用概率和统计方法进行分析、求解。

3）能定量地处理设计对象的失效概率和可靠度。

1.1.2　强度概率设计的步骤

1）提出设计任务，规定详细指标。

2）确定有关的设计变量及参数。

3）确定失效模式的判据。

4）得出应力公式。

5）确定每种失效模式下的应力分布。

6）确定强度计算公式。

7）确定每种失效模式下的强度分布。

8）确定每种致命失效模式下与应力公布和强度分布相关的可靠度。

9）确定零件的可靠度。

10）确定零件可靠度的置信度。

1.2　应力－强度分布干涉理论与可靠度的一般表达式

令 $f(S)$ 为应力分布的概率密度函数，$g(\delta)$ 为强度分布的概率密度函数，两者之间有干涉区存在，如图6.6-1所示。

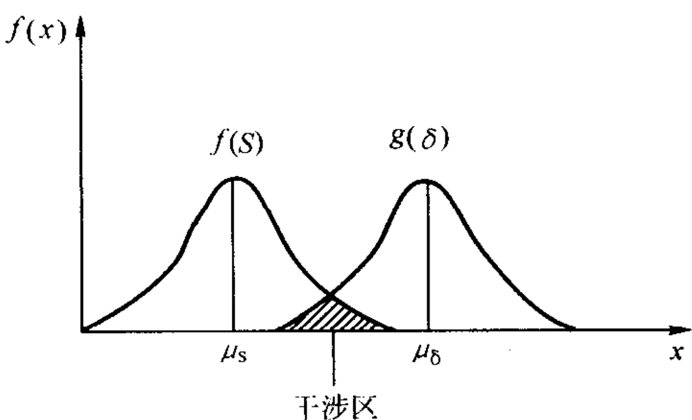


图6.6-1　应力－强度干涉原理

应力值位于宽度为 $\mathrm{d}S$ 的小区间内的概率为 $f(S_1)\,\mathrm{d}S$，强度 δ 大于应力 S_1 的概率为 $\int_{S_1}^{\infty} g(\delta)\,\mathrm{d}\delta$。应力出现在该小区间且材料不失效的概率为

$$\mathrm{d}R = f(S_1)\mathrm{d}S \times \int_{S_1}^{\infty} g(\delta)\mathrm{d}\delta$$

因而，材料在随机应力 S 作用下的可靠度为

$$R = \int_{-\infty}^{\infty} f(S)\left[\int_{S}^{\infty} g(\delta)\,\mathrm{d}\delta\right]\mathrm{d}S \tag{6.6-1}$$

或

$$R = \int_{-\infty}^{\infty} g(\delta)\left[\int_{-\infty}^{\delta} f(S)\,\mathrm{d}S\right]\,\mathrm{d}\delta \tag{6.6-2}$$

1.3　应力分布的确定

机械零件所受的工作应力 S 与其承受的载荷、温度、几何尺寸、物理特性、时间等参数有关，一般可表达为

$$S = f(L, T, G, p, t, m) \tag{6.6-3}$$

1.3.1　确定应力分布的步骤

1）确定零件的失效模式及其判据。

2）应力分析。

3）确定失效应力、应力修正系数及有关设计参数的分布。

4）确定可靠性计算用应力分布。

1.3.2　求解应力分布参数的代数方法

如果影响零件工作应力 S 的参数 $X_1 \sim (\mu_1, \sigma_1^2)$，$X_2 \sim (\mu_2, \sigma_2^2)$，$\cdots$，$X_n \sim (\mu_n, \sigma_n^2)$ 均为正态随机变量，则可以根据这些参数与应力的函数关系，把它们综合为仅含单一随机变量 Z 的应力函数 $S(Z) = f(X_1, X_2, \cdots, X_n)$，并确定其分布。

如果各随机变量的变异系数都小于0.1，即 $\sigma_i/\mu_i < 0.1$，且满足随机变量的多重性要求，则由中心极限定理可知，这个应力函数近似于正态分布。

综合过程是由两个随机变量开始，逐步进行。

表6.6-1为正态分布函数特征值计算公式。

表6.6-1　正态分布函数特征值计算公式

序号	函数形式	均值 μ_z	标准差 σ_z
1	$Z = c$	c	0
2	$Z = cx$	$c\mu_x$	$c\sigma_x$
3	$Z = cx + d$	$c\mu_x + d$	$c\sigma_x$
4	$Z = x \pm y$	$\mu_x \pm \mu_y$	$(\sigma_x^2 + \sigma_y^2 \pm 2\rho\sigma_x^2\sigma_y^2)^{1/2}$
5	$Z = xy$	$\mu_x\mu_y + \rho\sigma_x\sigma_y$	$[(\mu_x^2\sigma_y^2 + \mu_y^2\sigma_x^2 + \sigma_x^2\sigma_y^2)(1 + \rho^2)]^{1/2}$
6	$Z = x/y$	$\mu_x/\mu_y + \mu_x\sigma_y^2/\mu_y^3(\sigma_y/\mu_y - \rho\sigma_x/\mu_x)$	$\mu_x/\mu_y[\sigma_x^2/\mu_x^2 + \sigma_y^2/\mu_y^2 - 2\rho\sigma_x\sigma_y/(\mu_x\mu_y)]^{1/2}$
7	$Z = x^2$	μ_x^2 或 $\mu_x^2 + \sigma_x^2$	$(4\mu_x^2\sigma_x^2 + 2\sigma_x^4)^{1/2}$
8	$Z = x^3$	μ_x^3 或 $\mu_x^3 + 3\sigma_x^2\mu_x$	$3\sigma_x^6 + 8\sigma_x^4\mu_x^2 + 5\sigma_x^2\mu_x^4$
9	$Z = x^n$	μ_x^n	$n\mu_x^{n-1}\sigma_x$
10	$Z = x^{1/2}$	$(1/2)^{1/2}(4\mu_x^2 - 2\sigma_x^2)^{1/4}$	$(\mu_x - \frac{1}{2}\sqrt{4\mu_x^2 - 2\sigma_x^2})^{\frac{1}{2}}$
11	$Z = x^2 + y^2)^{1/2}$	$(\mu_x^2 + \mu_y^2)^{1/2}$	$\left(\frac{\mu_x^2\sigma_y^2 + \mu_y^2\sigma_x^2}{\mu_x^2 + \mu_y^2}\right)^{\frac{1}{2}}$

1.3.3 求解应力分布参数的矩方法

用矩法求随机变量 X 的函数 $f(X)$ 的均值及标准方差，是通过泰勒展开式来实现的。

1) 一维随机变量 设 $y=f(X)$ 为一维随机变量 X 的函数，X 的均值为 μ（已知）。将 $f(X)$ 在 $X=\mu$ 处展开

$$Y=f(x)=f(\mu)+(X-\mu)f'(\mu)+\frac{1}{2}(x-\mu)^2f(\mu)''+R \tag{6.6-4}$$

对上式取数学期望，有

$$E(y)=E[f(\mu)]+E[(X-\mu)f'(\mu)]+E[\frac{1}{2}(X-\mu)^2 f''(\mu)]+E[R]$$
$$\approx f(\mu)+\frac{1}{2}f''(\mu)\mathrm{var}(X)$$

$$E(y)\approx f(\mu)+\frac{1}{2}f''(\mu)\mathrm{var}(X) \tag{6.6-5}$$

对上式取方差，有

$$\mathrm{var}(y)=\mathrm{var}[f(\mu)]+\mathrm{var}[(X-\mu)f'(\mu)]+\mathrm{var}[R_1]$$
$$=\mathrm{var}[(X-\mu)][f'(\mu)]^2+\mathrm{var}[R_1]$$
$$=\mathrm{var}[X][f'(\mu)]^2 \tag{6.6-6}$$

2) 多维随机变量 设 $y=f(X)=F(X_1,X_2,\cdots,X_n)$，为相互独立的随机变量 $(X_1,X_2,\cdots,X_n)$ 的函数。在均值处展开

$$y=f(\mu_1,\mu_2,\cdots,\mu_n)+\sum_{i=1}^{n}\frac{\partial f(X)}{\partial X_i}\Big|_{X=\mu}\cdot(X_i-\mu_i)+\frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 f(X)}{\partial X_i\partial X_j}\Big|_{X=\mu}\cdot(X_i-\mu_i)(X_j-\mu_j)+R_n$$

$$E(y)\approx f(\mu_1,\mu_2,\cdots,\mu_n)+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2 f(X)}{\partial X_i^2}\Big|_{X=\mu}\cdot\mathrm{var}(X_i) \tag{6.6-7}$$

$$\mathrm{var}(y)\approx\sum_{i=1}^{n}\left\{\frac{\partial f(X)}{\partial X_i}\Big|_{X=\mu}\right\}^2\mathrm{var}(X_i)$$

1.4 随机变量函数的均值和标准差的近似计算

应力和强度常为一些随机变量的函数。已知各随机变量的均值和标准差，求其函数均值和标准差的近似方法如下。

1.4.1 泰勒展开法

设 n 维随机变量 x_1，x_2，…，x_n 的函数为

$$y=f(x_1,x_2,\cdots\cdots,x_n) \tag{6.6-8}$$

各随机变量的变异系数 $v_{x_i}=\dfrac{S_{x_i}}{\bar{x}_i}$ 很小，则函数的均值、函数的标准差

$$\bar{y}\approx f(\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_n)+\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2 y}{\partial x_i^2}\right)_0 S_{x_i}^2+\sum_{i=1}^{n-1}\times\sum_{j=i+1}^{n}\left(\frac{\partial^2 y}{\partial x_i\partial x_j}\right)_0\rho_{ij}S_{x_i}S_{x_j}$$
$$\approx f(\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_n)$$

$$S_y\approx\left[\sum_{i=1}^{n}\left(\frac{\partial y}{\partial x_i}\right)_0^2 S_{x_i}^2+2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\frac{\partial y}{\partial x_i}\right)_0\left(\frac{\partial y}{\partial x_j}\right)_0\rho_{ij}S_{x_i}S_{x_j}\right]^{\frac{1}{2}} \tag{6.6-9}$$

式中，脚标“0”为表示求导后自变量取均值；ρ_{ij} 为 x_i 与 x_j 的相关系数。

$$\rho_{ij}=\frac{E[x_i-\bar{x}_i(x_j-\bar{x}_j)]}{[E(x_i-\bar{x}_i)^2E(x_j-\bar{x}_j)^2]^{\frac{1}{2}}} \tag{6.6-10}$$

对 (x_i,x_j) 作 n 次独立观测的相关系数估计值。一般 $\rho_{ij}=-1\sim+1$。若 $x_i=ax_j+b$，$a\neq0$，则 x_i 与 x_j 线性相关。当 $a>0$，则 $\rho_{ij}=+1$；当 $a<0$，则 $\rho_{ij}=-1$。若 x_i 与 x_j 互相独立，则 $\rho_{ij}=0$。ρ_{ij} 的具体数值，就概念上判定为 +1、-1、0 常比较容易，而对非线性又不独立往往难以估计。在工程计算中，为了简单常假定为独立而取 $\rho_{ij}=0$，也有时折中取 $|\rho_{ij}|=0.5$。为避免由此产生的麻烦，有时可设法避开相关的麻烦。

1.4.2 变异系数法

对于单项式（没有加、减运算的式子）的函数，式 (6.6-8) 的具体形式为

$$y=a\prod_{i=1}^{n}x_i^{m_i}=ax_1^{m_1}x_2^{m_2}\cdots x_n^{m_n} \tag{6.6-11}$$

式中，a、m_i 为任意常数，可为正也可为负函数的均值。

$$\bar{y}=a\prod_{i=1}^{n}\bar{x}_i^{m_i}=a\bar{x}_1^{m_1}\bar{x}_2^{m_2}\cdots\cdots\bar{x}_n^{m_n} \tag{6.6-12}$$

函数的变异系数

$$v_y\approx\left(\sum_{i=1}^{n}m_i^2v_{x_i}^2+2\sum_{i=1}^{n-1}\times\sum_{j=i+1}^{n}m_im_j\rho_{ij}v_{x_i}v_{x_j}\right) \tag{6.6-13}$$

函数的标准差

$$S_y=\bar{y}v_y \tag{6.6-14}$$

1.4.3 基本函数法

本方法是将常用的函数作为基本函数，其均值和标准差的结论式列于表 6.6-2，应用时可直接搬用，对于较复杂的函数一般可化为这些基本函数的形式。

表 6.6-2 基本函数形式和近似结论式

序号	函数形式	均值 $\bar{y}$	标准差 s_y
1	$y=a$	a	0
2	$y=ax$	$a\bar{x}$	as_x
3	$y=a\pm x$	$a\pm\bar{x}$	s_x
4	$y=x^m$	$\bar{x}^m$	$\|m\|\bar{x}^{m-1}s_x$
5	$y=x_1\pm x_2$	$\bar{x}_1\pm\bar{x}_2$	$(s_{x_1}^2+s_{x_2}^2\pm2\rho_{12}s_{x_1}s_{x_2})^{\frac{1}{2}}$
6	$y=x_1x_2$	$\bar{x}_1\bar{x}_2+\rho_{12}s_{x_1}s_{x_2}$	$(\bar{x}_1^2s_{x_2}^2+\bar{x}_2^2s_{x_1}^2+2\rho_{12}\bar{x}_1\bar{x}_2s_{x_1}s_{x_2})^{\frac{1}{2}}$
7	$y=\dfrac{x_1}{x_2}$	$\dfrac{\bar{x}_1}{\bar{x}_2}+\dfrac{\bar{x}_1s_{x_2}}{\bar{x}_2^2}\left(\dfrac{s_{x_2}}{\bar{x}_2}-\dfrac{\rho_{12}s_{x_1}}{\bar{x}_1}\right)\approx\dfrac{\bar{x}_1}{\bar{x}_2}$	$\dfrac{\bar{x}_1}{\bar{x}_2}\left(\dfrac{s_{x_1}^2}{\bar{x}_1^2}+\dfrac{s_{x_2}^2}{\bar{x}_2^2}-\dfrac{2\rho_{12}s_{x_1}s_{x_2}}{\bar{x}_1\bar{x}_2}\right)^{\frac{1}{2}}$

三种求函数均值和标准差的近似法中，基本函数法只在简单情况才显得方便；变异系数法计算最简便，建议尽量采用；泰勒展开法可用于各式各样的函数形式。

例 1　某钢制拉杆，截面直径（$\bar{d}$，s_d）=（10，0.1）mm，杆长（$\bar{L}$，s_L）=（1 000，3.3）mm，受拉力（$\bar{F}$，s_F）=（40 000，2 000）N，弹性模量（$\bar{E}$，s_E）=（20 600，618）MPa，求拉杆受力后的伸长量。

解　由材料力学知，伸长量

$$\delta = \frac{4FL}{\pi d^2 E}$$

1）用泰勒展开法　用式（6.6-8）、式（6.6-9）各随机变量相互独立，$\rho_{ij}=0$，

$$\bar{\delta} = \frac{4\overline{FL}}{\pi \bar{d}^2 \bar{E}} = \frac{4\times 4\ 000\times 1\ 000}{\pi\times 10^2\times 20\ 600} = 24.72\ \text{mm}$$

$$s_\delta = [(\frac{\partial \delta}{\partial F})_0^2 s_F^2 + (\frac{\partial \delta}{\partial L})_0^2 s_L^2 + (\frac{\partial \delta}{\partial d})_0^2 s_d^2 + (\frac{\partial \delta}{\partial E})_0^2 s_E^2]^{\frac{1}{2}}$$

$$= [(\frac{4\bar{L}}{\pi \bar{d}^2\bar{E}})^2 s_F^2 + (\frac{4\bar{F}}{\pi\bar{d}^2\bar{E}})^2 s_L^2 + (\frac{-8\,\overline{FL}}{\pi\bar{d}^3\bar{E}})^2 s_d^2 + (\frac{-4\overline{FL}}{\pi\bar{d}^2\bar{E}^2})^2 s_E]^{\frac{1}{2}}$$

$$= [(\frac{4\times 1\ 000}{\pi\times 10^2\times 20\ 600})^2\times 2\ 000^2 + (\frac{4\times 4\ 000}{\pi\times 10^2\times 20\ 600})^2\times 3.3^2$$

$$(\frac{-8\times 40\ 000\times 1\ 000}{\pi\times 10^3\times 20\ 600})^2\times 0.1^2 + (\frac{-4\times 40\ 000\times 1\ 000}{\pi\times 10^2\times 20\ 600^2})^2\times 618^2]^{\frac{1}{2}}$$

$= 1.526$ mm

即

$$(\bar{\delta}, s_\delta) = \frac{4(\bar{F}, s_F)(\bar{L}, s_L)}{\pi(\bar{d}, s_d)(\bar{E}, s_E)} = (24.72, 1.526)\text{mm}$$

2）用变异系数法　用式（6.6-12）～式（6.6-14）$\rho_{ij}=0$。并先求出各随机变量的变异系数

$$v_d = \frac{s_d}{\bar{d}} = \frac{0.1}{10} = 0.01$$

$$v_L = \frac{s_L}{\bar{L}} = \frac{3.3}{1\ 000} = 0.003\ 3$$

$$v_F = \frac{s_F}{\bar{F}} = \frac{2\ 000}{40\ 000} = 0.05$$

$$v_E = \frac{s_E}{\bar{E}} = \frac{618}{20\ 600} = 0.03$$

则

$$\bar{\delta} = \frac{4}{\pi}\overline{FL}\,\bar{d}^{-2}E^{-1}$$

$$= \frac{4}{\pi}\times 4\ 000\times 1\ 000\times 10^{-2}\times 20\ 600^{-1}$$

$$= 24.72\ \text{mm}$$

$$v_\delta = (v_F^2 + v_L^2 + 2^2 v_d^2 + v_E^2)^{\frac{1}{2}}$$

$$= (0.05^2 + 0.003\ 3^2 + 2^2\times 0.01^2 + 0.03^2)^{\frac{1}{2}}$$

$$= 0.061\ 73$$

$$s_\delta = \bar{\delta} v_\delta = 24.72\times 0.061\ 73 = 1.526\ \text{mm}$$

即　（$\bar{\delta}$，s_δ）=（24.72，1.526）mm

3）用基本函数法　由于变量较多，故分步计算，按表 6.6-2，按序号 4

$$\bar{d}^2 = 10^2 = 100$$

$$s_{d^2} = 2\bar{d}s_d = 2\times 10\times 0.1 = 2$$

d^2E，按序号 6，$\rho_{12}=0$

$$\overline{d^2E} = \bar{d}^2\bar{E} = 10^2\times 20\ 600 = 2\ 060\ 000$$

$$s_{d^2E} = [(\bar{d}^2)^2 s_E^2 + \bar{E}^2 s_{d^2}^2]^{\frac{1}{2}}$$

$$= [(100)^2\times 0.03^2 + 20\ 600^2\times 2^2]^{\frac{1}{2}}$$

$$= 74\ 274$$

FL，按序号 6，$\rho_{12}=0$

$$\overline{FL} = \bar{F}\,\bar{l} = 40\ 000\times 1\ 000 = 40\ 000\ 000$$

$$s_{FL} = (\bar{F}^2 s_L^2 + \bar{L}^2 s_F^2)^{\frac{1}{2}}$$

$$= (40\ 000^2\times 3.3^2 + 1\ 000^2\times 2\ 000^2)^{\frac{1}{2}}$$

$$= 2\ 004\ 351$$

$\frac{FL}{d^2E}$，按序号 7，$\rho_{12}=0$

$$\overline{\frac{FL}{d^2E}} = \frac{\overline{FL}}{\overline{d^2E}} = \frac{40\ 000\ 000}{2\ 060\ 000} = 19.417\ 5$$

$$s_{(\frac{FL}{d^2E})} = \frac{\overline{FL}}{\overline{d^2E}}\ (\frac{s_{FL}^2}{\overline{FL}^2} + \frac{s_{d^2E}^2}{\overline{d^2E}^2})^{\frac{1}{2}}$$

$$= \frac{40\ 000\ 000}{2\ 060\ 000}\times\ (\frac{2\ 004\ 351^2}{40\ 000\ 000^2} + \frac{74\ 274^2}{2\ 060\ 000^2})^{\frac{1}{2}}$$

$$= 1.198\ 7$$

$\delta = \frac{4}{\pi}\times\frac{FL}{d^2E}$，按序号 2

$$\bar{\delta} = \frac{4}{\pi}\ (\overline{\frac{FL}{d^2E}}) = \frac{4}{\pi}\times 19.417\ 5 = 24.72\ \text{mm}$$

$$s_\delta = \frac{4}{\pi}s\ (\frac{FL}{d^2E}) = \frac{4}{\pi}\times 1.198\ 7 = 1.526\ \text{mm}$$

即（$\bar{\delta}$，s_δ）=（24.72，1.526）mm

1.4.4　求解应力分布参数的统计模拟方法

用统计模拟方法（又称 Monte Carlo 方法），可以通过对有关随机变量的统计抽样试验或随机模拟估计函数的统计量。

基本思路和步骤为：

1）构造概率模型；

2）定义随机变量；

3）通过模拟获得子样；

4）统计计算。

1.5　强度分布的确定

确定强度分布的步骤：

1）确定强度判据；

2）确定名义强度分布；

3）修正名义强度分布。

1.6　已知应力和强度分布时的可靠度计算

当应力 $S\sim(\mu_s,\sigma_s^2)$ 与强度 $\delta\sim(\mu_\delta,\sigma_\delta^2)$ 均为正态分布时，$y=\delta-S\sim(\mu_y,\sigma_y^2)$，其中，

$$\mu_y = \mu_\delta - \mu_s$$

$$\sigma_y^2 = \sigma_\delta^2 + \sigma_s^2$$

可靠度可表达为

$$R = \int_0^\infty \frac{1}{\sigma_y\sqrt{2\pi}}\exp[-\frac{1}{2}(\frac{y-\mu_y}{\sigma_y})^2]\mathrm{d}y \quad (6.6\text{-}15)$$

令

$$z = \frac{y-\mu_y}{\sigma_y}$$

则 z 为标准正态分布随机变量，且有

$$R = \int_{-\frac{\mu_\delta-\mu_S}{\sqrt{\sigma_\delta^2+\sigma_S^2}}}^\infty \frac{1}{\sqrt{2\pi}}\exp[-\frac{z^2}{2}]\mathrm{d}z \quad (6.6\text{-}16)$$

上式的可靠度 R 可通过查阅标准正态分布表获得，即

$$R = 1-\Phi\ (-\frac{\mu_\delta-\mu_S}{\sqrt{\sigma_\delta^2+\sigma_S^2}}) = \Phi\ (\frac{\mu_\delta-\mu_S}{\sqrt{\sigma_\delta^2+\sigma_S^2}}) \quad (6.6\text{-}17)$$

可靠性联结方程（或称耦合方程）：

$$z_0 = -\frac{\mu_\delta-\mu_S}{\sqrt{\sigma_\delta^2+\sigma_S^2}}$$

$z_R = -z_0$ 称为可靠性系数或可靠度指数。

2　概率设计常用的随机变量及其分布

概率设计的数学基础是概率论与数理统计。概率设计中将载荷、强度等设计变量作为随机变量，这些随机变量有各自的分布，因而概率设计涉及到对随机变量的样本数据进行统计处理，分布拟合，参数估计，可靠度计算，等等。在强度概率设计问题中，需要选择一个适当概率密度函数。本节介绍几个常用的密度函数，以及它们的可靠性与风险函数、概率密度函数、可靠性函数以及风险函数之间的关系，并对这些分布的均值和方差的求法进行了说明。

2.1　分布特征

概率密度函数可以分为离散型和连续型两类。离散型概率密度函数所取的值仅限于整数。连续型密度函数能够在一个具体的区域内取得所有的值。密度函数必须达到两条标准：

1) 对于所有 x 的值满足 $f(x) \geqslant 0$；

2) 对于连续型分布，$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$，对离散型的分布，$\sum_n f(x_n) = 1$。

累积分布函数是随机变量 X 小于某个具体值的概率 $P(X<x)$。对于连续型分布，累积的分布函数被定义为：

$$F(x) = \int_{-\infty}^{x} f(\tau)\mathrm{d}\tau \tag{6.6-18}$$

x 落于两个具体值 a、b 之间的概率是在概率密度函数与横座标轴之间从 a 到 b 的区域的面积，即 $P(a<X<b) = \int_a^b f(x)\mathrm{d}x$，等价于 $F(b)-F(a)$，如图6.6-2所示。

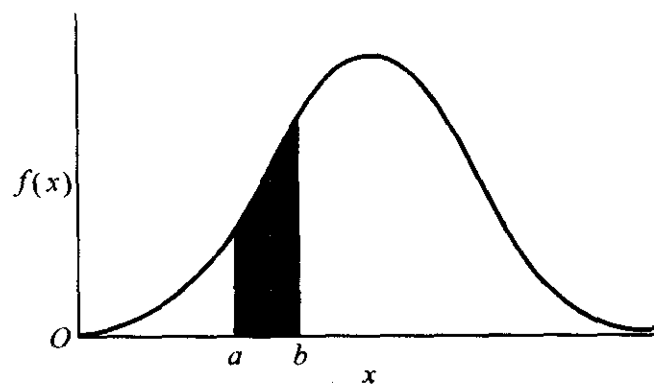


图6.6-2　x 落在两特定值之间的概率

对于离散型分布，x 取某些值的概率一定大于0。例如，在四次投掷硬币中，x 代表得到正面的次数，正面一次也没出现的概率是0.062 5，恰好出现一次正面的概率是0.25。该概率密度函数如图6.6-3所示。

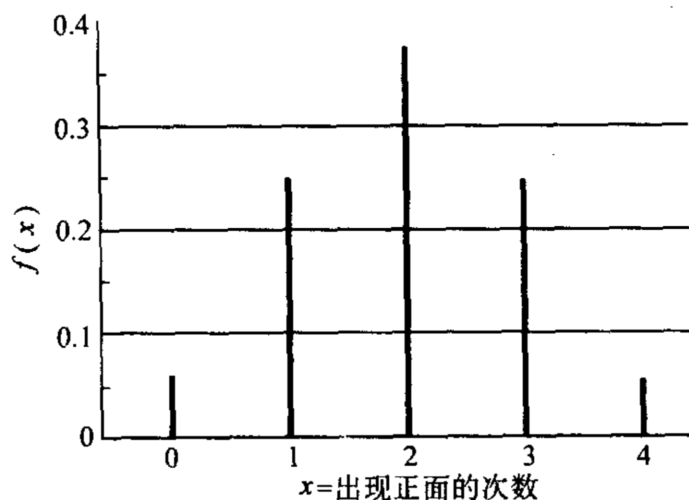


图6.6-3　离散概率密度函数

对于离散型分布的累积分布函数是：

$$F(x) = \sum_{\tau \leqslant x} f(\tau) \tag{6.6-19}$$

投掷硬币实验的累积分布函数如图6.6-4所示。

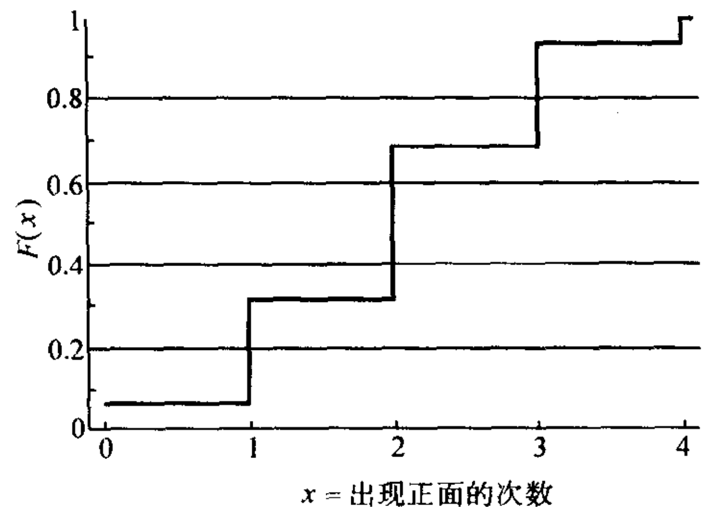


图6.6-4　离散累积分布函数

可靠性函数是累积分布函数 $F(x)$ 的补充，它由概率密度函数通过以下关系获得：

$$R(x) = 1 - F(x) = 1 - \int_{-\infty}^{x} f(\tau)\mathrm{d}\tau = \int_x^{\infty} f(\tau)\mathrm{d}\tau \tag{6.6-20}$$

可靠性函数与累积分布函数的关系如图6.6-5所示。

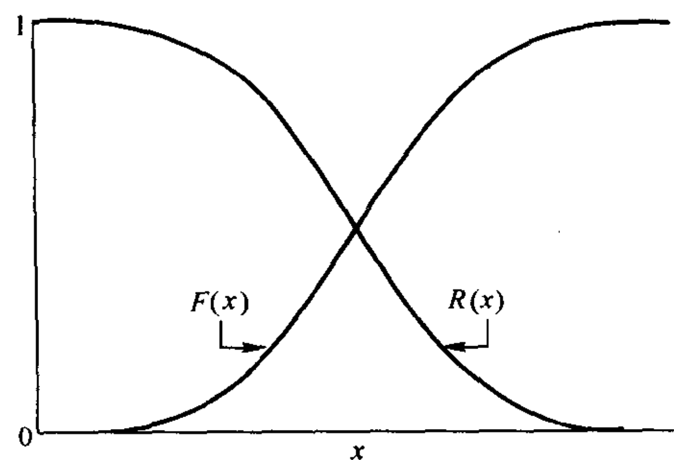


图6.6-5　可靠性函数和累积分布函数之间的关系

风险函数是即时失效率，是对失效可能性的度量。风险函数定义为：

$$h(x) = \frac{f(x)}{R(x)} \tag{6.6-21}$$

可靠性函数和概率密度函数能够通过风险函数使用表达式

$$R(t) = e^{-\int_{-\infty}^{x} h(\tau)\mathrm{d}\tau} \tag{6.6-22}$$

以及

$$f(x) = h(x)e^{-\int_{-\infty}^{x} h(\tau)\mathrm{d}\tau} \tag{6.6-23}$$

分布的期望值能够通过表达式

$$E(x) = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \mu \tag{6.6-24}$$

或

$$E(x) = \int_0^{\infty} R(x)\mathrm{d}x = \mu \tag{6.6-25}$$

分布的方差为

$$V(x) = E(x^2) - \mu^2 \tag{6.6-26}$$

其中 μ 是分布的均值，且

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x \tag{6.6-27}$$

分布特征值通常通过矩生成函数获得，如果 x 是一个连续型随机变量，那么原函数的第 n 次矩是

$$E(x^n) = \int_{-\infty}^{\infty} x^n f(x)\mathrm{d}x \tag{6.6-28}$$

如果 x 是一个离散型随机变量，那么原函数的第 n 次矩是

$$E(x^n) = \sum_x x^n(x)\mathrm{d}x \qquad (6.6\text{-}29)$$

原函数的一次矩是分布的均值，关于均值的第二次矩是方差，关于均值的第三次矩是分布的倾斜。如果一个单顶点分布有一个向右延伸的长尾巴，那就说明，这是右倾斜。均值的第四次矩是对峰态的测量。

2.2　正态分布

载荷、强度、寿命等是概率设计涉及的重要的指标。这些指标一般都是随机变量，有一定的取值范围，服从一定的统计分布。

分布的类型很多，确定载荷、强度或产品的寿命服从何种分布的方法有两种，一种是根据其物理背景判断。产品的寿命分布与其所承受的载荷情况、产品的结构特点及其物理、化学、力学性能有关，与产品发生失效的物理过程有关。例如，通过失效分析，证实产品的故障模式或失效机理与某种类型分布的物理背景相接近时，可基本判定产品寿命服从该类型的分布。另一种方法是失效数据，用统计推断的方法来判断属于何种分布。

2.2.1　正态分布

正态分布是一个连续型分布，也叫做高斯分布。它是最容易处理的可靠性密度函数之一。正态分布有两条有用的性质，一条是方差的可加性，另一条是随机变量总和的趋势与单个密度函数无关，当相加起来时符合正态分布（这又叫做中心极限定理或大数定理），这两条性质使得正态分布在对一个已建立的公差进行建模时变得非常有用。

正态可靠性密度函数定义为：

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],\ -\infty < x < \infty \qquad (6.6\text{-}30)$$

式中，μ 是分布均值，是对中心趋势或中点的度量；σ 是分布标准偏差，是对发散性的度量。

分布的均值由样本的均值估计得来，即：

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (6.6\text{-}31)$$

$$\hat{\sigma} = s = \sqrt{\frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}} \qquad (6.6\text{-}32)$$

另一个经常使用的求样本标准偏差的表达式是：

$$\hat{\sigma} = s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})}{(n-1)}} \qquad (6.6\text{-}33)$$

这两个公式在代数上是可以证明的。式（6.6-32）是计算时最经常使用的。

正态分布不是倾斜的，均值、中位数和众数都是平衡的。图 6.6-6 表达了正态概率密度函数。

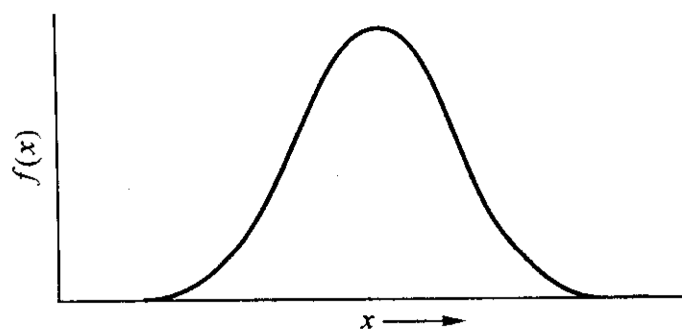


图 6.6-6　正态概率密度函数

2.2.2　标准正态分布

$\mu=0$，$\sigma^2=1$ 的正态分布称为标准正态分布，其概率密度函数为：

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \quad (-\infty < x < \infty) \qquad (6.6\text{-}34)$$

正态概率密度函数不能被直接积分。因此，可靠性和累积分布函数通常以标准正态分布的形式。通过以下公式实现向标准正态公布的转换：

$$z = \frac{x-\mu}{\sigma} \qquad (6.6\text{-}35)$$

标准正态随机偏差 z，均值是 0，标准差是 1。标准正态累积分布函数 Φ（z），已经以表格的形式在附表 1 列出，在所有的统计教材中都可以查到。

正态可靠性函数是：

$$R(x) = \int_x^\infty \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\mathrm{d}\tau$$
$$= 1-\Phi(z),\ -\infty < x < \infty \qquad (6.6\text{-}36)$$

其中，Φ（z）是标准正态累积分布函数。

标准正态分布函数与标准正态累积分布函数如图 6.6-7 所示。

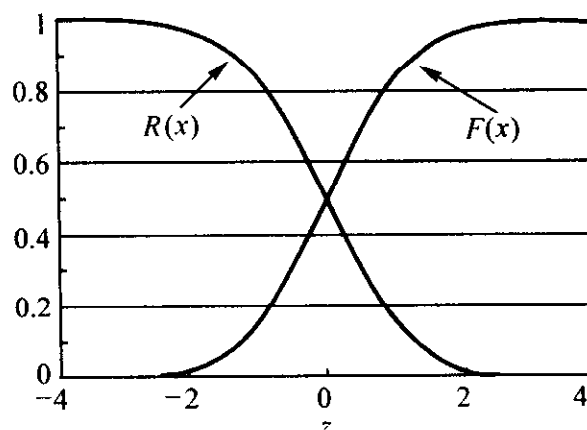


图 6.6-7　标准正态可靠性和累积分布函数

例 2　制造厂生产金属梁，平均承受力是 40 单位，方差为 36。如果承受力服从正态分布，那么 1）一根横梁具有超过 52 单位承受力的概率是多少？2）小于 43 单位的呢？

解　1）所求的解是正态概率密度函数以下 $x=52$ 以右的区域。

$$P(x > 52) = R(52) = \int_5^\infty 2\frac{1}{\sqrt{36}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-40}{\sqrt{36}}\right)^2\right]\mathrm{d}x$$

由下式变为标准型：

$$z = \frac{x-\mu}{\sigma} = \frac{52-40}{\sqrt{36}} = 2$$

由附表 1 得，从 $z=-\infty$ 到 $z=2$ 的标准正态分布曲线以区域是 Φ（2）$=0.977\,2$

R（52）$=1-0.977\,2=0.022\,8$

2）所求的解是正态概率密度函数以下 $x=43$ 以左的区域。

$$P(x < 43) = 1-R(43) = \int_{-\infty}^{43}\frac{1}{\sqrt{36}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-40}{\sqrt{36}}\right)^2\right]\mathrm{d}x$$

由下式变为标准型：

$$z = \frac{43-40}{\sqrt{36}} = 0.5$$

由附表 1 得，在标准正态分布曲线以下的区域从 $z=-\infty$ 到 $z=0.5$，Φ（0.5）$=0.691\,5$。

正态风险函数是单调递增的，定义为：

$$h(x) = \frac{f(x)}{R(x)} = \frac{\Phi(z)}{\sigma[1-\Phi(z)]},\ -\infty < x < \infty$$

其中，$\Phi(z)$ 是标准正态概率密度函数。

图6.6-8表达了标准正态风险函数。

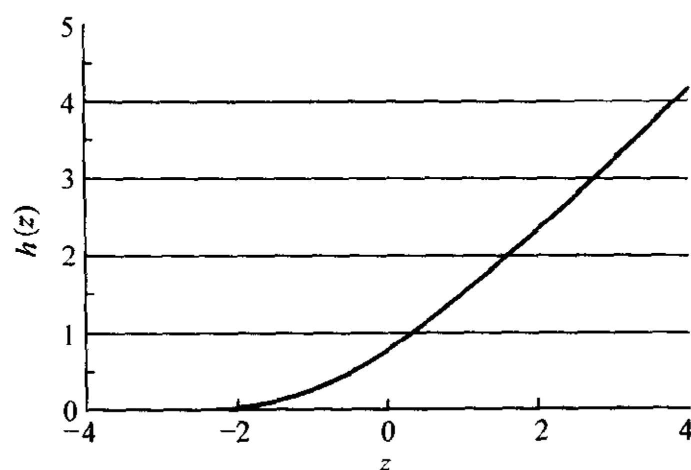


图6.6-8 标准正态风险函数

例3 已知失效时间分别为28、29、31、34和37 h，那么27 h的估计的即时失效率是多少?

解:

样本的均值是:

$$\hat{\mu}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{28+29+31+34+37}{5}=31.8$$

样本的标准偏差是:

$$\hat{\sigma}=\sqrt{\frac{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2}{n(n-1)}}=\sqrt{\frac{5\times(5\ 111)-(159)^2}{5\times(5-1)}}$$

$$=3.701$$

即时失效率由风险函数给出，即:

$$h(27)=\frac{\Phi(z)}{\hat{\sigma}[1-\Phi(z)]}=\frac{\Phi\left(\frac{27-31.8}{3.701}\right)}{3.701\times\left[1-\Phi\left(\frac{27-31.8}{3.701}\right)\right]}$$

$$=0.051\ 5$$

2.2.3 截尾正态分布

工程实际中有很多试验或观察数据近似服从正态分布。但正态分布的取值范围（$-\infty$到$+\infty$）不很符合实际情况。考虑到许多试验或观察数据无负值，因此用截尾正态分布来表示较为准确。截尾正态分布定义为:

若 X 是一个非负的随机变量，且密度函数为

$$f(x)=\frac{1}{\alpha\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\quad(0\leqslant x<\infty)\tag{6.6-37}$$

则称 X 服从截尾正态分布。式中，α 为"正规化常数"，它保证了 $\int_0^{-\infty}f(x)\mathrm{d}x=1$。

正态分布是最常用的分布，在机械概率设计中，可以用来描述零件的强度分布。从物理背景上讲，如果影响某个随机变量的独立因素很多，且不存在起决定作用的主导因素时，则该随机变量一般可用正态分布来描述。正态分布随机变量的取值范围从负无穷大直到正无穷大，从这一点来看，强度不可能是真正的正态分布，而只可能是截尾正态分布。

2.3 对数正态分布

正态分布和对数正态分布是紧密相关的。如果 x 是一个对数正态随机变量，那么变量 $y=\ln(x)$ 就是一个正态随机变量。对数正态概率密度函数是:

$$f(x)=\frac{1}{\sigma x\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right],x>0\tag{6.6-38}$$

对数正态分布也是连续型的，也是最广泛使用的寿命分布之一。它高度倾斜，当效果增强时，它使自己适应于具体情况。对数正态分布已经在金属失效循环、导体的寿命、轴承的寿命，以及许多其他现象的建模中使用。

对数正态分布的参数是位置参数 μ，以及比例参数 σ。比例参数有时会涉及到形状参数。位置参数能够通过以下表达式估计:

$$\hat{\mu}=\frac{\sum_{i=1}^{n}\ln x_i}{n}\tag{6.6-39}$$

式中，n 是样本容量。

比例参数由以下表达式估计:

$$\hat{\sigma}=\sqrt{\frac{n\sum_{i=1}^{n}\ln x_i^2-\left(\sum_{i=1}^{n}\ln x_i\right)^2}{n(n-1)}}\tag{6.6-40}$$

对数正态概率密度函数如图6.6-9所示。

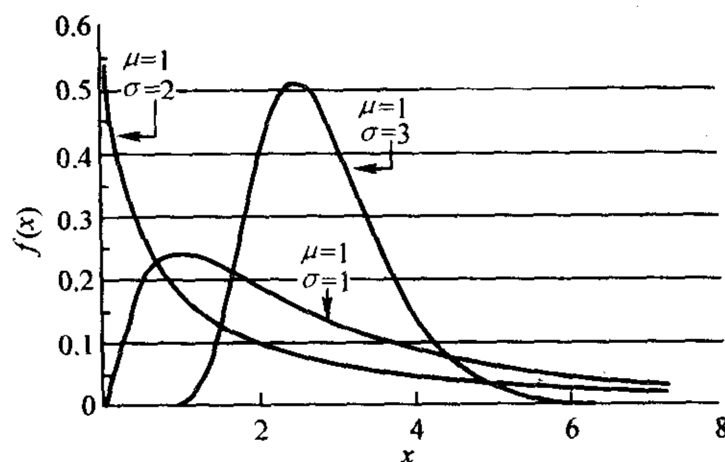


图6.6-9 对数正态概率密度函数

对数正态分布的均值是:

$$E(x)=\exp\left(\mu+\frac{\sigma^2}{2}\right)=T_{50}\exp\left(\frac{\sigma^2}{2}\right)\tag{6.6-41}$$

对数正态分布的方差是:

$$V(x)=[\exp(2\mu+\sigma^2)][\exp(\sigma^2)]=T_{50}\exp\left(\frac{\sigma^2}{2}\right)[\exp(\sigma^2)-1]\tag{6.6-42}$$

对数正态可靠性函数是:

$$R(x)=1-\Phi\left(\frac{\ln x-\mu}{\sigma}\right),\ x>0\tag{6.6-43}$$

对数正态可靠性函数如图6.6-10所示。

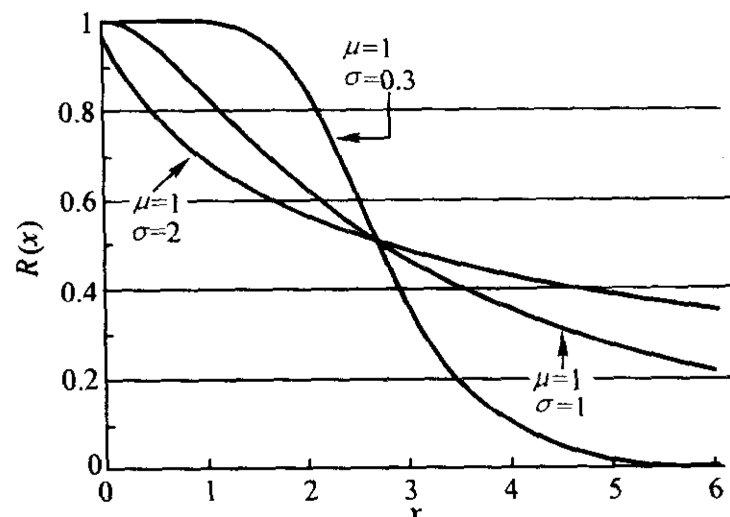


图6.6-10 对数正态可靠性函数

对数正态风险函数如下:

$$h(x)=\frac{f(x)}{R(x)}=\frac{\Phi\left(\frac{\ln x-\mu}{\sigma}\right)}{x\sigma\left[1-\Phi\left(\frac{\ln x-\mu}{\sigma}\right)\right]},x>0\tag{6.6-44}$$

对数正态风险函数与参数的值无关，它是先增后减，无限接近于 0 的。对于较大的 σ 值（$\sigma>1.5$）时，风险函数增长如此迅速，它最初出现在图像上时可能是递减的，但是，这并不是本来的情况。对数正态风险函数如图 6.6-11 所示。

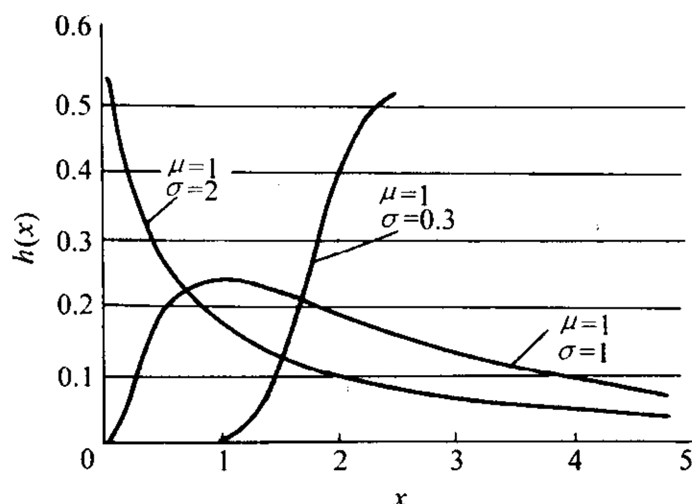


图 6.6-11 对数正态风险函数

2.4 威布尔（Weibull）分布

瑞典工程师威布尔（Weibull）采用“链式”模型来研究、描述结构强度和寿命问题，假设一个结构是由若干小元件（设为 n 个）串联而成，将结构看成是由 n 个环构成的一条链条，其强度（或寿命）取决于最薄弱环的强度（或寿命）。单个链的强度（或寿命）为一随机变量，设各环强度（或寿命）相互独立，分布相同，则求链强度（或寿命）的概率分布就变成求极小值分布问题，由此得出了威布尔分布函数。

由于威布尔分布是根据最弱环节模型或串联模型得到的，能充分反映材料缺陷和应力集中源对材料疲劳寿命的影响，所以，将它作为材料或零件的寿命分布模型或给定寿命下的疲劳强度模型是比较合适的。

威布尔分布是一个连续型分布，现在已经被广泛使用，特别是在可靠性领域。威布尔分布之所以流行，是因为其能应用于小容量样本的能力和它的适应性。

三参数威布尔分布记为 $X\sim W(\beta,\theta,\alpha,x)$，其中 β 为形状参数，θ 为尺度参数，α 为位置参数。这些变量的可接受范围是 $0<\beta<\infty$，$0<\theta<\infty$。

三参数威布尔分布的密度函数为

$$f(x)=\frac{\beta(x-\alpha)^{\beta-1}}{\theta^{\beta}}\exp\left[-\left(\frac{(x-\alpha)}{\theta}\right)^{\beta}\right],x\geqslant 0 \qquad (6.6\text{-}45)$$

三参数威布尔分布的分布函数为

$$F(x)=1-\exp\left[-\left(\frac{x-\alpha}{\theta}\right)^{\beta}\right],x>0 \qquad (6.6\text{-}46)$$

把三参数威布尔分布的位置参数 α 取为 0，则简化为两参数威布尔分布。其密度函数为

$$f(x)=\frac{\beta x^{\beta-1}}{\theta^{\beta}}\exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right],x\geqslant 0 \qquad (6.6\text{-}47)$$

两参数威布尔分布的分布函数为

$$F(x)=1-\exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right],x>0 \qquad (6.6\text{-}48)$$

通常，β 一般在 0.5～8 之间。随着 β 的增加，威布尔分布接近于 θ，方差接近于 0。由于在 θ 点有 63.2% 的总体失效，因此 θ 也被认为是特征寿命，而这些与 β 的值无关。

2.4.1 形状参数的作用

通过选择形状参数 β，威布尔概率密度函数可以呈现不同的形状。如图 6.6-12 所示。

根据 β 值的不同，威布尔分布能被应用于多种情形，能等价或近似于其他几个分布，例如，$\beta=1$，威布尔分布等同于指数分布；$\beta=2.5$，威布尔分布近似于对数正态分布；$\beta=3.6$，威布尔分布近似于正态分布。由于这个适应性，威布尔分布几乎能对所有观察到的失效率进行建模。

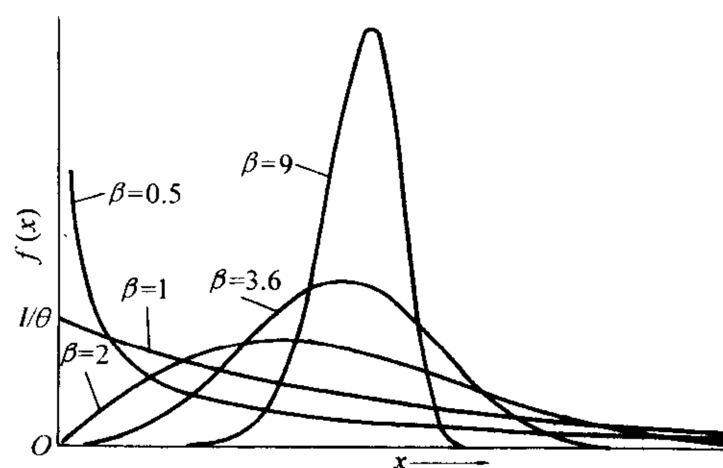


图 6.6-12 威布尔概率密度函数

2.4.2 威布尔分布的均值和方差

威布尔分布的均值是：

$$E(x)=\theta\Gamma\left(1+\frac{1}{\beta}\right) \qquad (6.6\text{-}49)$$

式中，$\Gamma(x)$ 是 γ 函数，在附表 5 中已经制成表格。

威布尔分布的方差是：

$$V(x)=\theta^{2}\left[\Gamma\left(1+\frac{2}{\beta}\right)-\Gamma^{2}\left(1+\frac{1}{\beta}\right)\right] \qquad (6.6\text{-}50)$$

如果 $\beta<1$，那么威布尔分布的均值将大于 θ。如果 $\beta=1$，威布尔分布的均值等于 1。如果 $\beta>1$，威布尔分布的均值小于 θ，且随着 x 的减小接近于 θ。随着 β 增长到无穷，威布尔分布的方差减小，且无限接近于 0。

2.4.3 威布尔可靠性函数与风险函数

威布尔可靠性函数是：

$$R(x)=\exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right],x>0 \qquad (6.6\text{-}51)$$

图 6.6-13 说明了 β 取不同值时，对可靠性函数的影响。

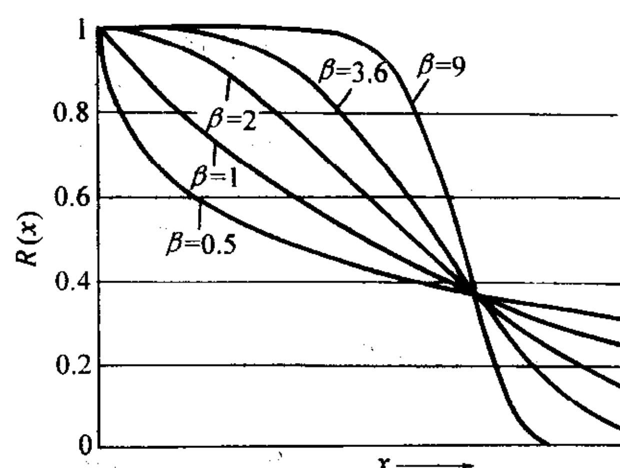


图 6.6-13 威布尔可靠性函数

对于 $\beta<1$ 的情况，可靠性开始迅速递减，然后趋于平缓。这是初期致命失效的结果。对于 $\beta=1$ 的情况，可靠性逐渐递减，结果是失效率趋于恒定。当 $\beta>1$ 时，可靠性开始缓慢递减，然后随着特征寿命趋近而递减。

威布尔风险函数是：

$$h(x)=\frac{\beta x^{\beta-1}}{\theta^{\beta}},x>0 \qquad (6.6\text{-}52)$$

β 对于风险函数的效果如图 6.6-14 所示。

在递减的失效率中，形状特征小于 1 种结果。如果 $\beta=1$，失效率恒等于 $1/\theta$，如果 $\beta>1$，失效率是递增的，β 的值越大，失效率递增得越迅速。

2.5 Ⅰ型极值分布（Gumbel 分布）

2.5.1 Ⅰ型极大值分布概率密度函数

Ⅰ型极大值分布概率密度函数为：

$$f(x)=\frac{1}{\sigma}e^{-\frac{x-\mu}{\sigma}}e^{-e^{-\frac{x-\mu}{\sigma}}}\quad(-\infty<x<\infty,\sigma>0,-\infty<\mu<\infty)\tag{6.6-53}$$

式中，μ 为极值分布的位置参数；σ 为极值分布的尺度参数。

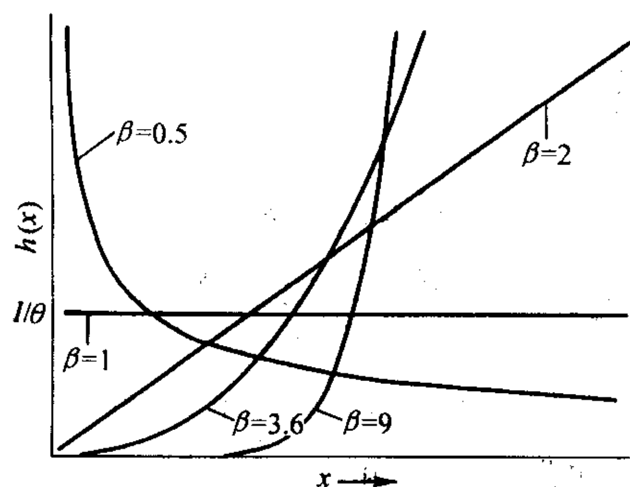


图 6.6-14　威布尔风险函数

极大值分布的数字特征为：$E(x)=\mu+0.5772\sigma$，$D(x)=1.644\sigma^2$

2.5.2　Ⅰ型极小值分布

Ⅰ型极小值分布概率密度函数为：

$$f(x)=\frac{1}{\sigma}e^{\frac{x-\mu}{\sigma}}e^{-e^{\frac{x-\mu}{\sigma}}}\quad(-\infty<x<\infty,0<\sigma<\infty,-\infty<\mu<\infty)\tag{6.6-54}$$

式中，μ 为位置参数；σ 为尺度参数。

极小值分布的数字特征为：$E(x)=\mu-0.5772\sigma$，$D(x)=1.644\sigma^2$。

Ⅰ型极值分布是对应于大量子样的最小值或最大值的分布，主要用来描述一个随机变量出现极小值或极大值的现象。这类问题很多，例如，建筑工程中结构构件抗力的最小值分布，结构载荷的最大值分布；机械工程中导致机械产品失效的强度或寿命的最小值分布，短期过载的最大值分布，串联系统的最弱元件，并联系统中最强元件等。

Ⅰ型极极小值分布与威布尔分布的关系为：设随机变量 T 服从威布尔分布，则 $X=\ln T$ 服从Ⅰ型最小极值分布。由于Ⅰ型最小极值分布与威布尔分布有这种关系，因此，可以用它作为威布尔分布的拟合优度检验。所以，极值分布也是寿命数据分析中经常用到的一种分布类型。

2.6　χ^2 分布、F 分布和 t 分布

χ^2 分布、F 分布和 t 分布首先被用于样本的分布而不是对寿命时间进行建模。在其他领域，这些函数被用于：

1）在一个区间内对失效的次数进行计数；
2）决定合适的区间；
3）限定置信极限；
4）获得关于总体均值的推论；
5）获得关于总体方差的推论。

2.6.1　χ^2 分布

如果 y_1、$y_2\cdots y_\nu$ 相互独立，服从标准正态分布，那么：

$$\chi^2=y_1^2+y_2^2+\cdots+y_\nu^2$$

即具有 ν 个自由度的 χ^2 分布（独立变量的平方相加）。χ^2 是 γ 分布当 $\lambda=2$，$\eta=2/\nu$ 的特殊情况。若失效次数是在给定区间定义的一个整数，χ^2 分布才可能对这种情况进行建模。

χ^2 概率密度函数是：

$$f(x)=\frac{1}{2^{(\nu/2)}\Gamma(\nu/2)}x^{(\nu/2-1)}e^{(-x/2)},x>0\tag{6.6-55}$$

χ^2 分布形状如图 6.6-15 所示。

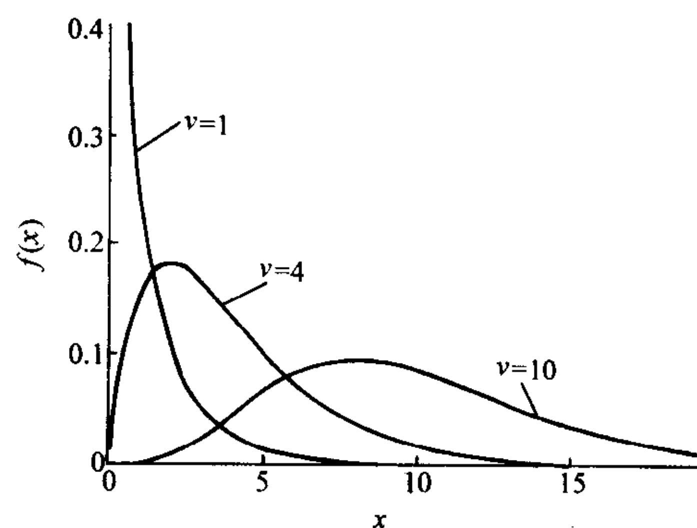


图 6.6-15　χ^2 概率密度函数

χ^2 分布的均值是：

$$E(x)=\nu\tag{6.6-56}$$

χ^2 分布的方差是：

$$V(x)=2\nu\tag{6.6-57}$$

χ^2 分布有几条有用的性质。

1）两个或两个以上独立的服从 χ^2 分布的变量之和仍然服从 χ^2 分布，且自由度等于各个变量自由度之和。

2）两个自由度数分别为 ν_1 和 ν_2，独立地服从 χ^2 分布的变量之比，再乘上 ν_1/ν_2 就得到了 F 分布。

3）如果两个变量服从 χ^2 分布，那么任意一个变量比上两个变量之和即是 β 分布。

4）随着 ν 增大，χ^2 分布的期望 ν 与方差 2ν 接近正交。

2.6.2　F 分布

F 分布用来推断方差和构建置信极限。给定两个独立 χ_1^2 变量和 χ_2^2，自由度为 ν_1 和 ν_2，则随机变量

$$F=\frac{\chi_1^2/\nu_1}{\chi_2^2/\nu_2}\tag{6.6-58}$$

有一个自由度为 ν_1 和 ν_2 的 F 分布，该变量的概率密度函数是：F 概率密度函数如图 6.6-16 所示。

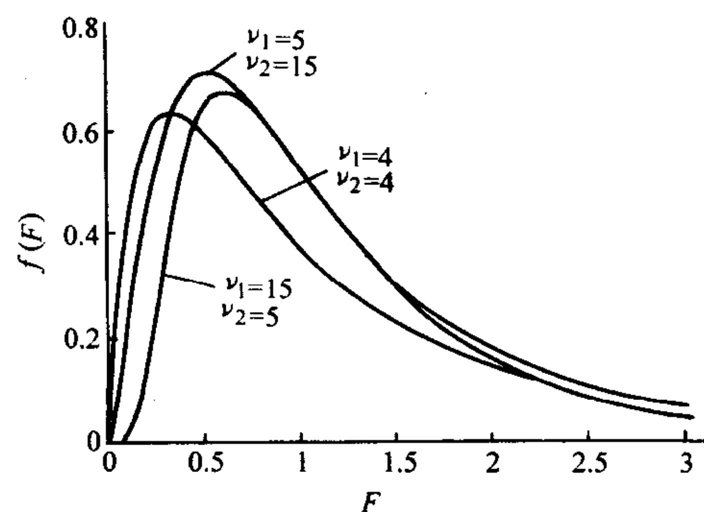


图 6.6-16　F 概率密度函数

由图 6.6-16 可知，F 分布曲线取决于自由度的分布顺序。F 分布分子中 χ^2 变量的自由度数首先被确定，然后是分母中 χ^2 变量的自由度数。如果 $F_\alpha(\nu_1,\nu_2)$ 代表自由度为 ν_1 和 ν_2，α 右部的 F 概率密度函数下方的区域，那么：

$$F_{1-\alpha}(\nu_1,\nu_2)=\frac{1}{F_\alpha(\nu_2,\nu_1)}\tag{6.6-59}$$

F 分布在附表 4 中列出。

F 分布的均值是：

$$E(F)=\frac{\nu_2}{\nu_2-2},\nu_2>2\tag{6.6-60}$$

注意，F 分布的均值与 ν_1 无关，F 分布的方差为：

$$V(F)=\frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)},\nu_2>4 \qquad (6.6\text{-}61)$$

如果 s_1^2 与 s_2^2 是从独立事件中选取的两个样本的方差，正态分布的方差为 σ_1^2 和 σ_2^2，样本的大小为 n_1，n_2，那么统计量

$$F=\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \qquad (6.6\text{-}62)$$

是以 $\nu_1=n_1-1$，$\nu_2=n_2-1$ 的 F 分布，这样，F 分布也以方差比分布著称。

2.6.3 *t* 分布

t 分布用于在方差未知或有效样本太少而不能使用正态分布时获得关于均值的推断和构建置信区间。对于 $n>30$ 的大子样，样本方差 s^2 是 σ^2 的一个合理估计量。对于 $n<30$ 的小型样本，使用 s^2 作为 σ^2 的估计量所引起的误差导致统计量

$$t=\frac{\overline{x}-\mu}{s/\sqrt{n}} \qquad (6.6\text{-}63)$$

t 分布是一随机变量

$$t=\frac{z}{\sqrt{\frac{x^2}{\nu}}} \qquad (6.6\text{-}64)$$

这里，z 是一个标准正态随机变量，x^2 是 γ 自由度的随机变量。

t 分布是以 1 为参数，γ 自由度的 F 分布的一个特例。t 概率密度函数是：

$$f(t)=\frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)\sqrt{\Pi\nu}}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2},-\infty<t<\infty \qquad (6.6\text{-}65)$$

t 概率密度函数如图 6.6-17 所示。

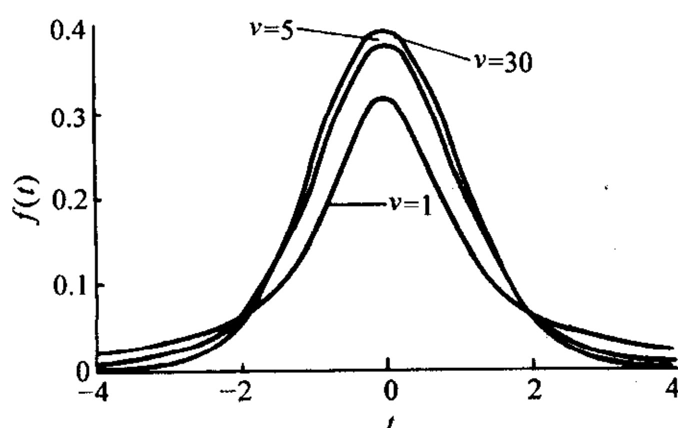


图 6.6-17 t 概率密度函数

由图 6.6-17 可知，t 分布的方差随 ν 的增加而减少，如果 ν 大于 30，即使 ν 增加，方差也几乎没有变化，t 分布实际上与标准正态分布完全相同。

t 分布的均值是：

$$E(t)=0 \qquad (6.6\text{-}66)$$

标准正态偏离量为 $z=\frac{x-\mu}{\sigma}$，中心极限法则确定，均值的标准偏差被倒转比例到了样本数字的方根上。所以，大小为 n 的一个样本，标准正态偏离为 $z=\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$。

t 分布方差是：

$$V(t)=\frac{\nu}{\nu-2},\ \nu>2 \qquad (6.6\text{-}67)$$

3 参数估计

参数估计有很多种方法，在可靠性工程中，应用较广泛的有：

1）极大似然估计法；
2）风险函数图；
3）概率图；
4）矩估计法。

一般需要参数估计量具有如下性质。

1）无偏性　期望值等于参数真实值。

2）最小方差　估计的方差越小，达到一定精度要求所需的子样容量越小，估计量越有效。具有最小方差的估计量是最有效的估计量。

3）一致性　随着子样容量的增加，估计的参数值越接近参数真值。

4）充分性　估计量要利用数据中所有有用的信息。

用于参数估计的方法主要依据数据或试验的类型以及分布形式。除以上列举的参数估计方法外，这一节还描述了截尾数据，介绍指数分布、正态分布、对数正态分布、威布尔分布和 γ 分布的参数估计方法。完整数据和截尾数据的参数估计方法分别介绍。

参数估计学是冗长且枯燥的。通常利用计算来帮助数据分析。Dodson 和 Kirkland 于 1994 年开发了参数估计的综合软件。这个软件能够进行最大似然估计、风险图、概率图和矩估计的计算。常用的分布有指数分布、威布尔分布、正态分布、对数正态分布和 γ 分布。

3.1 截尾数据

一个截尾数据的集合包括失效项与存活项的信息。考虑表 6.6-3 种的数据，在试验台上进行八个产品的试验；其中的三个失效了，五个到试验结束时还没有失效。

表 6.6-3 截尾数据的例子

失效时间	
35	60 +
45	60 +
55	60 +
60 +	60 +

很明显，失效三项的样本平均值及样本标准偏差不能用来估算这个案例中的正态分布参数。这三个样本平均值为 (35 + 45 + 55) /3 = 45。存活下来的五项中每项的失效时间都大于 60；真实的样本均值要比 45 大的多。

截尾用很多种方式描述。

1）Ⅰ型截尾　一个试验最初有多个产品受测，试验在一段预定时间后结束。一个例子是，在试验台上安置 100 个晶体管，并且在 1 000 h 后结束测试，而不管失效数据的多少如何。这就是众所周知的定时截尾。在这个例子当中，试验时间是固定的，而失效数据数量为随机变量。

2）Ⅱ型截尾　一个试验最初有多个产品受测，试验在一定数目的产品失效后结束。一个例子是在测试位置安装 100 个晶体管，并且在产生了 30 个失效数据后终止检测，无论消逝时间是多少。这就是众所周知的定数截尾或截断。在这个例子中，失效数据的数目被锁定，而检测时间是随机变量。

以上的例子包含右截尾。如果失效的时间是未知的，那么部件会在右侧被截尾，已知的是该部件在试验结束时还没有失效。如果已知一个部件在具体的时间或之前处于一个失效的状态，但确切的失效时间是不知道的，那么就是左截尾。

3）单一截尾　单一截尾是只有一个截尾点。如果 100 个传感器置于一个检测台上，检测在 1 000 h 后完成，在 1 000 h 处有一个截尾点。如果在测试 1 000 h 后有 20 个没有失效，测试 1 200 h 后还有 15 个未失，这时就有 2 个截尾点，结果数据是多重截尾的。如果确切的失效时间未知，但是在

一个时间区间内失效的次数被记录到了，那么这就是区间数据或成组数据。

3.2 极大似然估计

极大似然估计法是一种被广泛使用的估计参数的方法。其基本原理是，通过使已知样本出现的概率最大化的方法来确定参数值。

一个给定分布的似然函数是一个已知样本数据的概率的表达式。令 x_1，x_2，…，x_n 是服从概率密度函数 $f(x,\theta)$ 的独立随机变量，其中，θ 是惟一的分布参数。那么

$$L(x_1,x_2,\cdots x_n;\theta)=f(x_1,\theta)f(x_2,\theta)\cdots f(x_n,\theta) \quad (6.6\text{-}68)$$

就是随机变量的联合分布或似然函数。

$\hat{\theta}$的极大似然估计，是使似然函数最大化的 θ 值。该估计是渐进正态的。为了简化计算，通常先将似然函数取自然对数，然后再求其极大值。

估计量的方差能够通过求似然函数二阶偏微分的负转置矩阵来获得，该矩阵也叫局部信息矩阵。这些估计量是渐进正态的，通过局部信息矩阵获得的变量可用于计算置信区间。

3.3 风险图

风险图是一种进行参数估计的绘图方法。累积风险函数通常通过对数变形被变为线性表达式并绘图，得到的截距和斜率提供了进行分布参数估计所需的信息。

如果人工构建一个风险图，则需要特殊分布风险纸。通过使用风险纸，失效时间和累积风险函数估计量能被直接描绘出来。由于电脑和电子拓展表格所带来的方便，不再需要特殊的描点纸也能迅速而简单地做出所需的变形了。

3.4 概率图

概率图法是参数估计中的一种图形方法。对于假定的分布，累积分布函数通常利用对数变换转换成线性形式，并描点绘图。如果描点形成了一条直线，这个假定的分布就是可接受的，图像的斜率和截距提供了估计分布参数所需的信息。中值秩通常用于估计累积分布函数，虽然还有其他选择，诸如平均秩和 kaplan-Meier 乘积极限估计量。

如果手工绘制一张概率图，那么还要有一张具体的概率分布纸。利用这张概率纸，失效时间和累积分布函数的估计值可以直接绘出，在计算机和电子绘图仪的帮助下，必要的转换能够快速轻松地实现，专用图纸已不再需要。

3.5 矩估计

矩估计是基于样本矩与母体分布及其参数定义。一阶原点矩等于分布的平均值。

$$E(x)=\begin{cases}\sum\limits_{x}xf(x) & \text{如果 } x \text{ 是离散的}\\ \int_{-\infty}^{\infty}xf(x)\mathrm{d}x & \text{如果 } x \text{ 是连续的}\end{cases} \quad (6.6\text{-}69)$$

对于分布均值的二阶矩等于分布的方差。

$$E[(x-\mu)]^2=\begin{cases}\sum\limits_{x}x^2f(x)-\mu^2 & \text{如果 } X \text{ 是离散的}\\ \int_{-\infty}^{\infty}x^2f(x)\mathrm{d}x-\mu^2 & \text{如果 } X \text{ 是连续的}\end{cases} \quad (6.6\text{-}70)$$

其中，μ 为分布均值。

对于一些分布，诸如正态分布，这些矩提供了参数的直接估计量。对于其他分布，诸如威布尔、对数正态和伽玛分布，分布参数需要通过让样本矩等于理论矩来解出分布参数。所需矩的数目是由被估计的参数的数目来决定的。如果分布只有一个参数，这个参数能通过解式（6.6-69）样本均值方程来估计并解出参数。如果有三个参数需要估计，就要三个方程。分布的不对称性，也就是均值的三阶矩。通常同以上两个方程共同使用。

3.6 指数分布参数估计

指数分布进行参数估计的最简单的方法是极大似然法。极大似然法提供一个无偏估计，但不表明相符合的程度。由于作图的方法经常能够直观地表达符合程度，因此就经常与极大似然估计法一起使用。

3.6.1 极大似然估计

指数分布概率密度函数是：

$$f(x)=\frac{1}{\theta}e^{-x/\theta},x\geqslant 0 \quad (6.6\text{-}71)$$

对于参数 θ 的极大似然估计是：

$$\hat{\theta}=\frac{\sum\limits_{i=1}^{n}x_i}{r} \quad (6.6\text{-}72)$$

式中，x_i 为第 i 个数据点，它可能是失效或截尾点；n 为包括截尾和未截尾点的数据点的总数；r 为失效的数目。

此估计是无偏估计和最小方差估计。

例4 7个弹簧失效时经历的载荷循环数分别是：

30183 14871 35031 76321 43891 31650 12310

假设失效的循环数服从指数分布，估计平均寿命和平均失效率。

解：

平均寿命是：

$$\hat{\theta}=\frac{30\,183+14\,871+35\,031+76\,321+43\,891+31\,650+12\,310}{7}$$

$$=\frac{244\,257}{7}=348\,93.9 \text{ 循环}$$

平均失效率是平均寿命的倒数：

$$\hat{\lambda}=\frac{1}{34\,893.9}=0.000\,028\,7 \text{ 次失效/循环}$$

例5 如果例4中的数据代表7根弹簧的失效循环数，但另有10根弹簧测试了80 000个载荷循环而未失效的。估计平均失效时间和平均失效率。

解：

平均失效时间是：

$$\hat{\theta}=\frac{244\,257+10\times(80\,000)}{7}=149\,179.6 \text{ 周期}$$

平均失效率是：

$$\lambda=\frac{1}{149\,179.6}=0.000\,006\,7 \text{ 次失效/周期}$$

对于时间截尾实验，θ 的置信区间是：

$$\frac{2\sum\limits_{i=1}^{n}x_i}{\chi^2_{(\alpha/2,2r+2)}}\leqslant\theta\leqslant\frac{2\sum\limits_{i=1}^{n}x_i}{\chi^2_{(1-\alpha/2,2r)}} \quad (6.6\text{-}73)$$

注意，χ^2 上限与下限的自由度数不同。

例6 15个产品测试1 000 h。失效分别在第120、190、560、812 h发生。试为平均失效时间和平均失效率构建一个90%的置信区间。

解：

这是一个时间截尾实验，估计的平均寿命是：

$$\hat{\theta}=\frac{120+190+560+812+11\times 1\,000}{4}=\frac{12\,682}{4}=3\,170.5$$

当置信区间是90%时，$\alpha=0.1$，由 χ^2 分布表（附表2）知，$\chi^2_{(0.05,10)}=18.307$，$\chi^2_{(0.95,8)}=2.733$。则 θ 的90%置信区

间是：

$$\frac{2\times 12\,682}{18.307}\leqslant\theta\leqslant\frac{2\times 12\,682}{2.733}$$

$$1\,385.5\leqslant\theta\leqslant 9\,280.6$$

给定置信区间内的失效率是相同置信区间内平均失效时间的倒数，

$$\frac{1}{9\,280.6}\leqslant\lambda\leqslant\frac{1}{1\,385.5}$$

$$0.000\,107\,7\leqslant\lambda\leqslant 0.000\,721\,7$$

对于定数截尾实验，θ 的置信区间是：

$$\frac{2\sum_{i=1}^{n}x_i}{\chi^2_{(\alpha/2,2r)}}\leqslant\theta\leqslant\frac{2\sum_{i=1}^{n}x_i}{\chi^2_{(1-\alpha/2,2r)}} \tag{6.6-74}$$

注意，χ^2 上限与下限的自由度数相同。

例 7 12 个个体进行测试，失效时间分别发生在第 43、67、92、94、149 h，在第 149 h，剩余的 7 个对象停止测试，试为平均失效时间构建一个 95%的置信区间。

解 这是一个定数截尾试验，估计的平均寿命是：

$$\hat{\theta}=\frac{43+67+92+94+149+7\times 149}{5}=\frac{1\,488}{5}=297.6$$

当置信区间是 95% 时，$\alpha=0.05$，由附表 2 查知，$\chi^2_{(0.025,10)}=20.483$，$\chi^2_{(0.975,10)}=3.247$。则 θ 的 95% 置信区间是：

$$\frac{2\times 1\,488}{20.483}\leqslant\theta\leqslant\frac{2\times 1\,488}{3.247}$$

$$145.3\leqslant\theta\leqslant 916.5$$

对于无失效测试，单边置信下限简化为：

$$\frac{-nt}{\ln\alpha}\leqslant\theta \tag{6.6-75}$$

式中，t 是测试时间；ln 是自然对数符号；α 是显著性（95%限制时，$\alpha=0.05$）。

例 8 20 个样本进行 230 h 的测试无失效，当下置信限为 90%时，确定 θ 的值。

解 $$\frac{-20\times(230)}{\ln(0.1)}=1\,997.8\text{ h}$$

对于可靠性的置信区间是：

$$e^{-\frac{x}{\theta_L}}\leqslant R(x)\leqslant e^{-\frac{x}{\theta_U}} \tag{6.6-76}$$

式中，θ_L 是平均失效时间的置信下限；θ_U 是平均失效时间的置信上限。

置信区间的百分位是：

$$-\theta_L\ln(1-P)\leqslant x\leqslant-\theta_U\ln(1-P) \tag{6.6-77}$$

式中，P 是时间达到 x 之前失效的概率。

3.6.2 风险图

指数累积风险函数是：

$$H(x)=\frac{x}{\theta} \tag{6.6-78}$$

如果数据服从指数分布，那么根据式（6.6-78）描点，就能得到一个截距是 0，斜度为 $1/\theta$ 的直线。为了建立一个风险图，需要对 $H(x)$ 进行估计。对于一个包含 n 个点的从小到大排列的数列来说，第一点的排序是 n，第二个是 $n-1$ 等。

例 9 用例 7 中的数据描绘一条风险曲线。

解 表 6.6-4 计算了描绘风险曲线所需的值。曲线如图 6.6-18 所示。注意，截尾数据点并没有描出。最适合直线的截距是 0，斜率是 0.003 33。估计的均值是：$\theta=1/0.003\,33=300.3$。

表 6.6-4 对指数风险坐标点的列表

失效时间/h	转秩	$h(t)$	$H(t)$
43	12	0.083 3	0.083 3
67	11	0.090 9	0.174 2
92	10	0.100 0	0.274 2
94	9	0.111 1	0.385 4
149	8	0.125 0	0.510 4
149c	7		
149c	6		
149c	5		
149c	4		
149c	3		
149c	2		
149c	1		

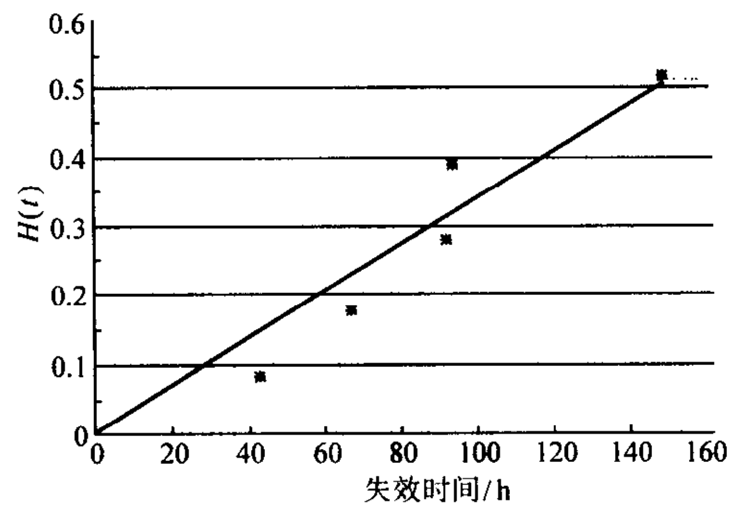


图 6.6-18 指数数据的风险图

3.6.3 概率图

指数累积分布函数是：

$$F(x)=1-e^{-\frac{x}{\theta}} \tag{6.6-79}$$

通过变形，上式能够变为线性表达式：

$$\ln\frac{1}{1-F(x)}=\frac{x}{\theta} \tag{6.6-80}$$

如果一组数据服从指数分布，则 $\ln\frac{1}{1-F(x)}$ 与 x 的关系曲线将会是以 0 为截距，$1/\theta$ 为斜率的直线。在进行绘图以前，需要对 $F(x)$ 进行估计。累积分布函数 $F(x)$ 经常通过中位秩进行估计，但是也可使用其他的估计，如平均秩或 Kaplan-Meier 乘积极限估计。中位秩对 $F(x)$ 进行估计是：

$$\hat{F}(x)=\frac{O_i-0.3}{n+0.4} \tag{6.6-81}$$

式中，O_i 是第 i 个数据点的次序数。

当包含截尾数据时，需要对失效次序进行调整；如果不包含截尾数据，那么原始的失效顺序 i 与调整的失效顺序是等价的。调整失效次序的步骤如下：考虑三个个体；第二个测试了 3 h，没有失效时测试就停止了，第二个进行测试并在 4 h 时失效，第三个进行测试，4.5 h 时失效。对这些数据，失效次序是不明确的。一个个体可能是第一个失效、第二个失效或第三个失效者。对失效次序进行调整是由以下表达式计算而来的：

$$O_i=O_{i-1}+I_i \tag{6.6-82}$$

式中，I_i 是第 i 次失效的增量，由表达式：

$$I_i = \frac{(n+1) - O_p}{1+c} \tag{6.6-83}$$

计算得来。其中，n 是包括截尾和未截尾的所有数据点的总数；c 是包括当前点的保留在数据中的点数；O_p 是前一个失效的次序。

关于 x 的曲线也可以用下式替换：

$$\ln \frac{1}{1-F(x)}$$

传统上是在专用概率纸上绘 $F(x)$ 与 x 的关系图。概率纸的优点是 $\ln\frac{1}{1-F(x)}$ 不必计算。电子数据表内置的统计函数能迅速对数据进行变形，使用线性回归的方法确定最合适的直线，并将结果绘制成图。

例 10　用例 7 中的数据构建一个概率图。

解　表 6.6-5 包含了计算所需的坐标点。概率坐标曲线如图 6.6-19 所示。通过原点的估计指数分布失效率的最合适直线的斜率是 0.003 04。分布的均值是：$\theta = 1/0.003\ 04 = 328.9$。

表 6.6-5　对指数概率坐标点的列表

失效时间/h	I_i	O_i	中值序列 $F(t)$	$\frac{1}{1-F(x)}$	$\ln\frac{1}{1-F(x)}$
43	1	1	0.056 5	1.059 8	0.058 1
67	2	2	0.137 1	1.158 9	0.147 5
92	3	3	0.217 7	1.278 4	0.245 6
94	4	4	0.298 4	1.425 3	0.354 4
149	5	5	0.379 0	1.610 4	0.476 5
149c					
149c					
149c					
149c					
149c					
149c					
149c					

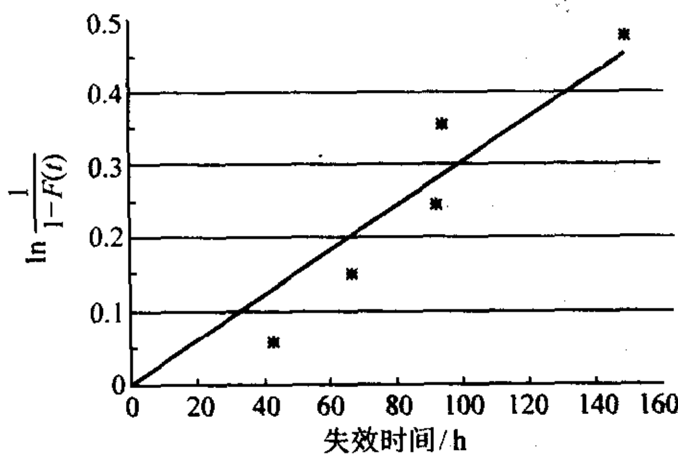


图 6.6-19　指数概率图

通常，可靠性置信限被加在概率图上。上、下置信限用 5%和 95%的秩近似。这些秩能从秩表中得到，或用以下表达式计算：

$$w_\alpha = \frac{\frac{j}{n-j+1}}{F_{1-\alpha,2(n-j+1),2j} + \frac{j}{n-j+1}}, \quad \alpha \leqslant 0.5 \tag{6.6-84}$$

$$w_\alpha = \frac{\frac{j}{n-j+1}F_{\alpha,2j,2(n-j+1)}}{1+\frac{j}{n-j+1}F_{\alpha,2j,2(n-j+1)}}, \quad \alpha > 0.5 \tag{6.6-85}$$

式中，w_α 是 $100(1-\alpha)$%非参数置信极限；j 是失效次序；n 是包括截尾和未截尾的数据点的总数；$F_{\alpha,v1,v2}$ 是来自 F 分布的临界值。

当遇见多重截尾数据时，修正的失效次序可能不是整数，秩值必须用插值的方法获得。秩值不是对应于相应的失效时间描绘出的。任何失效时间与通过数据得到的最佳直线的偏差被认为是抽样误差，秩值对应的时间是通过平移 x 轴，直到和最适合的直线相交而得到。这个绘图位置是：

$$x_i = \theta \ln\left[\frac{1}{1-F(x_i)}\right] \tag{6.6-86}$$

例 11　用前面包含置信限的例子描绘一个概率图。

解　表 6.6-6 显示了 5%和 95%秩以及其对前面例子的绘图位置。对本例的包括置信极限的概率图如图 6.6-20 所示。

表 6.6-6　包括置信限的指数概率图数据的列表

失效时间/h	5%序列	$\ln\frac{1}{1-5\%序列}$	95%序列	$\ln\frac{1}{1-95\%序列}$	描点位置
43	0.004 3	0.004 3	0.220 9	0.249 6	19.1
67	0.030 5	0.031 0	0.338 7	0.413 5	48.5
92	0.071 9	0.074 6	0.438 1	0.576 4	80.8
94	0.122 9	0.1311	0.527 3	0.749 3	116.6
149	0.181 0	0.199 7	0.609 1	0.939 3	156.7
149c					
149c					
149c					
149c					
149c					
149c					
149c					

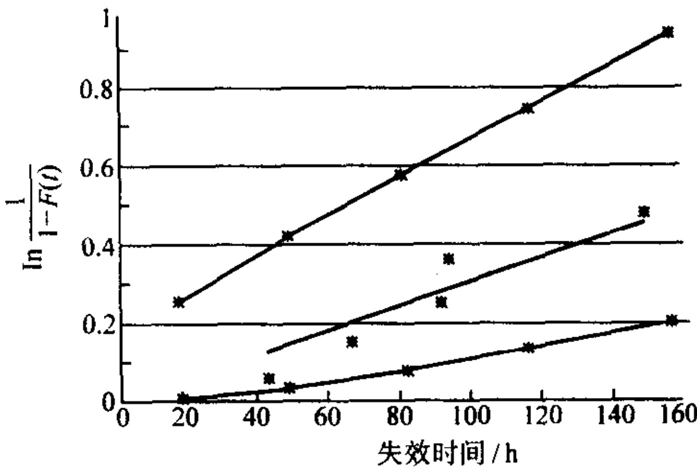


图 6.6-20　用置信极限进行指数概率描点

3.7　正态分布

正态概率密度函数是：

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], -\infty < x < \infty \tag{6.6-87}$$

式中，μ 是分布的均值；σ 是分布的标准差。

如果没有进行截尾,那么分布的均值由表达式:

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad (6.6\text{-}88)$$

估计。其中,n 是样本容量。

如果没有进行截尾,那么分布的标准差由表达式:

$$\hat{\sigma} = \sqrt{\frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}} \quad (6.6\text{-}89)$$

然而,当有截尾数据时,参数的估计就变得复杂了。当有截尾时,对正态分布的三种常用的参数估计方法是:①极大似然估计;②风险图;③概率图。以下对三种方法分别介绍。

3.7.1 极大似然估计

正态分布的极大似然估计方程是:

$$\frac{\partial L}{\partial \mu} = \frac{r}{\sigma}\left[\frac{\bar{x}-\mu}{\sigma} + \sum_{i=1}^{k}\frac{h(x_i)}{r}\right] = 0 \quad (6.6\text{-}90)$$

$$\frac{\partial L}{\partial \sigma} = \frac{r}{\sigma}\left[\frac{s^2 + (\bar{x}-\mu)^2}{\sigma^2} - 1 + \sum_{i=1}^{k}\frac{z(x_i)h(x_i)}{r}\right] = 0 \quad (6.6\text{-}91)$$

式中,r 是失效的次数;k 是截尾的次数;$\bar{x}$ 是失效的样本均值;s 是失效的样本标准差;$z(x_i)$ 是标准偏量,$z(x_i)=\frac{x_i-\mu}{\sigma}$;$h(x_i)$ 是在第 i 点的风险函数的估计:

$$h(x_i) = \frac{\phi[z(x_i)]}{\sigma[1-\Phi(z(x_i))]}$$

其中,$\phi[z(x_i)]$ 是在第 i 点估计的标准正态概率密度函数;$\Phi(z(x_i))$ 是在第 i 点估计的标准正态累积分布函数。

注意,如果没有截尾,那么这些表达式将还原为样本均值和样本标准偏差。

要解方程式(6.6-90)和式(6.6-91),需要使用迭代方法。迭代方法是一个基于 Taylor 级数展开的标准方法,它包括反复估计参数,直到达到所要求的精确度。用以下表达式估计:

$$\hat{\mu}_j = \hat{\mu}_{j-1} + h \quad (6.6\text{-}92)$$

$$\hat{\sigma}_j = \hat{\sigma}_{j-1} + k \quad (6.6\text{-}93)$$

式中,h 是分布均值的纠正系数;k 是分布标准偏差的纠正系数。

对于每次重复,纠正系数由以下表达式估计得到:

$$h\frac{\partial^2 L}{\partial \mu^2} + k\frac{\partial^2 L}{\partial \mu \partial \sigma} = -\frac{\partial L}{\partial \mu} \quad (6.6\text{-}94)$$

$$h\frac{\partial^2 L}{\partial \mu \partial \sigma} + k\frac{\partial^2 L}{\partial \sigma^2} = -\frac{\partial L}{\partial \sigma} \quad (6.6\text{-}95)$$

其中

$$\frac{\partial^2 L}{\partial \mu^2} = -\frac{r}{\sigma^2}\left[1 + \sum_{i=1}^{k}\frac{A_i}{r}\right] \quad (6.6\text{-}96)$$

$$\frac{\partial^2 L}{\partial \mu \partial \sigma} = -\frac{r}{\sigma^2}\left[\frac{2(\bar{x}-\mu)}{\sigma} + \sum_{i=1}^{k}\frac{B_i}{r}\right] \quad (6.6\text{-}97)$$

$$\frac{\partial^2 L}{\partial \sigma^2} = -\frac{r}{\sigma^2}\left\{\frac{3[s^2 + (\bar{x}-\mu)^2]}{\sigma^2} + \sum_{i=1}^{k}\frac{C_i}{r}\right\} \quad (6.6\text{-}98)$$

且,$A_i = h(x_i)[h(x_i) - z(x_i)]$

$B_i = h(x_i) + z(x_i)A_i$

$C_i = z(x_i)[h(x_i) + B_i]$

估计的参数是渐近正态的。估计的方差可以通过求局部信息矩阵的逆来得到。

$$F = \begin{bmatrix} -\frac{\partial^2 L}{\partial \mu^2} & -\frac{\partial^2 L}{\partial \mu \partial \sigma} \\ -\frac{\partial^2 L}{\partial \mu \partial \sigma} & -\frac{\partial^2 L}{\partial \sigma^2} \end{bmatrix} \quad (6.6\text{-}99)$$

求逆后,方差为:

$$F^{-1} = \begin{bmatrix} \text{var}(\hat{\mu} & \text{cov}(\hat{\mu},\hat{\sigma}) \\ \text{cov}(\hat{\mu},\hat{\sigma}) & \text{var}(\hat{\sigma}) \end{bmatrix} \quad (6.6\text{-}100)$$

被估计参数的大约 100%(1-α)的置信区间是:

$$\hat{\mu} - K_{\alpha/2}\sqrt{\text{var}(\hat{\mu})} \leqslant \mu \leqslant \hat{\mu} + K_{\alpha/2}\sqrt{\text{var}(\hat{\mu})} \quad (6.6\text{-}101)$$

$$\frac{\hat{\sigma}}{\exp\left[\frac{K_{\alpha/2}\sqrt{\text{var}(\hat{\sigma})}}{\hat{\sigma}}\right]} \leqslant \sigma \leqslant \hat{\sigma}\exp\left[\frac{K_{\alpha/2}\sqrt{\text{var}(\hat{\sigma})}}{\hat{\sigma}}\right] \quad (6.6\text{-}102)$$

式中,$K_{\alpha/2}$ 是标准正态概率密度函数的反函数。

这些置信区间都是大约的,但是它们随着样本容量的增大而接近于精确。

可靠性的置信区间能够通过使用以下表达式获得:

$$\text{var}(\hat{z}) \approx \left(\frac{\text{var}(\hat{\mu}) + \hat{z}^2\text{var}(\hat{\sigma}) + 2\hat{z}\text{cov}(\hat{\mu},\hat{\sigma})}{\hat{\sigma}^2}\right) \quad (6.6\text{-}103)$$

$$\hat{z} - K_{\alpha/2}\sqrt{\text{var}(\hat{z})} \leqslant z \leqslant \hat{z} + K_{\alpha/2}\sqrt{\text{var}(\hat{z})} \quad (6.6\text{-}104)$$

$$1-\Phi(\hat{z}+K_{\alpha/2}\sqrt{\text{var}(\hat{z})}) \leqslant R(x) \leqslant 1-\Phi(\hat{z}-K_{\alpha/2}\sqrt{\text{var}(\hat{z})}) \quad (6.6\text{-}105)$$

百分位点的置信区间能够通过以下表达式求得:

$$\text{var}(\hat{x}) \approx \text{var}(\hat{\mu}) + \hat{z}^2\text{var}(\hat{\sigma}) + 2\hat{z}\text{cov}(\hat{\mu},\hat{\sigma}) \quad (6.6\text{-}106)$$

$$\hat{x} - K_{\alpha/2}\sqrt{\text{var}(\hat{x})} \leqslant x \leqslant \hat{x} + K_{\alpha+2}\sqrt{\text{var}(\hat{x})} \quad (6.6\text{-}107)$$

极大似然估计很复杂,经常使用计算机进行程序化的计算。一些商用软件可用于此类计算,如由 Dodson 和 Kirlnd(1994年)开发的软件。

3.7.2 风险图

正态累积风险函数是:

$$H(x) = -\ln\left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right] \quad (6.6\text{-}108)$$

式中,$\Phi(x)$ 是标准正态累积分布函数。

通过对方程式(6.6-108)进行变形,存活时间能够用累积风险函数,表示:

$$x = \mu + \sigma\Phi^{-1}\left[1-e^{-H(x)}\right] \quad (6.6\text{-}109)$$

式中,$\Phi^{-1}(x)$ 是标准正态累积分布函数的反函数。

通过绘制 x 与 $\Phi^{-1}[1-e^{-H(x)}]$ 的关系曲线可以看到,y 轴的截距等于 μ,斜率等于 σ。风险函数 $h(x)$ 由顺序失效量的翻转序列的反函数估计得来,$H(x)$ 是 $h(x)$ 值的累积。截尾的数据点用于计算秩,但不用于风险图。

另一种做法是在专用的风险座标纸上绘制 x 与 $H(x)$ 的关系曲线。风险纸的优点是不必计算 $\Phi^{-1}[1-e^{-H(x)}]$ 的值。然而,由于现在可以方便地应用电子数据表,风险纸的缺点就比其优点更突出了。

用电子数据表借助其内置的统计函数能迅速地进行数据变换,用线性回归的方法确定一条最佳的直线,并将结果绘制出来。

例12 使用风险图法应用多重截尾数据确定正态分布的参数。以“c”结尾的表示的是截尾的数据。

失效时间

150c	183c	235
157c	209	235
167c	216c	248c
179	217c	257

解 建立表6.6-7是为了获得描绘点所需要的数据。

表 6.6-7　正态分布风险图所用数据的列表

失效时间	逆秩	$h(t)$	$H(t)$	$1-e^{-H(t)}$	$\Phi^{-1}[1-e^{-H(t)}]$
150c	12				
157c	11				
167c	10				
179	9	0.111 1	0.111 1	0.105 2	-1.252 7
183c	8				
209	7	0.142 9	0.254 0	0.224 3	-0.757 8
216c	6				
217c	5				
235	4	0.250 0	0.504 0	0.395 9	-0.264 0
235	3	0.333 3	0.837 3	0.567 1	0.169 1
248c	2				
257	1	1.000 0	1.837 3	0.840 8	0.997 6

现在，五个失效时间能够用其与 $\Phi^{-1}[1-e^{-H(x)}]$ 对应的关系进行描点。

如图 6.6-21 所示，使用线性回归的方法能够找到穿过数据点的最适合曲线。拟合曲线的 y 截距对应本例中的 $\mu=230.3$ 进行了估计，直线的斜率给出 $\sigma=32.9$ 的估计值。

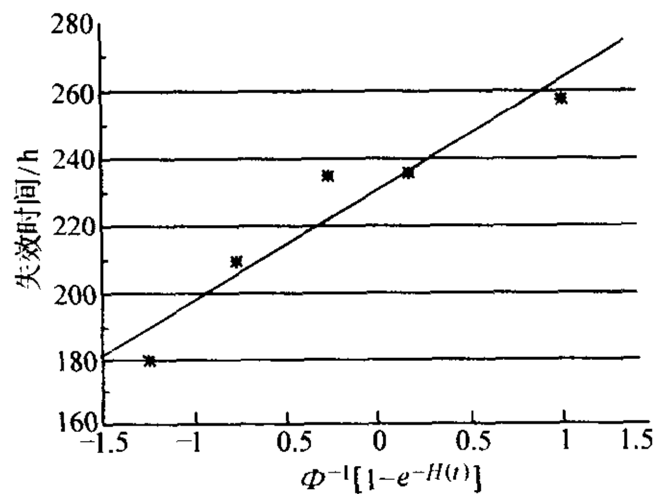


图 6.6-21　正态分布风险图

3.7.3　概率图法

通过对正态累积分布函数进行变形，能够获得一个表达式，即：

$$x=\mu+\sigma\Phi^{-1}[F(x)] \tag{6.6-110}$$

式中，$F(x)$ 是正态累积分布函数；$\Phi^{-1}(x)$ 是标准正态累积分布函数的反函数。

通过 x 与 $\Phi^{-1}[F(x)]$ 的关系曲线可以看到，y 轴的截距等于 μ，斜率等于 σ。累积分布函数 $F(x)$ 经常通过中位秩的估计得来，但也使用其他的估计方法，如平均秩和 Kaplan-Meier 极限估计法。中位秩使用本章 3.6.3 中的式 (6.6-82) ~式 (6.6-84) 进行估计。

对 x 与 $\Phi^{-1}[F(x)]$ 之间关系的一种替换是在专用概率纸上绘制 x 与 $F(x)$ 的关系。概率纸的优点是不必计算 $\Phi^{-1}[F(x)]$ 的值。然而，应用电子数据表及其内置的统计函数能迅速地进行数据变形，用线性回归的方法确定一条最佳的直线，并将结果绘制出来。

例 13　使用概率图法在给定以下数据的情况下确定正态分布的参数。以“c”结尾的表示的是截尾的数据。

失效时间

150c	183c	235
157c	209	235
167c	216c	248c
179	217c	257

解　表 6.6-8 是为绘制数据所建立的表格。数据被描绘在图 6.6-22 上，估计的斜率 σ 的值是 34.8，μ 的值是 235.3。

表 6.6-8　正态分布概率描点所用数据的列表

失效时间	I_i	O_i	中位秩 $F(x)$	$\Phi^{-1}[F(x)]$
150c				
157c				
167c				
179	1.300 0	1.300 0	0.080 6	-1.400 8
183c				
209	1.462 5	2.762 5	0.198 6	-0.846 7
216c				
217c				
235	2.047 5	4.810 0	0.363 7	-0.348 6
235	2.047 5	6.857 5	0.528 8	0.072 3
248c				
257	3.071 3	9.928 8	0.776 5	0.760 5

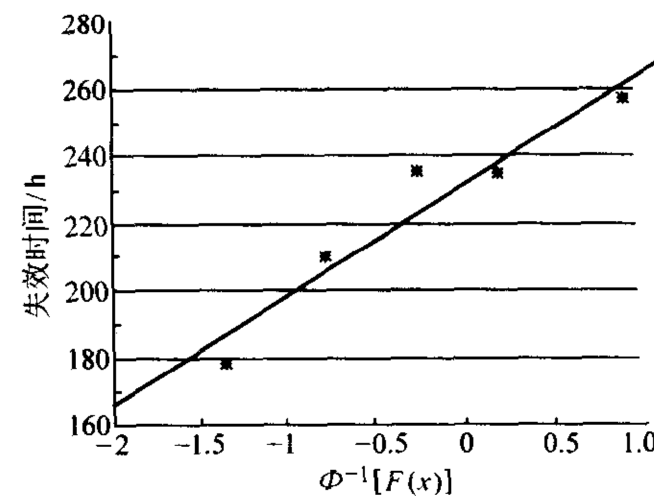


图 6.6-22　正态概率描点

可靠性的非参数置信极限可以以 5% 和 95% 的秩加在正态概率图中，这点已经在本章 3.6.3 中做了解释。5% 和 95% 的标准正态反函数的绘图位置是：

$$x=\mu+\sigma\Phi^{-1}[F(x)] \tag{6.6-111}$$

注意，$F(x)$ 是中位秩。

例 14　为前一个例子构建一个包括置信极限的概率图。

解　表 6.6-9 包括了获得置信极限的计算值以及它们的绘图位置。

表 6.6-9　正态概率描点所用的置信极限计算值

失效时间	失效次序	中位秩	$\Phi^{-1}[F(x)]$	5%秩	插值的5%秩	Φ^{-1} 5%秩	95%秩	插值的95%秩	Φ^{-1} 95%秩	绘图位置
	1			0.004 3			0.220 9			
179	1.3	0.080 6	-1.400 8		0.012 1	-2.253 3		0.256 2	-0.655 0	186.552 16

续表 6.6-9

失效时间	失效次序	中位秩	$\Phi^{-1}[F(x)]$	5%秩	插值的5%秩	Φ^{-1} 5%秩	95%秩	插值的95%秩	Φ^{-1} 95%秩	绘图位置
	2			0.030 5			0.338 7			
209	2.762 5	0.198 6	-0.846 7		0.062 0	-1.537 9		0.414 5	-0.216 0	205.834 84
	3			0.071 9			0.438 1			
	4			0.122 9			0.527 3			
235	4.81	0.363 7	-0.348 6		0.170 0	-0.954 3		0.593 6	0.236 8	223.168 73
	5			0.181 0			0.609 1			
	6			0.245 3			0.684 8			
253	6.857 5	0.528 8	0.072 3		0.305 3	-0.509 3		0.744 7	0.658 0	237.816 04
	7			0.315 2			0.754 7			
	9			0.472 7			0.819 0			
257	9.928 8	0.776 5	0.760 5		0.555 5	0.139 7		0.920 4	1.407 5	261.765 4
	10			0.561 9			0.928 1			

包括置信极限的概率图如图 6.6-23 所示。

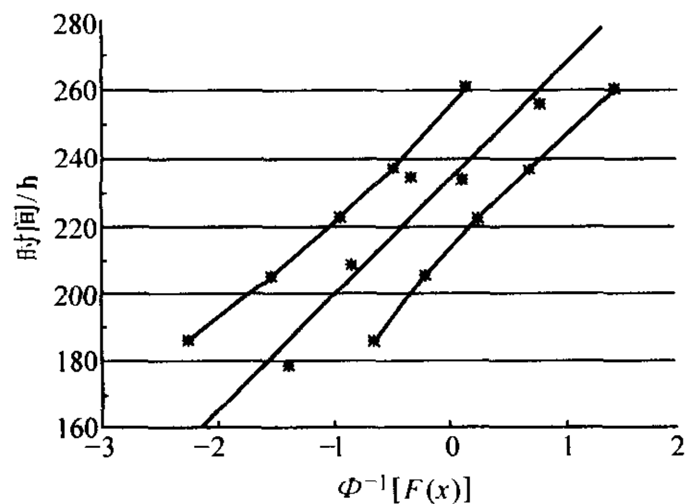


图 6.6-23 用置信极限进行正态概率图

3.8 对数正态分布

对数正态概率密度函数是：

$$f(x)=\frac{1}{\sigma x\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right],\quad x>0 \tag{6.6-112}$$

式中，μ 是位置参数；σ 是形状参数。

如果 x 是一个服从对数正态分布的随机变量，那么 $y=\ln(x)$ 就是一个正态分布随机变量。位置参数等于数据点对数的均值，形状参数等于数据点对数的标准偏差。因此，并不需要像处理一个单独的分布那样处理对数正态分布。通过取得数据点的对数，即在前面几节讲到的处理正态分布的方法，在这里能够被用于估计对数正态分布的参数。

3.9 威布尔分布

威布尔概率密度函数是：

$$f(x)=\frac{\beta x^{\beta-1}}{\theta^{\beta}}\exp\left(\frac{x}{\theta}\right)^{\beta},\quad x\geqslant 0 \tag{6.6-113}$$

式中，β 是形状参数；θ 是比例参数。

有的时候，三参数威布尔分布比两参数威布尔分布具有更好的适应性。两分布的不同在于位置参数 α，随着 α 值的不同，分布沿 x 轴变换。根据定义，当 $x<\alpha$ 时，失效的概率是 0。尽管位置取得负值是异常的，这说明个体在测试之前就已经失效了。三参数威布尔分布是：

$$f(x)=\frac{\beta(x-\alpha)^{\beta-1}}{\theta^{\beta}}\exp\left(\frac{x-\alpha}{\theta}\right)^{\beta},\quad x\geqslant\alpha \tag{6.6-114}$$

有四种对威布尔分布的参数进行估计的方法，分别是：①极大似然估计；②风险坐标法；③概率坐标法；④配合矩法。

3.9.1 极大似然估计

威布尔分布的极大似然估计方程是：

$$\frac{1}{r}\sum_{i=1}^{r}\ln(x_i)=\left[\sum_{i=1}^{n}x_i^{\beta}\ln(x_i)\right]\left[\sum_{i=1}^{n}x_i^{\beta}\right]^{-1}-\frac{1}{\beta} \tag{6.6-115}$$

$$\hat{\theta}=\left[\frac{1}{r}\sum_{i=1}^{n}x_i^{\hat{\beta}}\right]^{1/\hat{\beta}} \tag{6.6-116}$$

式中，r 是失效的次数；n 是包括检测和未检测点的数据点的总数。

解方程式（6.6-115）需要用反复的方法。被估计的参数没有正态特征。估计的方差能够通过求自身信息矩阵的逆来得到，即：

$$F=\begin{bmatrix}-\dfrac{\partial^2 L}{\partial\beta^2} & -\dfrac{\partial^2 L}{\partial\beta\partial\theta}\\ -\dfrac{\partial^2 L}{\partial\beta\partial\theta} & -\dfrac{\partial^2 L}{\partial\theta^2}\end{bmatrix} \tag{6.6-117}$$

对数似然方程的第二部分偏差是：

$$\frac{\partial^2 L}{\partial\beta^2}=\sum_r\left[-\frac{1}{\beta}-\left(\frac{x_i}{\theta}\right)^{\beta}\ln^2\left(\frac{x_i}{\theta}\right)\right]+\sum_k\left[-\left(\frac{x_i}{\theta}\right)^{\beta}\ln^2\left(\frac{x_i}{\theta}\right)\right] \tag{6.6-118}$$

$$\frac{\partial^2 L}{\partial\theta^2}=\sum_r\left[\frac{\beta}{\theta^2}-\left(\frac{x_i}{\theta}\right)^{\beta}\left(\frac{\beta}{\theta^2}\right)(\beta+1)\right]+\sum_k\left[-\left(\frac{x_i}{\theta}\right)^{\beta}\left(\frac{\beta}{\theta^2}\right)(\beta+1)\right] \tag{6.6-119}$$

$$\frac{\partial^2 L}{\partial\beta\partial\theta}=\sum_r\left\{-\frac{1}{\theta}+\left(\frac{x_i}{\theta}\right)^{\beta}\left(\frac{1}{\theta}\right)\left[\beta\ln\left(\frac{x_i}{\theta}\right)+1\right]\right\}+\sum_k\left\{\left(\frac{x_i}{\theta}\right)^{\beta}\left(\frac{1}{\theta}\right)\left[\beta\ln\left(\frac{x_i}{\theta}+1\right)\right]\right\} \tag{6.6-120}$$

式中，$\sum_r$ 代表所有失效的总和；$\sum_k$ 代表所有测试点的总和。

被估计参数的方差是：

$$F^{-1}=\begin{bmatrix}\mathrm{var}(\hat{\beta}) & \mathrm{cov}(\hat{\beta},\hat{\theta})\\ \mathrm{cov}(\hat{\beta},\hat{\theta}) & \mathrm{var}(\hat{\theta}\end{bmatrix} \tag{6.6-121}$$

被估计参数约100%（1 - α）的置信区间是：

$$\frac{\hat{\beta}}{\exp\left(\frac{K_{\alpha/2}\sqrt{\mathrm{var}(\hat{\beta})}}{\hat{\beta}}\right)} \leqslant \beta \leqslant \hat{\beta} + \hat{\beta}\exp\left(\frac{K_{\alpha/2}\sqrt{\mathrm{var}(\hat{\beta})}}{\hat{\beta}}\right) \quad (6.6\text{-}122)$$

$$\frac{\hat{\theta}}{\exp\left(\frac{K_{\alpha/2}\sqrt{\mathrm{var}(\hat{\theta})}}{\hat{\theta}}\right)} \leqslant \hat{\theta} + \hat{\theta}\exp\left(\frac{K_{\alpha/2}\sqrt{\mathrm{var}(\hat{\theta})}}{\hat{\theta}}\right) \quad (6.6\text{-}123)$$

式中，$K_{\alpha/2}$是标准正态概率密度函数的倒数。

这些置信区间是近似的，但其随着样本容量的增大而接近精确。

可靠性的置信区间能够通过以下表达式求得：

$$\exp[-\exp(u + K_{\alpha/2}\sqrt{\mathrm{var}(u)})] \leqslant R(x) \leqslant \exp[-\exp(u - K_{\alpha/2}\sqrt{\mathrm{var}(u)})] \quad (6.6\text{-}124)$$

$$u = \beta\,[\ln(x) - \ln(\theta)] \quad (6.6\text{-}125)$$

$$\mathrm{var}(\hat{u}) \approx \beta^2\left[\left(\frac{\mathrm{var}(\hat{\theta})}{\theta^2}\right) + \left(\frac{\hat{u}\mathrm{var}(\hat{\beta})}{\beta^4}\right) - \left(\frac{2\hat{u}\mathrm{cov}(\hat{\beta},\hat{\theta})}{\hat{\beta}^2\hat{\theta}}\right)\right] \quad (6.6\text{-}126)$$

百分数性质的置信区间能够通过以下表达式求得：

$$e^{y_L} \leqslant \hat{x} \leqslant e^{y_U} \quad (6.6\text{-}127)$$

$$\hat{x} = \hat{\theta}\,[-\ln(1-p)]^{1/\hat{\beta}} \quad (6.6\text{-}128)$$

$$y_L = \ln\hat{\theta} + \frac{\ln[-\ln(1-p)]}{\hat{\beta}} - K_\alpha\sqrt{\mathrm{var}(\hat{y})} \quad (6.6\text{-}129)$$

$$y_U = \ln\hat{\theta} + \frac{\ln[-\ln(1-p)]}{\hat{\beta}} + K_\alpha\sqrt{\mathrm{var}(\hat{y})} \quad (6.6\text{-}130)$$

$$\mathrm{var}(\hat{y}) = \frac{\mathrm{var}(\hat{\theta})}{\hat{\theta}^2} + \frac{\{\ln[-\ln(1-p)]\}^2\mathrm{var}(\hat{\beta})}{\beta^4} - \frac{2\{\ln[-\ln(1-p)]\}\mathrm{cov}(\hat{\theta},\hat{\beta})}{\hat{\beta}^2\hat{\theta}} \quad (6.6\text{-}131)$$

极大似然估计很复杂，经常使用计算机进行程序化的计算。一些商用软件可用于此类计算，如由 Dodson 和 Kirland（1994年）开发的软件。

3.9.2　风险坐标法

威布尔累积风险函数是：

$$H(x) = -\ln[1 - F(x)] \quad (6.6\text{-}132)$$

替换掉 $F(x)$ 并变形，可得到一个线性表达式：

$$\ln H(x) = \beta\ln x - \beta\ln\theta \quad (6.6\text{-}133)$$

通过对 $\ln H(x)$ 对应的 $\ln x$ 进行描点得到的斜度可对 β 进行估计。这个描点在 y 轴截距可以对 $\beta\ln\theta$ 进行估计。因此 θ 可以由表达式

$$\hat{\theta} = \exp\left(-\frac{y_0}{\hat{\beta}}\right) \quad (6.6\text{-}134)$$

进行估计，其中，y_0 是风险坐标曲线的 y 轴截距。

风险函数 $h(x)$ 由顺序失效翻转序列的逆估计得来；累积风险函数 $H(x)$ 是 $h(x)$ 值的累积。可以在专门的威布尔风险坐标纸上直接用 $H(x)$ 对应的 x 替换 $\ln H(x)$ 对应的 $\ln x$。风险纸的优点是不必进行对数变形的计算。对电子扩展表格来说，这些变形容易计算，并不需要风险纸。电子扩展表格也能够通过数据构建一条适合的直线，并将其容易地打印出来。

例15　使用以下表格中的多个测量数据，确定威布尔分布的参数。以“c”结尾表明是截尾的数据。

失效时间

309c	229
386	104c
180	217c
167c	168
122	138

解　建立表6.6-10是为了获得所需的描点数据。

表6.6-10　威布尔分布风险描点列表

失效时间	翻转序列	$h(t)$	$H(t)$	$\ln H(t)$	$\ln t$
104	10				
122	9	0.111 1	0.111 1	-2.197 2	4.804 0
138	8	0.125 0	0.236 1	-1.443 5	4.927 3
167c	7				
168	6	0.166 7	0.402 8	-0.909 4	5.124 0
180	5	0.200 0	0.602 8	-0.506 2	5.193 0
217c	4				
229	3	0.333 3	0.936 1	-0.066 0	5.433 7
309c	2				
386	1	1.000 0	1.936 1	0.660 7	5.955 8

表中最后两列的数据能被用于描点。所描的图像如图6.6-24所示。通过数据求得（用线性回归的方法）最佳直线的斜率等于2.34，对 β 进行了估计。通过数据得到的最佳直线 y 轴的截距是 -13.004。对威布尔分布比例参数的估计是：

$$\hat{\theta} = \exp\left(-\frac{-13.004}{2.34}\right) = 259.1$$

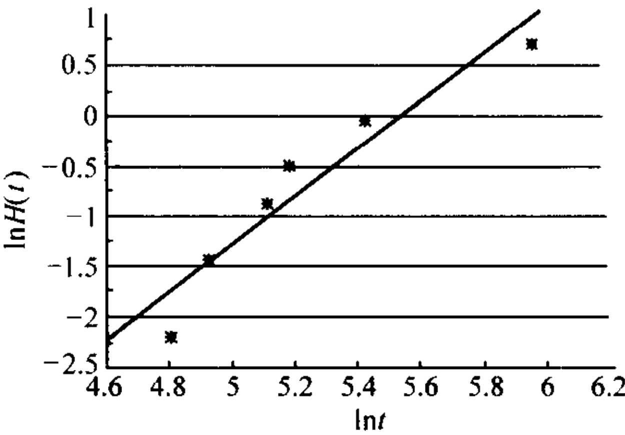


图6.6-24　威布尔分布风险坐标描点

3.9.3　概率坐标法

通过两次对威布尔累积分布函数取对数并变形得：

$$\ln\left[\ln\left(\frac{1}{1-F(x)}\right)\right] = \beta\ln x - \beta\ln\theta \quad (6.6\text{-}135)$$

通过对 $\ln\left[\ln\left(\frac{1}{1-F(x)}\right)\right]$ 对应的 $\ln x$ 进行描点，使直线服从于点的分布，就能对威布尔分布的参数进行估计。所得到直线的斜率可以对 β 进行估计，y 轴截距能够对 θ 进行估计，即：

$$\hat{\theta} = \exp\left(-\frac{y_0}{\hat{\beta}}\right) \quad (6.6\text{-}136)$$

累积分布函数 $F(x)$ 经常通过对中值序列的估计得来，但也使用其他的估计方法，如均值序列和 Kaplan - Meier

产品极限估计法。中间序列使用本章 3.6.3 中的式（6.6-82）~式（6.6-84）进行估计。专门的概率纸可用于概率坐标法。使用概率纸杜绝了描点之前对数据进行变形的需要，但这一切显然没有使用电子扩展表格方便和迅速。

例 16 对例 15 中的数据使用概率描点法确定威布尔分布的参数。

解 构建表 6.6-11 是为了获得描点所需的数据。

表 6.6-11 威布尔分布概率描点列表

失效时间	I_i	O_i	中位秩，$\hat{F}(t)$	$\ln\left[\ln\left(\frac{1}{1-F(x)}\right)\right]$	$\ln t$
104c					
122	1.100 0	1.100 0	0.076 9	-2.525 2	4.804 0
138	1.100 0	2.200 0	0.182 7	-1.600 8	4.927 3
167c					
168	1.257 1	3.457 1	0.303 6	-1.016 7	5.124 0
180	1.257 1	4.714 3	0.424 5	-0.593 4	5.193 0
217c					
229	1.571 4	0.575 5	0.575 5	-0.154 4	5.433 7
309c					
386	2.357 1	8.642 9	0.802 2	0.482 7	5.955 8

表中最后两列的数据能被用于描点。所描的图像如图 6.6-25 所示。通过数据求得（用线性回归的方法）最佳直线的斜率等于 2.41，对 β 进行了估计。通过数据得到的最佳直线 y 轴的截距是 -13.55。对威布尔分布比例参数的估计是：

$$\hat{\theta}=\exp\left(-\frac{-13.55}{2.41}\right)=276.6$$

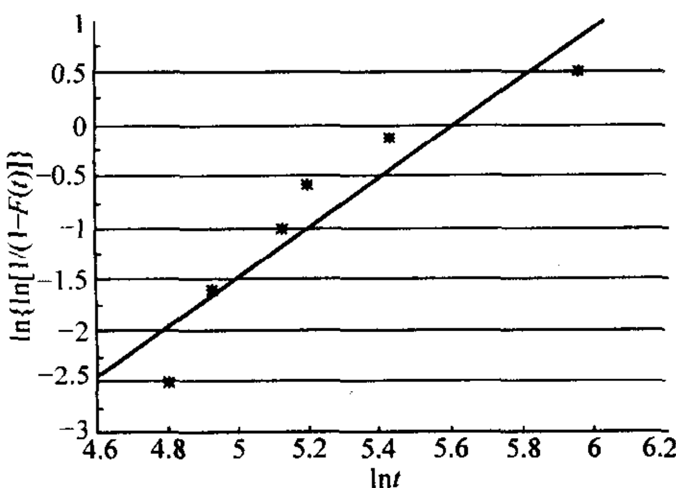


图 6.6-25 威布尔分布概率

置信极限能以 5% 和 95% 序列被加在此坐标描点中，详细介绍见本章 3.6.3。5% 和 95% 序列的描点位置见表达式：

$$x_i=\theta\left[\ln\left(\frac{1}{1-F(x)}\right)\right]^{1/\beta} \tag{6.6-137}$$

3.9.4 配合矩法

威布尔分布的均值被定义为：

$$\mu=\theta\Gamma\left(1+\frac{1}{\beta}\right) \tag{6.6-138}$$

威布尔分布的方差是：

$$\sigma^2=\vartheta^2\left[\Gamma\left(1+\frac{1}{\beta}\right)-\Gamma^2\left(1+\frac{1}{\beta}\right)\right] \tag{6.6-139}$$

威布尔分布的参数能够通过使以上的两个方程等于样本数据的均值和方差的方法获得。

例 17 用配合矩的方法确定威布尔分布的均值和方差。失效时间分别为：34、56、71、43 h 和 99 h。

解 五个失效时间样本的均值是 60.6，样本的方差是 655.3。用配合矩的方法获得了以下两个表达式：

$$60.6=\theta\Gamma\left(1+\frac{1}{\beta}\right)$$

$$655.3=\vartheta^2\left[\Gamma\left(1+\frac{1}{\beta}\right)-\Gamma^2\left(1+\frac{1}{\beta}\right)\right]$$

解了这两个表达式也同时给出了 β 和 θ 的估计值分别约为 2.54 和 68.3。

3.10 选择恰当的分布函数的意义

在使用一个分布之前，应先进行测试。有许多种合适的测试，有些是对应于具体分布的，有些是通用的。两个广泛使用的比较适合的测试是 χ^2 测试和 Kolmogorov-Smirnov 测试，都值针对可靠性分析的有限的值的。χ^2 测试要求样本容量较大（$n\geqslant25$），且不能与多个测试数据一起使用，Kolmogorov-Smirnov 测试通常没有专门的分布测试那样有效。

一个能用于多个测试数据的通用适合测试是 Hollander-Proschan 测试。该测试最好与一个例子结合进行说明，以表 6.6-12 中的数据为例，这些数据服从威布尔失效时间分布，对形状参数和比例参数的极大似然估计是 2.974 和 203.3。

表 6.6-12 适合的数据举例

失效时间			
42.1+	105.9	151.3+	195.6+
77.8	117.0	157.3	207.0
83.3+	126.9	163.8	215.3+
88.7	138.7	177.2+	217.4
101.8	148.9	194.3+	258.8+

以“+”结尾的数值表示该个体在失效之前就已经结束了测试。

为了检测形状参数和比例参数分别为 2.974 和 203.3 的威布尔分布对这些数据足够适合，应采取以下步骤。

1）每个未检测的数据使用 Kaplan-Meier 产品极限方法估计其可靠性。Kaplan-Meier 可靠性估计是：

$$\hat{R}(x_i)=\prod_{j=1}^{i}\left(\frac{n-j}{n-j+1}\right)^{Q_j} \tag{6.6-140}$$

式中，$Q_j=1$ 代表未检测的数据，$Q_j=0$ 代表检测的数据；n 是包括检测和未检测的观测值的样本容量。

2）对于每个观测值，无论是最后的、检测的还是未检测的，都要计算 Kaplan-Meier 可靠性估计的增量。除了最后的观测值以外，检测的观测值的增量都是 0。Kaplan-Meier 增量是：

$$\hat{f}(x_i)=\frac{1}{n}\prod_{j=1}^{i-1}\left(\frac{n-j+1}{n-j}\right)^{1-Q_j} \tag{6.6-141}$$

除了最后观测值以外，如果是测量的值，Kaplan-Meier 增量能够用以下表达式进行估计：

$$\hat{f}(x)=\hat{R}(x_p)-\hat{R}(x_i) \tag{6.6-142}$$

式中，$\hat{R}(x_p)$ 是对大多数以前未检测点的 Kaplan-Meier 可靠性估计。

3）用已经被测试合适的分布计算每个观测值的可靠性。对于这种情况，可靠性的表达式是：

$$R_0(x)=e^{-\left(\frac{x}{203.3}\right)^{2.974}}$$

4）对每个观测值都要计算 Kaplan-Meier 可靠性估计和 Kaplan-Meier 增量的结果。

5）对每个观测值都要计算使用对以前点和当前点的适合性而进行测试的分布的可靠性的不同。对第一个观测值来说，这等于1减去使用为适合性而进行测试的分布的可靠性。

6）对每个观测值，用$n/(n-i+1)$乘以在步骤5中获得的值。

1）～6）的计算步骤在表6.6-13中进行了摘要。

表6.6-13　1）～6）的计算摘要

a 次序	b x	c $\frac{n-i}{n-i+1}$	d $\hat{R}(x)$	e 跳动	f $R_0(x)$	g (e) * (f)	h $R_0^4(x)$	i $R_0^4(x_{i-1})$ $R_0^4(x_i)$	j $\frac{n-i}{n-i+1}$ * (i)
1	42.1　c			0.000 0	0.990 8	0.000 0	0.963 7	0.036 3	0.036 3
2	77.8	0.947 4	0.947 4	0.052 6	0.944 2	0.049 7	0.794 7	0.169 0	0.177 9
3	83.3　c			0.000 0	0.932 0	0.000 0	0.754 6	0.040 1	0.044 5
4	88.7　c			0.000 0	0.918 6	0.000 0	0.712 2	0.042 4	0.049 9
5	101.8	0.937 5	0.888 2	0.059 2	0.880 0	0.052 1	0.599 7	0.112 5	0.140 6
6	105.9	0.933 3	0.828 9	0.059 2	0.866 1	0.051 3	0.562 7	0.037 0	0.049 4
7	117.0	0.928 6	0.769 7	0.059 2	0.824 2	0.048 8	0.461 4	0.101 3	0.144 7
8	126.9	0.923 1	0.710 5	0.059 2	0.781 8	0.046 3	0.373 5	0.087 9	0.135 2
9	138.7	0.916 7	0.651 3	0.059 2	0.725 6	0.043 0	0.277 2	0.096 3	0.160 5
10	148.9	0.909 1	0.592 1	0.059 2	0.672 9	0.039 8	0.205 1	0.072 1	0.131 2
11	151.3　c			0.000 0	0.660 1	0.000 0	0.189 9	0.015 2	0.030 5
12	157.3	0.888 9	0.526 3	0.065 8	0.627 3	0.041 3	0.154 9	0.035 0	0.077 8
13	163.8	0.875 0	0.460 5	0.065 8	0.591 0	0.038 9	0.122 0	0.032 9	0.082 2
14	177.2　c			0.000 0	0.514 5	0.000 0	0.070 1	0.051 9	0.148 3
15	194.3　c			0.000 0	0.417 3	0.000 0	0.030 3	0.039 8	0.132 5
16	195.6　c			0.000 0	0.410 0	0.000 0	0.028 3	0.002 1	0.008 2
17	207.0	0.750 0	0.345 4	0.115 1	0.348 2	0.040 1	0.014 7	0.013 6	0.067 9
18	215.3　c			0.000 0	0.305 5	0.000 0	0.008 7	0.006 0	0.039 9
19	217.4	0.500 0	0.172 7	0.172 7	0.290 5	0.050 9	0.007 6	0.001 1	0.011 3
20	258.8　c			0.172 7	0.128 7	0.022 2 0.524 4	0.003	0.007 3	0.146 0 1.814 7

7）第g列的总和用C表示，第j列的总和用Y表示。例如，$C=0.524\,4$，$Y=1.814\,7$

8）用以下表达式估计C的标准偏差：

$$\hat{\sigma}=\sqrt{\frac{Y}{16}} \tag{6.6-143}$$

在这个例子中，$\hat{\sigma}=\sqrt{\frac{1.814\,7}{16}}=0.336\,8$，注意，$C$是近似的标准正态分布。

9）计算统计量

$$C^{*}=\frac{\sqrt{n}\,(C-0.5)}{\hat{\sigma}} \tag{6.6-144}$$

式中，n是包括测试和未测试的观测值的总数。在此例中：

$$C^{*}=\frac{\sqrt{20}\,(0.524\,4-0.5)}{0.336\,8}=0.324\,0$$

10）确定C^{*}是否落在测试所能接收的区域内。对于显著性$\mu=0.05$，所能接受的区域是$\pm z_{\alpha/2}=\pm 1.96$。

在此例中，计算统计量落在所能接受的区域内，形状参数为2.974，比例参数为203.3的威布尔分布被认为对数据来说是足够的模型。

4　随机变量分布类型的假设检验

4.1　分布类型的假设检验

分布类型的判断原则上有理论法和统计法两种。理论法是根据失效机理制定的数学模型或根据某种分布的性质推导出来的。例如，失效率为常数的寿命分布为指数分布；失效由"最弱"环节决定的寿命分布为极值分布；受很多独立随机因素和的影响，且没有一个因素起主导作用，这种分布为正态分布等。统计法是根据大量试验数据经统计求得的。很多同类性能在以往大量试验的基础上已经验证了其分布。例如，几何尺寸、材料性能、硬度等多服从正态分布；金属的疲劳寿命则服从对数正态分布或威布尔分布等。

下面仅介绍统计法。在使用统计法时，对分布不明的情况应做大样本的试验以判定其分布类型；对已有经验参考的情况则可做较小样本的试验，假设其分布类型再进行相应的拟合性检验。下面给出通用的χ^2检验法和$K-S$检验法。

4.1.1　χ^2检验法

χ^2检验法一般只用于大样本。χ^2检验法是计算理论频数与实际频数间的差异，将检验统计量χ^2的观测值与临界值$\chi^2_\alpha(\nu)$比较。满足下列条件，接受原假设；否则，拒绝原假设。

$$\chi^2=\sum_{i=1}^{k}\frac{(\nu_i-np_i)^2}{np_i}\leqslant \chi^2_\alpha(k-m-1) \tag{6.6-145}$$

式中，n为样本大小；k为分组数，按样本大小宜取$k=7\sim14$；ν_i为第i组的实际频数，$\nu_i\geqslant5$；p_i为第i组的理论频数（概率）；m为未知参数的数目；α为显著性水平；$\chi^2_\alpha(\cdot)$为临界值，查附表2。

例 18　220 个某产品的失效时间记录列于表 6.6-14 中。试检验该产品的寿命是否服从指数分布。

表 6.6-14　某产品失效时间的数据记录

时间 t/h	0 ~ 100	~ 200	~ 300	~ 400	~ 500	~ 600	~ 700	~ 800	~ 900
失效数 r_i	39	50	35	32	28	18	12	4	2

解　假设该产品的寿命服从指数分布，参数 λ 未知。取组中值作为该组时间的代表值 t_i，则 λ 的点估计

$$\hat{t} = \frac{1}{n}\sum_{i=1}^{k} t_i\nu_i = \frac{1}{220}(50\times 39 + 150\times 50 + \cdots + 850\times 2) = 293\ \text{h}$$

$$\hat{\lambda} = \frac{1}{\hat{t}} = \frac{1}{293}\quad 1/\text{h}$$

假设 H_0：$F(t) = 1 - e^{\frac{1}{293}}$。

为了使用 χ^2 检验法，首先按规定分组。由于每组中实际频数不宜少于 5，故将前 7 段时间各作为一组，最后两段时间合为一组。总计组数 $k = 8$，正好在 7 ~ 14 范围内。其他计算见表 6.6-15。

$$\chi^2 = \sum_{i=1}^{k}\frac{(\nu_i - np_i)^2}{np_i} = 36.905$$

取显著性水平 $\alpha = 0.10$，由 $\nu = k - m - 1 = 8 - 1 - 1 = 6$，查附表 2

$$\chi^2_\alpha(\nu) = \chi^2_{0.10}(6) = 10.64$$

由于 $\chi^2 > \chi^2_{0.10}$ (6)，故拒绝原假设，既不能认为该产品的寿命服从指数分布。

表 6.6-15　例 18 的列表计算

组号 i	$\nu_i = r_i$	$p_i = (1 - e^{-\frac{t_i}{293}}) - (1 - e^{-\frac{t_{i-1}}{293}})$	$np_i = 220p_i$	$\nu_i - np_i$	$(\nu_i - np_i)^2$	$\frac{(\nu_i - np_i)^2}{np_i}$
1	39	0.282 7	62.194	-23.194	537.962	8.650
2	50	0.205 5	45.210	4.790	22.944	0.507
3	35	0.146 1	32.140	2.860	8.180	0.254
4	32	0.103 9	22.858	9.142	83.576	3.656
5	28	0.073 8	16.236	11.764	138.392	8.524
6	18	0.052 5	11.550	6.450	41.603	3.602
7	12	0.037 3	8.206	3.794	14.394	1.754
8	6	0.091 7	20.174	-14.174	200.90	9.958
$\sum$						36.905

4.1.2　*K - S* 检验法

K - S 检验法（亦称 *d* 检验法）比 χ^2 检验法精确，而且还适用于小样本的情况。但是，*K - S* 检验法要求所检验的分布中不含未知参数。当指定分布中含有未知参数时，对某些分布应该用专门的临界值表。

K - S 检验法是将 n 个试验数据由小到大的次序排列。根据假设的分布，计算每个数据对应的 F_0 (x_i)，将其与经验分布函数 F_n (x_i) 相比较。其中，差的最大绝对值就是检验统计量 D_n 的观测值。将 D_n 与临界值 $D_{n,\alpha}$ 比较。满足下列条件，接受原假设；否则，拒绝原假设。

$$D_n = \sup_{-\infty < x < \infty} |F_n(x) - F_0(x)| = \max\{d_i\} \leqslant D_{n,\alpha}\quad (6.6\text{-}146)$$

式中，F_0 (x) 为原假设的分布函数；F_n (x) 为经验分布函数；$D_{n,\alpha}$ 为临界值，查表 6.6-16。

$$F_n(x) = \begin{cases} 0, & x < x_1 \\ \frac{i}{n}, & x_i < x \leqslant x_{i-1} \\ 1, & x > x_n \end{cases}\quad (6.6\text{-}147)$$

$$d_i = \max\left[F_0(x_i) - \frac{i-1}{n}, \frac{i}{n} - F_0(x_i)\right]\quad (6.6\text{-}148)$$

表 6.6-16　*K - S* 检验临界值表

n \ α	0.20	0.10	0.05	0.02	0.01
1	0.900 00	0.950 00	0.975 00	0.990 00	0.995 00
2	0.683 77	0.776 39	0.841 89	0.900 00	0.929 29
3	0.564 81	0.636 04	0.707 60	0.784 56	0.829 00
4	0.492 65	0.565 22	0.623 94	0.688 87	0.734 24
5	0.446 98	0.509 45	0.563 28	0.627 18	0.668 53

续表 6.6-16

n \ α	0.20	0.10	0.05	0.02	0.01
6	0.410 37	0.467 99	0.519 26	0.577 41	0.616 61
7	0.381 48	0.436 07	0.483 42	0.538 44	0.575 81
8	0.358 31	0.409 62	0.454 27	0.506 54	0.541 79
9	0.339 10	0.387 46	0.430 01	0.479 60	0.513 32
10	0.322 60	0.368 66	0.409 25	0.456 62	0.488 93
11	0.308 29	0.352 42	0.391 22	0.436 70	0.467 70
12	0.295 77	0.338 15	0.375 43	0.419 18	0.449 05
13	0.284 70	0.325 49	0.361 43	0.403 62	0.432 47
14	0.274 81	0.314 17	0.348 90	0.389 70	0.417 62
15	0.265 88	0.303 97	0.337 60	0.377 13	0.404 20
16	0.257 78	0.294 72	0.327 33	0.365 71	0.392 01
17	0.250 39	0.286 27	0.317 96	0.355 28	0.380 86
18	0.243 60	0.278 51	0.309 36	0.345 69	0.370 62
19	0.237 35	0.271 36	0.301 43	0.336 85	0.361 17
20	0.231 56	0.264 73	0.294 08	0.328 66	0.352 41
21	0.226 17	0.258 58	0.287 24	0.321 04	0.344 27
22	0.221 15	0.252 83	0.280 87	0.313 94	0.336 66
23	0.216 45	0.247 46	0.274 90	0.307 28	0.329 54
24	0.212 05	0.242 42	0.269 31	0.301 04	0.322 86
25	0.207 90	0.237 68	0.264 04	0.295 16	0.316 57
26	0.203 99	0.233 20	0.259 07	0.289 62	0.310 64
27	0.200 30	0.228 98	0.254 38	0.284 38	0.305 02
28	0.196 80	0.224 97	0.249 93	0.279 42	0.299 71
29	0.193 48	0.221 17	0.245 71	0.274 71	0.294 66
30	0.190 32	0.217 56	0.241 70	0.270 23	0.289 87

例 19 某合金 9 个试件测得的强度极限为 453，436，429，419，405，416，432，423，440 MPa。检验该合金的强度极限是否服从均值 $\mu=28$ MPa，标准差 $\sigma=15$ MPa 的正态分布。

解 令该合金的强度极限 $\sigma_b=X$，将数据按由小到大次序排列。假设 X 服从正态分布，分布函数

$$F(x)=\int_{-\infty}^{x}\frac{1}{15\sqrt{2\pi}}e^{-\frac{(x-428)^2}{2\times15^2}}dx=\Phi(\frac{x-428}{15})$$

式中的 $\Phi(\cdot)$ 查附表 1。计算结果见表 6.6-17。

由上述中计算结果知，D_n 的观测值按式（6.6-146）

$$D_n=\max\{d_i\}=0.1009$$

取显著性水平 $\alpha=0.10$，由表 6.6-16 查得 $D_{n,\alpha}=0.38764$。由于 $D_n<D_{n,\alpha}$，故接受原假设，即认为该合金的强度极限服从 $\mu=28$ MPa，$\sigma=15$ MPa 的正态分布。

表 6.6-17 例 6.6-19 的列表计算

序号 i	x_i	$F(x)=\Phi(\frac{x-428}{15})$	$\frac{i-1}{n}$	$\frac{i}{n}$	d_i
1	405	0.063 0	0.000	0.111	0.030 1
2	416	0.211 9	0.111	0.222	0.100 9
3	419	0.274 3	0.222	0.333	0.058 7
4	423	0.370 7	0.333	0.444	0.073 3
5	429	0.527 9	0.444	0.556	0.083 9
6	432	0.606 4	0.556	0.667	0.060 6
7	436	0.709 1	0.667	0.778	0.076 1
8	440	0.788 1	0.778	0.889	0.100 9
9	453	0.952 5	0.889	1.000	0.063 5

4.1.3 回归分析法

回归分析法就是图解分析法的解析。在直角坐标纸上描得几个试验点 (x_1, y_1)，(x_2, y_2)，…，(x_n, y_n)，如图 6.6-26 所示。按最小二乘原理确定直线

$$\hat{y}=\hat{A}+\hat{B}x \qquad (6.6\text{-}149)$$

它反映出试验点散布状态的一条最佳直线，称为回归直线。斜率 $\hat{B}$称为回归系数，截距 $\hat{A}$为常数项。

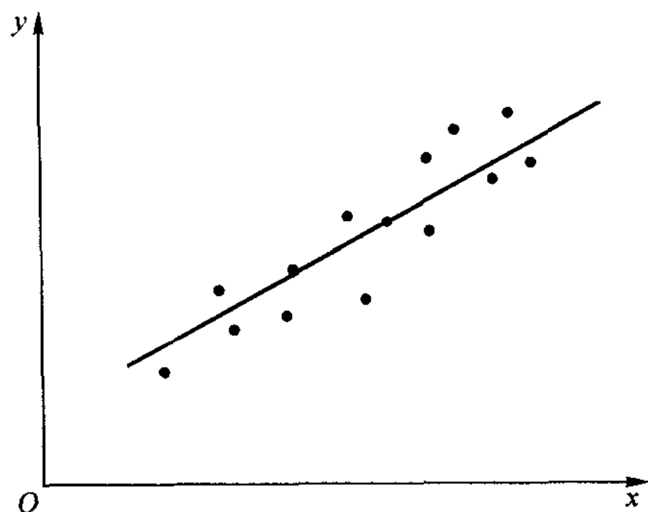


图 6.6-26 x，y 散点图和回归直线

$$\hat{B}=\frac{\sum_{i=1}^{n}x_iy_i-n\overline{xy}}{\sum_{i=1}^{n}x_i^2-n\bar{x}} \qquad (6.6\text{-}150)$$

$$\hat{A}=\bar{y}-\hat{B}\bar{x} \qquad (6.6\text{-}151)$$

式中

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \qquad (6.6\text{-}152)$$

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i \qquad (6.6\text{-}153)$$

各试验点是否在一直线上，即是否具有线性相关的关系，可用相关系数检验法进行检验。相关系数计算公式为：

$$\hat{\rho}=\frac{\sum_{i=1}^{n}x_iy_i-n\overline{xy}}{\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}y_i^2-n\bar{y}\right)^{\frac{1}{2}}} \qquad (6.6\text{-}154)$$

当 $|\hat{\rho}|>\rho_\alpha$，则认为具有线性相关的关系。ρ_α 是显著性水平为 α 时的相关系数起码值，见表 6.6-18。查表时取自由度 $\nu=n-2$。

表 6.6-18 相关系数起码值

ν \ α	0.10	0.05	0.02	0.01	0.001
5	0.669 4	0.754 5	0.832 9	0.874 5	0.950 7
6	0.621 5	0.706 7	0.788 7	0.834 3	0.924 9
7	0.582 2	0.666 4	0.749 8	0.797 7	0.898 2
8	0.549 4	0.631 9	0.715 5	0.764 6	0.872 1
9	0.521 4	0.602 1	0.685 1	0.734 8	0.847 1
10	0.497 3	0.576 0	0.658 1	0.707 9	0.823 3
11	0.476 2	0.552 9	0.633 9	0.683 5	0.801 0
12	0.457 5	0.532 4	0.612 0	0.661 4	0.780 0
13	0.440 9	0.513 9	0.592 3	0.641 1	0.760 3
14	0.425 9	0.497 3	0.574 2	0.622 6	0.742 0
15	0.412 4	0.482 1	0.557 7	0.605 5	0.724 6
16	0.400 0	0.468 3	0.542 5	0.589 7	0.708 4
17	0.388 7	0.455 5	0.528 5	0.575 1	0.693 2
18	0.378 3	0.443 8	0.515 5	0.561 4	0.678 7
19	0.368 7	0.432 9	0.503 4	0.548 7	0.665 2
20	0.359 8	0.422 7	0.492 1	0.536 8	0.652 4

对于可靠性分析中常用的概率分布，其分布函数与自变量之间一般在直角坐标上并不成线性关系，因此应先进行适当的变换。几种常用概率分布的变换关系列于表 6.6-19 中。

在使用回归分析法时，首先将试验获得的 n 个数据按由小到大的次序排列，即 $t_1<t_2<\cdots<t_n$，取中位秩作为各试验点相应的分布函数，即

$$\hat{F}(t_i)\approx\frac{i-0.3}{n+0.4} \qquad (6.6\text{-}155)$$

表 6.6-19 几种常用概率分布的变换关系

名称	分布函数 $\hat{F}(t_i)\approx\frac{i-0.3}{n+0.4}$	y	x	B	A
指数分布	$1-e^{-\lambda t}$	$\ln\frac{1}{1-F(t)}$	t	λ	0
威布尔分布	$1-e^{-\left(\frac{t-a}{b}\right)^k}$	$\ln\ln\frac{1}{1-F(t)}$	$\ln(t-a)$	k	$-k\ln b$
正态分布	$\int_{-\infty}^{t}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$	$\frac{t-\mu}{\sigma}=\Phi^{-1}[F(t)]$	t	$\frac{1}{\sigma}$	$-\frac{\mu}{\sigma}$
对数正态分布	$\int_{0}^{t}\frac{1}{t\sigma\sqrt{2\pi}}e^{-\frac{(\ln t-\mu)^2}{2\sigma^2}}dt$	$\frac{\ln t-\mu}{\sigma}=\Phi^{-1}[F(t)]$	$\ln t$	$\frac{1}{\sigma}$	$-\frac{\mu}{\sigma}$

假设一种分布，按表 6.6-19 进行变换后即可用式（6.6-149）～式（6.6-155）进行计算。若相关系数检验通过则接受原假设。估计得 B、A 后再按表 6.6-19 中关系估计原分布函数的参数。

例 20 某合金材料在某应力水平做疲劳寿命试验，10

个试件的疲劳寿命分别为 211，229，272，276，295，303，332，354，382，409 千次。试进行分布类型的判断并进行参数估计。

解 一般金属疲劳寿命多较好地服从对数正态分布，故先假设该材料疲劳寿命服从对数正态分布。按表 6.6-19

$$y=\frac{\ln N-\mu}{\sigma},\ x=\ln N,\ B=\frac{1}{\sigma},\ A=-\frac{\mu}{\sigma}$$

计算结果见表 6.6-20。

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{57.046}{10}=5.704\,6$$

$$\overline{y}=\frac{1}{n}\sum_{i=1}^{n}y_i=0$$

相关系数检验，按式（6.6-154）

表 6.6-20 例 20 的计算过程

序号	N_i	$x_i=\ln N_i$	$\hat{F}(N_i)\approx\frac{i-0.3}{n+0.4}$	$y_i=\Phi^{-1}[\hat{F}(N_i)]$	x_i^2	y_i^2	x_iy_i
1	211	5.352	0.067	-1.500	28.642	2.250	-8.028
2	229	5.434	0.162	-0.985	29.525	0.970	-5.352
3	272	5.606	0.295	-0.645	31.425	0.416	-3.616
4	276	5.620	0.356	-0.370	31.589	0.137	-2.080
5	295	5.687	0.452	-0.120	32.342	0.014	-0.682
6	303	5.714	0.548	0.120	32.647	0.014	0.686
7	332	5.805	0.644	0.370	33.700	0.137	2.148
8	354	5.869	0.741	0.645	34.445	0.416	3.786
9	382	5.945	0.838	0.985	35.348	0.970	5.856
10	409	6.014	0.933	1.500	36.165	2.250	9.021
$\sum$		57.046		0	325.828	7.574	1.739

$$\hat{\rho}=\frac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\left(\sum_{i=1}^{n}x_i^2-n\overline{x}^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}y_i^2-n\overline{y}^2\right)^{\frac{1}{2}}}=\frac{1.739-0}{(325.828-10\times5.704\,6)^{\frac{1}{2}}(7.574-0)^{\frac{1}{2}}}=0.994\,9$$

由表 6.6-18，当 $\nu=n-2=8$，$\alpha=0.05$ 查得 $\rho_{0.05}=0.631\,9$。$\hat{\rho}>\rho_\alpha$，故接受疲劳寿命服从对数正态分布的假设。

按式（6.6-150）

$$\hat{B}=\frac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum_{i=1}^{n}x_i^2-n\overline{x}}=\frac{1.739-0}{325.828-10\times5.704\,6^2}=4.311$$

按式（6.6-151）

$$\hat{A}=\overline{y}-\hat{B}\overline{x}=0-4.311\times5.704\,6=-24.592\,5$$

估计分布参数 μ 和 σ，由表 6.6-19 知：$B=\frac{1}{\sigma}$，$A=-\frac{\mu}{\sigma}$，故

$$\hat{\sigma}=\frac{1}{\hat{B}}=\frac{1}{4.311}=0.232$$

$$\hat{\mu}=-\hat{A}\hat{\sigma}=-(-24.592\,5\times0.232)=5.704\,6$$

4.2 正态及对数正态分布的分析法

正态分布的概率密度

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\tag{6.6-156}$$

对数正态分布的概率密度

$$f(y)=\frac{1}{\sigma y\sqrt{2\pi}}e^{\frac{(\ln y-\mu)^2}{2\sigma^2}}\tag{6.6-157}$$

由对数正态分布与正态分布的关系知，若随机变量 $Y\sim ln(\mu,\sigma^2)$，则 $X=\ln Y\sim N(\mu,\sigma^2)$。故取

$$x=\ln y\tag{6.6-158}$$

则正态分布的所有分析方法都可用于对数正态分布。但要注意，进行对数正态分布的分析时，必须将其数据按式（6.6-158）取成对数。

4.2.1 拟合性检验

对于样本容量不大、分布参数 μ、σ 未知时，若用两参数的点估计 $\overline{x}$，s_x 代替时，则假设

$$F_0(x;\overline{x},s_x^2)=\int_{-\infty}^{x}\frac{1}{s_x\sqrt{2\pi}}e^{-\frac{(x-\overline{x})^2}{2s_x^2}}dx\tag{6.6-159}$$

与 $K-S$ 检验法类似，满足下列条件则接受原假设，否则拒绝原假设。

$$\tilde{D}_n=\sup_{-\infty<x<\infty}|F_0(x;\overline{x},s_x^2)-F_n(x)|=\max\{\tilde{d}_i\}\leqslant\tilde{D}_{n,\alpha}\tag{6.6-160}$$

式中，$F_n(x)$ 为经验分布函数，同式（6.6-148）；$\tilde{D}_{n,\alpha}$ 为临界值，查表 6.6-21。

$$\tilde{d}_i=\max\left[F_0(x;\overline{x},s_x^2)-\frac{i-1}{n},\frac{i}{n}-F_0(x;\overline{x},s_x^2)\right]\tag{6.6-161}$$

表 6.6-21 $\tilde{D}_n$ 的临界值 $\tilde{D}_{n,\alpha}$（正态分布）

n \ α	0.20	0.15	0.10	0.05	0.01
4	0.300	0.319	0.352	0.381	0.417
5	0.285	0.299	0.315	0.337	0.405
6	0.265	0.277	0.294	0.319	0.364
7	0.247	0.258	0.276	0.300	0.348
8	0.233	0.244	0.261	0.285	0.331
9	0.223	0.233	0.249	0.271	0.311
10	0.215	0.224	0.239	0.258	0.294
11	0.206	0.217	0.230	0.249	0.284
12	0.199	0.212	0.223	0.242	0.275
13	0.190	0.202	0.214	0.234	0.268
14	0.183	0.194	0.207	0.227	0.261

续表 6.6-21

n \ α	0.20	0.15	0.10	0.05	0.01
15	0.177	0.187	0.201	0.220	0.257
16	0.173	0.182	0.195	0.213	0.250
17	0.169	0.177	0.189	0.206	0.245
18	0.166	0.173	0.184	0.200	0.239
19	0.163	0.169	0.179	0.195	0.235
20	0.160	0.166	0.174	0.190	0.231
25	0.142	0.147	0.158	0.173	0.200
30	0.131	0.136	0.144	0.161	0.187

例 21　对某钢材进行静强度试验，9 个试件的强度极限按由小到大次序分别为 625，650，656，659，661，662，663，668，672 MPa。检验该钢材强度极限是否服从正态分布。

解　假设该钢材的强度极限服从正态分布。由于分布参数未知，故先进行估计。由式（6.6-162）、式（6.6-163）

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{9}(625+650+656+659+661+662+663+668+672) = 657.3\ \text{N/mm}^2$$

$$s_x = \left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{\frac{1}{2}} = \left\{\frac{1}{9-1}\left[(625-657.3)^2+(650-657.3)^2+\cdots+(672-657.3)^2\right]\right\}^{\frac{1}{2}} = 13.69\ \text{N/mm}^2$$

假设：

$$F_0(x;\bar{x},s_x^2) = \int_{-\infty}^{x}\frac{1}{13.69\sqrt{2\pi}}\mathrm{e}^{\frac{(x-657.3)^2}{2\times13.69^2}}\mathrm{d}x$$

列表计算 $\tilde{d}_i$，见表 6.6-22。Φ（·）查附表 1。

表 6.6-22　例 21 的列表计算

序号 i	x_i/(N/mm²)	$F_0(x_i)=\Phi(\frac{x_i-657.3}{13.69})$	$\frac{i-1}{9}$	$\frac{i}{9}$	$\tilde{d}_i$
1	625	0.009 137	0.000 00	0.111 11	0.101 97
2	650	0.298 1	0.111 11	0.222 22	0.196 99
3	656	0.464 1	0.222 22	0.333 33	0.241 88
4	659	0.547 8	0.333 33	0.444 44	0.214 47
5	661	0.606 4	0.444 44	0.555 56	0.161 96
6	662	0.633 1	0.555 56	0.666 67	0.077 54
7	663	0.662 8	0.666 67	0.777 78	0.114 98
8	668	0.782 3	0.777 78	0.888 89	0.106 59
9	672	0.857 7	0.888 89	1.000 00	0.142 30

由上表中的计算结果知，$\tilde{D}_n$ 的观察值按式（6.6-160）

$$\tilde{D}_n = \max\{d_i\} = 0.241\ 88$$

取显著性水平 $\alpha=0.10$，由表 6.6-21 查得 $\tilde{D}_{n,\alpha}=0.249$。因为 $\tilde{D}_n < \tilde{D}_{n,\alpha}$，故接受原假设，即认为该钢材的强度极限服从正态分布。

例 22　某金属材料在某应力水平用 10 个试件做弯曲疲劳试验，其寿命循环次数 N 分别为 125 000，132 000，135 000，138 000，141 000，147 000，154 000，161 000，164 000，182 000。检验该寿命分布是否服从对数正态分布。

解　假设该金属材料的疲劳寿命服从对数正态分布。按式（6.6-158）求对数寿命 $x_i=\ln N_i$ = 11.736，11.791，11.813，11.835，11.857，11.898，11.945，11.989，12.008，12.112。按式（6.6-162）、式（6.6-163）估计分布参数：

$$\bar{x} = \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i = \frac{1}{10}(11.736+11.791+\cdots+12.112) = 11.898$$

$$s_x = \hat{\sigma} = \left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{\frac{1}{2}} = \left\{\frac{1}{10-1}\left[(11.736-11.898)^2+(11.791-11.898)^2+\cdots+(12.112-11.898)^2\right]\right\}^{\frac{1}{2}} = 0.115$$

列表计算 $\tilde{d}_i$，见表 6.6-23。Φ(·)查附表 1。

表 6.6-23　例 22 的列表计算

序号 i	N_i	x_i	$F_0(N_i)=\Phi(\frac{x_i-11.898}{0.115})$	$\frac{i-1}{10}$	$\frac{i}{10}$	$\tilde{d}_i$
1	125 000	11.736	0.009 137	0.0	0.1	0.079 47
2	132 000	11.791	0.298 1	0.1	0.2	0.076 06
3	135 000	11.813	0.464 1	0.2	0.3	0.070 10
4	138 000	11.836	0.547 8	0.3	0.4	0.108 09
5	141 000	11.857	0.606 4	0.4	0.5	0.139 27
6	147 000	11.898	0.633 1	0.5	0.6	0.100 00
7	154 000	11.945	0.662 8	0.6	0.7	0.058 61
8	161 000	11.989	0.782 3	0.7	0.8	0.085 62
9	164 000	12.008	0.857 7	0.8	0.9	0.069 40
10	182 000	12.112		0.9	1.0	0.068 61

由上表中的计算结果知，$\tilde{D}_n$ 的观察值按式（6.6-160）

$$\tilde{D}_n = \max\{d_i\} = 0.139\ 27$$

取显著性水平 $\alpha=0.10$，由表 6.6-21 查得 $\tilde{D}_{n,\alpha}=0.239$。因为 $\tilde{D}_n<\tilde{D}_{n,\alpha}$，故接受原假设，即认为该金属材料的疲劳寿命服从对数正态分布。

4.2.2　完全样本的参数估计

正态分布均值的点估计

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i \tag{6.6-162}$$

标准差的点估计

$$\hat{\sigma} = s_x = \left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{\frac{1}{2}} \tag{6.6-163}$$

均值 μ 的双侧置信的区间估计。置信下限 μ_L 和置信上限 μ_U 分别为

$$\mu_L = \bar{x}_L = \bar{x} - \frac{s_x}{\sqrt{n}}t_{\frac{\alpha}{2}}(\nu) \tag{6.6-164}$$

$$\mu_U = \bar{x}_U = \bar{x} + \frac{s_x}{\sqrt{n}}t_{\frac{\alpha}{2}}(\nu) \tag{6.6-165}$$

均值 μ 的单侧置信下限

$$\mu_L = \bar{x}_L = \bar{x} - \frac{s_x}{\sqrt{n}}t_{\alpha}(\nu) \tag{6.6-166}$$

式中，α 为显著性水平，$1-\alpha$ 为置信水平；n 为样本大小；ν 为自由度，当标准差 σ 已知，s_x 用 σ 代替，则 $\nu=\infty$；当标准差为点估计 s_x，则 $\nu=n-1$；t（ν）为 t 分布的分位

数，查附表3。

标准差的双侧置信区间的估计。置信下限 σ_L 和置信上限 σ_U 分别为

$$\sigma_L = s_{xL} = \left[\frac{n-1}{\chi^2_{\frac{\alpha}{2}}(\nu)}\right]^{\frac{1}{2}} s_x \quad (6.6\text{-}167)$$

$$\sigma_U = s_{xU} = \left[\frac{n-1}{\chi^2_{1-\frac{\alpha}{2}}(\nu)}\right]^{\frac{1}{2}} s_x \quad (6.6\text{-}168)$$

标准差 σ 的单侧置信上限

$$\sigma_U = s_{xU} = \left[\frac{n-1}{\chi^2_{1-\alpha}(\nu)}\right]^{\frac{1}{2}} s_x \quad (6.6\text{-}169)$$

式中，n 为样本大小；ν 为自由度，当均值 μ 已知时，$\nu = n$；当 μ 未知时，$\nu = n-1$。α 为显著性水平；$1-\alpha$ 为置信水平；$\chi^2(\nu)$ 为 χ^2 分布的分位数，查附表2。

若随机变量 y 服从对数正态分布，即 $Y \sim ln(\mu, \sigma^2)$，则 $\ln Y \sim N(\mu, \sigma^2)$，故取

$$x = \ln y$$

则式（6.6-162）～式（6.6-169）均可用于对数正态分布。

4.2.3 截尾寿命试验的参数估计

1）极大似然估计　当样本较大时，正态分布均值和标准差的极大似然估计是具有良好性质的估计量。从寿命服从正态分布的总体中抽取 n 个试样进行定时截尾寿命试验，试验到 x_0 时结束，共失效 r 个，失效时间为

$$x_1 \leqslant x_2 \leqslant \cdots \leqslant x_r \leqslant x_0$$

则寿命的均值 μ 和标准差 σ 的极大似然估计分别为

$$\hat{\mu} = \bar{x} = \bar{x}_r + \frac{s_x^2 - s_r^2}{d} \quad (6.6\text{-}170)$$

$$\hat{\sigma} = s_x = \frac{d}{g\left(D, \frac{r}{n}\right)} \quad (6.6\text{-}171)$$

式中，$\bar{x}_r = \frac{1}{r}\sum_{i=1}^{r} x_i$；$s_r^2 = \frac{1}{r}\sum_{i=1}^{r}(x_i - \bar{x}_r)^2$；$d = x_0 - \bar{x}_r$；$D = \frac{d^2}{d^2 + s_r^2}$；$g\left(D, \frac{r}{n}\right)$ 查表6.6-24。

寿命服从对数正态分布时，将寿命 y 取成对数，即 $x = \ln y$，则可用式（6.6-170）、式（6.6-171）估计其参数。

2）最佳线性无偏估计　寿命服从正态分布时，均值 μ 和标准差 σ 的最佳线性无偏估计（简记为BLUE）分别为

$$\hat{\mu} = \bar{x} = \sum_{j=1}^{r} D'(n, r, j)\, x_j \quad (6.6\text{-}172)$$

$$\hat{\sigma} = s_x = \sum_{j=1}^{r} C'(n, r, j)\, x_j \quad (6.6\text{-}173)$$

式中，n 为样本大小；r 为失效数；j 为寿命由小到大排列的次序；x_j 为第 j 个寿命值；$D'(n, r, j)$ 为 μ 的最佳线性无偏估计系数，查表6.6-25；$C'(n, r, j)$ 为 σ 的最佳线性无偏估计系数，查表6.6-25。

寿命服从对数正态分布时，将寿命 y 取对数，即 $x = \ln y$，则可用式（6.6-172）、式（6.6-173）估计其参数。

3）简单线性无偏估计　寿命服从正态分布时，均值 μ 和标准差 σ 的简单线性无偏估计（简记为GLUE）分别为

$$\hat{\mu} = \bar{x} = x_r - E(Y_{r,n})\,\hat{\sigma} \quad (6.6\text{-}174)$$

$$\hat{\sigma} = s_x = \frac{1}{nk_{r,n}}\left[r x_r - \sum_{j=1}^{r} x_j\right] \quad (6.6\text{-}175)$$

式中，n 为样本大小；r 为失效数；j 为寿命由小到大排列的次序；x_j 为第 j 个寿命值；$nk_{r,n}$ 为系数，查表6.6-26；$E(Y_{r,n})$ 为系数，查表6.6-26。

寿命服从对数正态分布时，将寿命 y 取对数，即 $x = \ln y$，则可用式（6.6-174）、式（6.6-175）估计其参数。

4.2.4 可靠寿命和可靠度的估计

寿命服从正态分布时，指定可靠度 R，可靠寿命的点估计

$$\hat{x}(R) = \bar{x} - z_R s_x \quad (6.6\text{-}176)$$

式中，z_R 按指定的可靠度 R 查表6.6-27或附表1。

指定寿命 x 的可靠度 R 的点估计

$$\hat{R}(x) = \Phi(z_R) \quad (6.6\text{-}177)$$

式中，$z_R = \frac{\bar{x} - x}{s_x}$；$\Phi(\cdot)$ 为标准正态分布函数，按 z_R 查表6.6-27或附表1。

指定可靠度 R_L 的可靠寿命置信下限

$$x_L(R) = \bar{x} - z_{R\gamma} s_x \quad (6.6\text{-}178)$$

式中，$z_{R\gamma}$ 为单侧置信限系数，按指定的可靠度和置信水平由下式求取。

$$z_{R\gamma} = \frac{z_R + z_\gamma\left[\frac{1}{n}\left(1 - \frac{z_\gamma^2}{2n-2}\right) + z_R^2\right]}{1 - \frac{z_\gamma^2}{2n-2}} \quad (6.6\text{-}179)$$

式中，z_R 按指定的可靠度 R 查表6.6-27或附表1；z_γ 按指定的置信水平查表6.6-28；n 为样本大小，本式用于 $n \geqslant 5$。

指定寿命 x_L 的可靠度置信下限

$$R_L = \Phi(z_{RL}) \quad (6.6\text{-}180)$$

$$z_{RL} \approx z_{R\gamma} - z_\gamma\left(\frac{1}{n} + \frac{z_{R\gamma}^2}{2n-2}\right)^{\frac{1}{2}} \quad (6.6\text{-}181)$$

$$z_{R\gamma} = \frac{\bar{x} - x_L}{s_x} \quad (6.6\text{-}182)$$

式中，z_γ 按指定的置信水平查表6.6-28；n 为样本大小，本式用于 $n \geqslant 5$。

应该指出，这里的 x 并不限于寿命，也可以是服从正态分布的其他特性值，例如材料的机械强度等。

若随机变量 y 服从对数正态分布，即 $y \sim ln(\mu, \sigma^2)$，则 $\ln y \sim N(\mu, \sigma^2)$，故正态分布可靠寿命及可靠度的估计完全适用于对数正态分布，只是将对数正态分布的数据取成对数后计算即可。

例23　某钢材的强度极限服从正态分布，11个试件测得的强度极限为608，622，630，638，642，648，652，660，673，688 MPa。求均值和标准差的点估计和置信水平 γ = 80% 的双侧置信限，失效概率为0.10时强度极限的点估计和置信水平为90%的单侧置信下限。

解　本例为完全样本试验。

均值 μ 的点估计，按式（6.6-162）

表6.6-24　极大似然估计用表（正态及对数正态分布）$g\left(D, \frac{r}{n}\right)$

D \ r/n	0.01	0.10	0.30	0.50	0.70	0.90	0.99
0.02	0.119 803	0.128 283	0.134 006	0.137 323	0.139 833	0.141 919	0.142 766
0.04	0.158 428	0.174 515	0.185 809	0.192 542	0.197 737	0.202 128	0.203 929
0.08	0.184 441	0.207 584	0.224 359	0.234 588	0.242 612	0.249 489	0.252 337

续表 6.6-24

D \ r/n	0.01	0.10	0.30	0.50	0.70	0.90	0.99
0.08	0.204 192	0.233 957	0.256 138	0.269 941	0.280 928	0.290 466	0.294 451
0.10	0.220 083	0.256 112	0.283 642	0.301 093	0.315 177	0.327 552	0.332 765
0.12	0.233 342	0.275 319	0.308 144	0.329 317	0.346 631	0.362 023	0.368 559
0.14	0.244 675	0.292 320	0.330 395	0.355 365	0.376 043	0.394 640	0.402 600
0.16	0.254 529	0.307 598	0.350 884	0.379 724	0.403 906	0.425 900	0.435 392
0.18	0.263 221	0.321 489	0.369 945	0.402 732	0.430 560	0.456 157	0.467 293
0.20	0.270 961	0.334 228	0.387 825	0.424 625	0.456 254	0.485 670	0.498 573
0.22	0.277 929	0.345 996	0.404 705	0.445 617	0.481 180	0.514 643	0.529 443
0.24	0.284 238	0.356 937	0.420 728	0.465 822	0.505 485	0.543 238	0.560 078
0.26	0.289 986	0.367 159	0.436 006	0.485 363	0.529 291	0.571 593	0.590 625
0.28	0.295 262	0.376 751	0.450 631	0.504 332	0.552 696	0.599 821	0.621 214
0.30	0.300 118	0.385 787	0.464 675	0.522 807	0.575 785	0.628 029	0.651 965
0.32	0.304 609	0.394 321	0.478 201	0.540 851	0.598 630	0.656 310	0.682 987
0.34	0.308 773	0.402 410	0.491 262	0.558 518	0.621 293	0.684 749	0.714 388
0.36	0.312 652	0.410 097	0.503 902	0.575 856	0.643 832	0.713 431	0.746 271
0.38	0.316 279	0.417 416	0.516 159	0.592 905	0.666 297	0.742 436	0.778 742
0.40	0.319 675	0.424 401	0.528 066	0.609 701	0.688 735	0.771 842	0.811 910
0.42	0.322 866	0.431 079	0.539 653	0.626 274	0.711 190	0.801 731	0.845 886
0.44	0.325 872	0.437 474	0.550 944	0.642 655	0.733 702	0.832 183	0.880 790
0.46	0.328 701	0.443 609	0.561 962	0.658 867	0.756 312	0.863 284	0.916 751
0.48	0.331 379	0.449 502	0.572 728	0.674 935	0.779 057	0.895 124	0.953 909
0.50	0.333 907	0.455 171	0.583 258	0.690 879	0.801 975	0.927 798	0.992 418
0.52	0.336 312	0.460 630	0.593 572	0.706 719	0.825 104	0.961 408	1.032 447
0.54	0.338 592	0.465 894	0.603 680	0.722 473	0.848 480	0.996 065	1.074 191
0.56	0.340 763	0.470 975	0.613 598	0.738 158	0.872 141	1.031 894	1.117 867
0.58	0.342 832	0.475 883	0.623 336	0.753 790	0.896 127	1.069 027	1.163 726
0.60	0.344 803	0.480 631	0.632 907	0.769 384	0.920 476	1.107 617	1.212 056
0.62	0.346 672	0.485 229	0.642 320	0.784 954	0.945 230	1.147 833	1.263 194
0.64	0.348 472	0.489 686	0.651 584	0.800 514	0.970 432	1.189 866	1.317 533
0.66	0.350 196	0.493 994	0.660 709	0.816 077	0.996 129	1.233 935	1.375 548
0.68	0.351 851	0.498 185	0.669 700	0.831 657	1.022 367	1.280 293	1.437 805
0.70	0.353 437	0.502 251	0.678 567	0.847 265	1.049 200	1.329 233	1.504 991
0.72	0.354 954	0.506 200	0.687 316	0.862 915	1.076 682	1.381 097	1.577 953
0.74	0.356 417	0.510 039	0.695 952	0.878 617	1.104 872	1.436 281	1.657 757
0.76	0.357 823	0.513 776	0.704 482	0.894 385	1.133 836	1.495 273	1.745 749
0.78	0.359 175	0.517 410	0.712 912	0.910 229	1.163 643	1.558 645	1.843 674
0.80	0.360 478	0.520 952	0.721 246	0.926 163	1.194 371	1.627 102	1.953 839
0.82	0.361 740	0.524 400	0.729 489	0.942 197	1.226 107	1.701 509	2.079 374
0.84	0.362 946	0.527 764	0.373 646	0.958 345	1.258 944	1.782 952	2.224 653
0.86	0.364 118	0.531 044	0.745 721	0.974 618	1.292 983	1.872 804	2.396 015
0.88	0.365 248	0.534 242	0.753 717	0.991 030	1.328 346	1.972 845	2.603 073
0.90	0.366 343	0.537 366	0.761 640	1.007 592	1.365 162	2.085 430	2.861 294
0.92	0.367 398	0.540 416	0.769 492	1.024 319	1.403 579	2.213 746	3.197 522
0.94	0.368 417	0.543 395	0.777 277	1.041 223	1.443 764	2.362 240	3.663 765
0.96	0.369 408	0.546 309	0.784 998	1.058 319	1.485 911	2.537 354	4.379 018
0.98	0.370 365	0.549 155	0.792 658	1.075 622	1.530 241	2.748 837	5.707 102

表 6.6-25 最佳线性无偏估计用表（正态及对数正态分布）

n	r	j	$C'(n, r, j)$	$D'(n, r, j)$	n	r	j	$C'(n, r, j)$	$D'(n, r, j)$
5	2	1	-1.497 1	-0.741 1	5	4	1	-0.511 7	0.125 2
5	2	2	1.497 1	1.741 1	5	4	2	-0.166 8	0.183 0
					5	4	3	0.027 4	0.214 7
5	3	1	-0.789 6	-0.063 8	5	4	4	0.651 1	0.477 1
5	3	2	-0.212 1	0.149 8					
5	3	3	0.981 7	0.914 0	5	5	1	-0.372 4	0.200 0

续表 6.6-25

n	r	j	$C'(n, r, j)$	$D'(n, r, j)$
5	5	2	−0.135 2	0.200 0
5	5	3	0.000 0	0.200 0
5	5	4	0.135 2	0.200 0
5	5	5	0.372 4	0.200 0
6	2	1	−1.598 8	−1.026 1
6	2	2	1.598 8	2.026 1
6	3	1	−0.824 4	−0.215 9
6	3	2	−0.276 0	0.064 9
6	3	3	1.100 4	1.151 1
6	4	1	−0.552 8	0.018 5
6	4	2	−0.209 1	0.122 6
6	4	3	−0.029 0	0.176 1
6	4	4	0.790 9	0.682 8
6	5	1	−0.409 7	0.118 3
6	5	2	−0.168 5	0.151 0
6	5	3	−0.040 6	0.168 0
6	5	4	0.074 0	0.182 8
6	5	5	0.544 8	0.379 9
6	6	1	−0.317 5	0.166 7
6	6	2	−0.138 6	0.166 7
6	6	3	−0.043 2	0.166 7
6	6	4	0.043 2	0.166 7
6	6	5	0.138 6	0.166 7
6	6	6	0.315 7	0.166 7
7	2	1	−1.681 2	−1.273 3
7	2	2	1.681 2	2.273 3
7	3	1	−0.868 2	−0.347 4
7	3	2	−0.326 9	−0.013 5
7	3	3	1.195 1	1.360 9
7	4	1	−0.584 8	−0.073 6
7	4	2	−0.242 8	0.067 7
7	4	3	−0.071 7	0.137 5
7	4	4	0.899 4	0.868 6
7	5	1	−0.437 0	0.046 5
7	5	2	−0.194 3	0.107 2
7	5	3	−0.071 8	0.137 5
7	5	4	0.031 2	0.162 6
7	5	5	0.670 9	0.546 2
7	6	1	−0.344 0	0.108 8
7	6	2	−0.161 0	0.129 5
7	6	3	−0.068 1	0.140 0
7	6	4	0.011 4	0.148 7
7	6	5	0.090 1	0.157 1
7	6	6	0.471 6	0.315 9
7	7	1	−0.277 8	0.142 9
7	7	2	−0.135 1	0.1429
7	7	3	−0.062 5	0.1429
7	7	4	0.000 0	0.142 9
7	7	5	0.062 5	0.142 9
7	7	6	0.135 1	0.142 9
7	7	7	0.277 8	0.142 9
8	2	1	−1.750 2	−1.491 5
8	2	2	1.750 2	2.491 5
8	3	1	−0.904 5	−0.463 2
8	3	2	−0.369 0	−0.085 5
8	3	3	1.273 5	1.548 7
8	4	1	−0.611 0	−0.154 9
8	4	2	−0.270 7	0.017 6
8	4	3	−0.106 1	0.100 1
8	4	4	0.987 8	1.037 2
8	5	1	−0.458 6	−0.016 7
8	5	2	−0.215 6	0.067 7
8	5	3	−0.097 0	0.108 4
8	5	4	0.000 2	0.141 3
8	5	5	0.770 9	0.6993
8	6	1	−0.363 8	0.056 9
8	6	2	−0.178 8	0.096 2
8	6	3	−0.088 1	0.115 3
8	6	4	−0.013 2	0.130 9
8	6	5	0.057 0	0.145 1
8	6	6	0.586 8	0.455 5
8	7	1	−0.297 8	0.099 7
8	7	2	−0.151 5	0.113 9
8	7	3	−0.079 6	0.120 8
8	7	4	−0.020 0	0.126 5
8	7	5	0.036 4	0.131 8
8	7	6	0.095 1	0.1370
8	7	7	0.417 5	0.270 4
8	8	1	−0.247 6	0.125 0
8	8	2	−0.129 4	0.125 0
8	8	3	−0.071 3	0.125 0
8	8	4	−0.023 0	0.125 0
8	8	5	0.023 0	0.125 0
8	8	6	0.071 3	0.125 0
8	8	7	0.129 4	0.125 0
8	8	8	0.247 6	0.125 0
9	2	1	−1.809 2	−1.686 8
9	2	1	1.809 2	2.686 8
9	3	1	−0.935 5	−0.566 4
9	3	2	−0.404 7	−0.152 1
9	3	3	1.340 2	1.718 5

续表 6.6-25

n	r	j	C′ (n, r, j)	D′ (n, r, j)
9	4	1	-0.633 0	-0.227 2
9	4	2	-0.294 4	-0.028 4
9	4	3	-0.134 8	0.064 4
9	4	4	1.062 2	1.191 2
9	5	1	-0.476 6	-0.073 1
9	5	2	-0.233 5	0.031 6
9	5	3	-0.118 1	0.080 9
9	5	4	-0.025 6	0.119 9
9	5	5	0.853 7	0.840 8
9	6	1	-0.379 7	0.010 4
9	6	2	-0.193 6	0.066 0
9	6	3	-0.104 8	0.092 3
9	6	4	-0.033 3	0.113 3
9	6	5	0.031 7	0.132 0
9	6	6	0.679 7	0.586 0
9	7	1	-0.312 9	0.060 2
9	7	2	-0.164 7	0.087 6
9	7	3	-0.093 8	0.100 6
9	7	4	-0.036 4	0.111 0
9	7	5	0.016 0	0.120 4
9	7	6	0.067 8	0.129 4
9	7	7	0.523 9	0.390 9
9	8	1	-0.263 3	0.091 5
9	8	2	-0.142 1	0.101 8
9	8	3	-0.084 1	0.106 7
9	8	4	-0.037 0	0.110 6
9	8	5	0.006 2	0.114 2
9	8	6	0.049 2	0.117 7
9	8	7	0.095 4	0.121 2
9	8	8	0.375 7	0.236 5
9	9	1	-0.223 7	0.111 1
9	9	2	-0.123 3	0.111 1
9	9	3	-0.075 1	0.111 1
9	9	4	-0.036 0	0.111 1
9	9	5	0.000 0	0.111 1
9	9	6	0.036 0	0.111 1
9	9	7	0.075 1	0.111 1
9	9	8	0.123 3	0.111 1
9	9	9	0.223 7	0.111 1
10	2	1	-1.860 6	-0.863 4
10	2	2	1.860 6	2.863 4
10	3	1	-0.962 5	-0.659 6
10	3	2	-0.435 7	-0.213 8
10	3	3	1.398 1	1.873 4
10	4	1	-0.652 0	-0.292 3
10	4	2	-0.315 0	-0.070 9
10	4	3	-0.159 3	0.030 5
10	4	4	1.126 3	1.332 7
10	5	1	-0.491 9	-0.124 0
10	5	2	-0.249 1	-0.001 6
10	5	3	-0.136 2	0.054 9
10	5	4	-0.047 2	0.099 0
10	5	5	0.924 3	0.971 8
10	6	1	-0.393 0	-0.031 6
10	6	2	-0.206 3	0.038 3
10	6	3	-0.119 2	0.070 7
10	6	4	-0.050 1	0.096 2
10	6	5	0.011 1	0.118 5
10	6	6	0.757 6	0.707 8
10	7	1	-0.325 2	0.024 4
10	7	2	-0.175 8	0.063 6
10	7	3	-0.105 8	0.081 8
10	7	4	-0.050 2	0.096 2
10	7	5	-0.000 6	0.108 9
10	7	6	0.046 9	0.120 7
10	7	7	0.610 7	0.504 5
10	8	1	-0.275 3	0.060 5
10	8	2	-0.152 3	0.080 4
10	8	3	-0.094 7	0.089 8
10	8	4	-0.048 8	0.097 2
10	8	5	-0.007 7	0.103 7
10	8	6	0.031 9	0.109 9
10	8	7	0.072 2	0.116 1
10	8	8	0.474 6	0.342 4
10	9	1	-0.234 6	0.084 3
10	9	2	-0.133 4	0.092 1
10	9	3	-0.085 1	0.095 7
10	9	4	-0.046 5	0.098 6
10	9	5	-0.011 9	0.101 1
10	9	6	0.021 5	0.103 6
10	9	7	0.055 9	0.106 6
10	9	8	0.093 7	0.108 5
10	9	9	0.342 3	0.210 1
10	10	1	-0.204 4	0.100 0
10	10	2	-0.117 2	0.100 0
10	10	3	-0.076 3	0.100 0
10	10	4	-0.043 6	0.100 0
10	10	5	-0.014 2	0.100 0
10	10	6	0.014 2	0.100 0
10	10	7	0.043 6	0.100 0
10	10	8	0.076 3	0.100 0
10	10	9	0.117 2	0.100 0
10	10	10	0.204 4	0.100 0
11	2	1	-1.906 5	-2.024 5
11	2	2	1.906 5	3.024 5
11	3	1	-0.986 2	-0.744 5
11	3	2	-0.463 6	-0.271 2

续表 6.6-25

n	r	j	C′ (n, r, j)	D′ (n, r, j)
11	3	3	1.449 2	2.015 7
11	4	1	-0.668 7	-0.351 6
11	4	2	-0.333 2	-0.110 4
11	4	3	-0.180 7	-0.001 6
11	4	4	1.182 5	1.463 6
11	5	1	-0.505 5	-0.170 2
11	5	2	-0.262 7	-0.032 3
11	5	3	-0.151 9	0.030 3
11	5	4	-0.065 7	0.078 6
11	5	5	0.985 7	1.093 7
11	6	1	-0.404 5	-0.069 8
11	6	2	-0.217 5	0.012 8
11	6	3	-0.131 7	0.050 4
11	6	4	-0.064 7	0.079 7
11	6	5	-0.006 1	0.104 9
11	6	6	0.824 6	0.822 0
11	7	1	-0.335 7	-0.008 2
11	7	2	-0.185 4	0.041 5
11	7	3	-0.116 3	0.064 2
11	7	4	-0.062 1	0.082 0
11	7	5	-0.014 6	0.097 4
11	7	6	0.029 9	0.111 6
11	7	7	0.684 2	0.611 6
11	8	1	-0.285 2	0.032 0
11	8	2	-0.161 0	0.060 9
11	8	3	-0.103 8	0.074 1
11	8	4	-0.058 9	0.084 5
11	8	5	-0.019 4	0.093 5
11	8	6	0.017 8	0.102 0
11	8	7	0.054 5	0.110 1
11	8	8	0.556 2	0.443 0
11	9	1	-0.246 3	0.059 2
11	9	2	-0.141 7	0.074 4
11	9	3	-0.093 4	0.081 4
11	9	4	-0.055 5	0.086 9
11	9	5	-0.022 0	0.091 7
11	9	6	0.009 5	0.096 2
11	9	7	0.040 9	0.100 5
11	9	8	0.073 6	0.104 9
11	9	9	0.434 9	0.304 7
11	10	1	-0.214 9	0.078 1
11	10	2	-0.125 6	0.084 1
11	10	3	-0.084 3	0.086 9
11	10	4	-0.051 9	0.089 1
11	10	5	-0.023 3	0.091 0
11	10	6	0.003 8	0.092 8
11	10	7	0.030 9	0.094 5
11	10	8	0.059 3	0.096 3
11	10	9	0.091 1	0.098 2
11	10	10	0.314 9	0.189 1
11	11	1	-0.188 3	0.090 9
11	11	2	-0.111 5	0.090 9
11	11	3	-0.076 0	0.090 9
11	11	4	-0.048 1	0.090 9
11	11	5	-0.023 4	0.090 9
11	11	6	0.000 0	0.090 9
11	11	7	0.023 4	0.090 9
11	11	8	0.048 1	0.090 9
11	11	9	0.076 0	0.090 9
11	11	10	0.111 5	0.090 9
11	11	11	0.188 3	0.090 9
12	2	1	-1.947 4	-2.172 8
12	2	2	1.947 4	3.172 8
12	3	1	-1.007 5	-0.822 5
12	3	2	-0.487 4	-0.324 9
12	3	3	1.494 8	2.147 4
12	4	1	-0.683 6	-0.405 9
12	4	2	-0.349 3	-0.147 2
12	4	3	-0.199 6	-0.032 1
12	4	4	1.232 4	1.585 2
12	5	1	-0.517 1	-0.212 5
12	5	2	-0.274 9	-0.058 9
12	5	3	-0.165 9	0.007 0
12	5	4	-0.082 0	0.050 9
12	5	5	1.039 9	1.207 5
12	6	1	-0.414 6	-0.104 8
12	6	2	-0.227 4	0.010 9
12	6	3	-0.142 8	0.031 3
12	6	4	-0.077 4	0.063 7
12	6	5	-0.021 0	0.091 5
12	6	6	0.383 3	0.929 2
12	7	1	-0.344 8	-0.038 2
12	7	2	-0.193 9	0.021 0
12	7	3	-0.122 5	0.047 7
12	7	4	-0.074 6	0.068 4
12	7	5	-0.026 7	0.086 1
12	7	6	0.015 5	0.102 2
12	7	7	0.747 9	0.712 8
12	8	1	-0.293 7	0.005 7
12	8	2	-0.168 6	0.042 8
12	8	3	-0.111 9	0.059 5
12	8	4	-0.067 8	0.072 4
12	8	5	-0.029 6	0.083 6
12	8	6	0.005 8	0.093 8
12	8	7	0.040 0	0.103 6
12	8	8	0.625 9	0.538 6

续表 6.6-25

n	r	j	C′ (n, r, j)	D′ (n, r, j)
12	9	1	-0.254 5	0.036 0
12	9	2	-0.148 7	0.058 1
12	9	3	-0.100 7	0.068 2
12	9	4	-0.063 3	0.075 9
12	9	5	-0.030 8	0.082 7
12	9	6	-0.000 7	0.088 8
12	9	7	0.028 6	0.094 8
12	9	8	0.058 2	0.100 6
12	9	9	0.511 9	0.395 0
12	10	1	-0.223 2	0.057 4
12	10	2	-0.132 4	0.069 3
12	10	3	-0.091 1	0.074 7
12	10	4	-0.059 0	0.078 9
12	10	5	-0.031 0	0.082 5
12	10	6	-0.005 0	0.085 9
12	10	7	0.020 3	0.089 1
12	10	8	0.046 1	0.092 3
12	10	9	0.073 3	0.095 6
12	10	10	0.402 0	0.274 5
12	11	1	-0.197 2	0.072 6
12	11	2	-0.118 5	0.077 5
12	11	3	-0.082 7	0.079 6
12	11	4	-0.054 8	0.081 3
12	11	5	-0.030 5	0.082 8
12	11	6	-0.007 9	0.084 2
12	11	7	0.014 2	0.085 5
12	11	8	0.036 7	0.086 8
12	11	9	0.060 8	0.088 2
12	11	10	0.088 1	0.089 6
12	11	11	0.291 9	0.171 9
12	12	1	-0.174 8	0.083 3
12	12	2	-0.106 1	0.083 3
12	12	3	-0.074 9	0.083 3
12	12	4	-0.050 6	0.083 3
12	12	5	-0.029 4	0.083 3
12	12	6	-0.009 7	0.083 3
12	12	7	0.009 7	0.083 3
12	12	8	0.029 4	0.083 3
12	12	9	0.050 6	0.083 3
12	12	10	0.074 9	0.083 3
12	12	11	0.106 1	0.083 3
12	12	12	0.174 8	0.083 3
13	2	1	-1.984 5	-2.310 1
13	2	2	1.984 5	3.310 1
13	3	1	-1.026 6	-0.894 6
13	3	2	-0.509 4	-0.375 3
13	3	3	1.536 0	2.269 9
13	4	1	-0.696 9	-0.456 1
13	4	2	-0.363 8	-0.181 7
13	4	3	-0.216 5	-0.061 0
13	4	4	1.277 3	1.698 8
13	5	1	-0.527 6	-0.251 6
13	5	2	-0.285 9	-0.087 6
13	5	3	-0.178 5	-0.015 1
13	5	4	-0.096 4	0.040 0
13	5	5	1.088 4	1.314 3
13	6	1	-0.423 6	-0.137 1
13	6	2	-0.236 3	-0.033 0
13	6	3	-0.152 8	0.013 2
13	6	4	-0.088 8	0.048 4
13	6	5	-0.034 1	0.078 4
13	6	6	0.935 5	1.030 1
13	7	1	-0.352 8	-0.065 9
13	7	2	-0.201 5	0.002 0
13	7	3	-0.133 9	0.032 2
13	7	4	-0.081 9	0.055 3
13	7	5	-0.037 4	0.075 0
13	7	6	0.003 2	0.092 8
13	7	7	0.804 2	0.808 5
13	8	1	-0.301 1	-0.018 5
13	8	2	-0.175 4	0.025 9
13	8	3	-0.119 1	0.045 7
13	8	4	-0.075 8	0.061 0
13	8	5	-0.038 6	0.074 0
13	8	6	-0.004 6	0.085 7
13	8	7	0.027 8	0.096 8
13	8	8	0.686 7	0.629 4
13	9	1	-0.261 6	0.014 4
13	9	2	-0.154 9	0.043 0
13	9	3	-0.107 1	0.055 7
13	9	4	-0.070 3	0.065 5
13	9	5	-0.038 6	0.073 9
13	9	6	-0.009 5	0.081 6
13	9	7	0.018 2	0.088 8
13	9	8	0.045 6	0.095 8
13	9	9	0.578 1	0.481 3
13	10	1	-0.230 1	0.038 0
13	10	2	-0.138 2	0.055 5
13	10	3	-0.097 0	0.063 3
13	10	4	-0.065 3	0.069 3
13	10	5	-0.037 9	0.074 5
13	10	6	-0.012 8	0.079 2
13	10	7	0.011 3	0.083 6
13	10	8	0.035 2	0.088 0
13	10	9	0.059 8	0.092 4
13	10	10	0.475 0	0.356 4

表 6.6-26　简单线性无偏估计用表（正态及对数正态分布）

n	r	$E(Y_{r,n})$	$nk_{r,n}$	n	r	$E(Y_{r,n})$	$nk_{r,n}$
20	4	−0.921 0	1.643 1	35	5	−1.112 30	2.021 3
	8	−0.314 9	4.906 5		10	−0.604 3	5.560 2
	12	0.187 0	9.669 4		15	−0.215 1	10.204 9
	16	0.745 4	17.253 1		20	0.142 8	16.288 4
					25	0.520 8	24.623 5
22	5	−0.815 3	2.279 6		30	0.997 9	37.573 1
	10	−0.169 9	6.715 8				
	15	0.405 6	13.638 1	40	5	−1.203 3	1.961 4
	20	1.188 2	27.132 7		10	−0.709 9	5.323 2
					15	−0.349 8	9.617 0
24	5	−0.876 8	2.223 9		20	−0.031 2	15.025 8
	10	−0.261 6	6.445 2		25	0.284 2	21.969 5
	15	0.261 6	12.724 4		30	0.631 8	31.375 2
	20	0.876 8	23.267 6		35	1.083 8	45.894 1
26	5	−0.931 7	2.176 0	45	5	−1.271 7	1.912 7
	10	−0.341 0	6.223 0		10	−0.797 9	5.137 1
	15	0.143 9	12.032 7		15	−0.459 1	9.174 3
	20	0.667 9	20.986 2		20	−0.167 1	14.128 4
					25	0.111 1	20.247 0
28	5	−0.981 2	2.134 3		30	0.398 3	28.010 8
	10	−0.411 0	6.036 2		35	0.723 8	38.451 1
	15	0.044 4	11.485 7		40	1.155 8	54.501 2
	20	0.509 8	19.420 5				
	25	1.137 0	33.346 2	50	5	−1.331 1	1.872 0
					10	−0.873 2	4.985 8
30	5	−1.026 1	2.097 4		15	−0.550 8	8.825 5
	10	−0.473 3	5.876 5		20	−0.278 1	13.449 0
	15	−0.041 5	11.039 0		25	−0.025 0	19.013 5
	20	0.382 4	18.256 4		30	0.226 5	25.806 3
	25	0.894 4	29.587 5		35	0.493 5	34.360 7
					40	0.802 3	45.807 7
					45	1.218 5	63.357 7

表 6.6-27　z_R 与可靠度 R 的关系（正态分布）$R=\Phi(z_R)$

R	$0.9^{1}0$	$0.9^{2}0$	$0.9^{3}0$	$0.9^{4}0$	$0.9^{5}0$	$0.9^{6}0$	$0.9^{7}0$	$0.9^{8}0$	$0.9^{9}0$	$0.9^{10}0$
z_R	1.282	2.326	3.090	3.719	4.625	4.753	5.199	5.612	5.997	6.36
R	$0.9^{11}0$	$0.9^{12}0$	$0.9^{13}0$	$0.9^{14}0$	$0.9^{15}0$	$0.9^{16}0$	$0.9^{17}0$	$0.9^{18}0$	$0.9^{19}0$	$0.9^{20}0$
z_R	6.70	7.03	7.34	7.65	7.94	8.22	8.49	8.75	9.01	9.26

注：$0.9^{4}0=0.999\,90$，其他类似。

表 6.6-28　不同置信水平 γ 时的 z_γ

γ	0.50	0.60	0.70	0.80	0.90	0.95	0.99	0.995	0.999
z_γ	0.00	0.253 4	0.524 4	0.841 6	1.282	1.645	2.326	2.576	3.090

$$\hat{\mu}=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{11}(608+622+\cdots+688)=647.91\ \text{MPa}$$

标准差的点估计，按式（6.6-163）

$$\hat{\sigma}=s_x=\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{\frac{1}{2}}$$

$$=\left\{\frac{1}{11-1}\left[(608-647.91)^2+(622-647.91)^2+\cdots+(688-647.91)^2\right]\right\}^{\frac{1}{2}}=23.33\ \text{MPa}$$

均值的双侧置信限，按式（6.6-164）、式（6.6-165），并由附表 3 查得 $t_{\frac{\alpha}{2}}(\nu)=t_{\frac{0.2}{2}}(11-1)=1.372$，故得

$$\bar{x}_L=\bar{x}-\frac{s_x}{\sqrt{n}}t_{\frac{\alpha}{2}}(\nu)=647.91-\frac{23.33}{\sqrt{11}}\times1.372=638.26\ \text{MPa}$$

$$\bar{x}_U=\bar{x}+\frac{s_x}{\sqrt{n}}t_{\frac{\alpha}{2}}(\nu)=647.91+\frac{23.33}{\sqrt{11}}\times1.372=657.56\ \text{MPa}$$

标准差的双侧置信限，按式（6.6-167）、式（6.6-168），并由附表 2 查得 $\chi^2_{\frac{\alpha}{2}}(\nu)=\chi^2_{\frac{0.2}{2}}(11-1)=15.987$，$\chi^2_{1-\frac{\alpha}{2}}(\nu)=\chi^2_{1-\frac{0.2}{2}}(11-1)=4.865$，故得

$$s_{xL}=\left[\frac{n-1}{\chi^2_{\frac{\alpha}{2}}(\nu)}\right]^{\frac{1}{2}}s_x=\left[\frac{11-1}{15.987}\right]^{\frac{1}{2}}\times23.33=18.45\ \text{MPa}$$

$$s_{xU}=\left[\frac{n-1}{\chi^2_{1-\frac{\alpha}{2}}(\nu)}\right]^{\frac{1}{2}}s_x=\left[\frac{11-1}{4.865}\right]^{\frac{1}{2}}\times23.33=33.45\ \text{MPa}$$

失效概率 $F=0.10$ 的强度极限，可借用可靠寿命与可靠度的关系式。这时相当于 $R=1-F=0.90$ 时的强度极限。点估计可用式（6.6-176）来求，并由 $R=0.90$ 查表 6.6-27 得 $z_R=1.28$，故

$$\hat{x}_F=\hat{x}(R)=\bar{x}-z_R s_x=647.91-1.28\times23.33=618.05\ \text{MPa}$$

失效概率 $F=0.10$，置信水平 $\gamma=90\%$ 强度极限的单侧置信下限可用式（6.6-178）来求，并由 $R=1-F=0.90$ 查表 6.6-27 得 $z_{R\gamma}=2.011\,29$，故

$$x_{FL}=x_L(R)=\bar{x}-z_{R\gamma}s_x=647.91-2.011\,29\times23.33=600.99\ \text{MPa}$$

例 24　某金属材料在某应力水平用 10 个试件做弯曲疲劳试验，其中 8 个试件失效的循环次数分别为 125 000，132 000，135 000，138 000，141 000，145 000，148 000，152 000次，其他 2 个试件超过 152 000 次未失效就停止了试验。已知该金属材料疲劳寿命服从对数正态分布。试用极大似然法估计分布参数及可靠度 R 为 0.90 时的寿命点估计。

解　这是定数截尾试验。先求出式（6.6-170）、式（6.6-171）中所需的各值

$$\bar{x}_r=\frac{1}{r}\sum_{i=1}^{r}x_i=\frac{1}{r}\sum_{i=1}^{r}\ln N_i$$

$$=\frac{1}{8}(\ln125\,000+\ln132\,000+\cdots+\ln152\,000)$$

$$=11.844$$

$$s_r^2=\frac{1}{r}\sum_{i=1}^{r}(x_i-\bar{x}_r)^2=\frac{1}{r}\sum_{i=1}^{r}(\ln N_i-\bar{x}_r)^2$$

$$=\frac{1}{8}\left[(\ln125\,000-11.844)^2+(\ln132\,000-11.844)^2+\cdots+(152\,000-11.844)^2\right]=0.029$$

$$d=x_0-\bar{x}_r=\ln N_0-\bar{x}_r=\ln152\,000-11.844=0.088$$

$$D=\frac{d^2}{d^2+s_r^2}=\frac{0.088^2}{0.088^2+0.029^2}=0.211$$

由表 6.6-24，按 $\frac{r}{n}=\frac{8}{10}=0.8$，查得 $g\left(D,\frac{r}{n}\right)=0.48157$，代入式（6.6-170）、式（6.6-171）得

$$\hat{\mu}=\bar{x}_r+\frac{s_x^2-s_r^2}{d}=11.844+\frac{0.183^2-0.029^2}{0.088}=11.895$$

$$\hat{\sigma}=s_x=\frac{d}{g\left(D,\ \frac{r}{n}\right)}=\frac{0.088}{0.481\ 57}=0.183$$

求可靠寿命先用式（6.6-178），并按 $R=0.9$ 由表6.6-27查得 $z_R=1.28$，故得

$$\hat{x}\ (0.90)\ =\bar{x}-z_R s_x=11.895-1.28\times0.183=11.66$$

再按式（6.6-159）进行反变换得

$$\hat{N}\ (0.90)\ =e^{\hat{x}(0.90)}=e^{11.66}=115\ 844$$

例25　数据同例24。试用最佳线性无偏估计法求分布参数。

解　按式（6.6-172）、式（6.6-173）列表计算，见表6.6-29表中系数 D'（10，8，j）和 C'（10，8，j）的数值由表6.6-25查得。

表6.6-29　例25的列表计算

j	N_j	$x_j=\ln N_j$	$D'(10,8,j)$	$D'(10,8,j)x_j$	$C'(10,8,j)$	$C'(10,8,j)x_j$
1	125 000	11.736	0.060 5	0.710 0	−0.275 3	−3.230 9
2	132 000	11.791	0.080 4	0.948 0	−0.152 3	−1.795 8
3	135 000	11.813	0.089 8	1.060 8	−0.094 7	−1.118 7
4	138 000	11.835	0.097 2	1.150 4	−0.048 8	−0.577 5
5	141 000	11.857	0.103 7	1.229 6	−0.007 7	−0.091 3
6	145 000	11.884	0.109 9	1.306 1	0.031 9	0.379 1
7	148 000	11.905	0.116 1	1.382 2	0.072 2	0.859 5
8	152 000	11.932	0.342 4	4.085 5	0.474 6	5.662 9

$$\hat{\mu}=\bar{x}=\sum_{j=1}^{r}D'\ (10,\ 8,\ j)\ x_j$$
$$=0.710\ 0+0.948\ 0+\cdots+4.085\ 5=11.872\ 6$$

$$\hat{\sigma}=s_x=\sum_{j=1}^{r}C'\ (10,\ 8,\ j)\ x_j$$
$$=-3.230\ 9-1.795\ 8-\cdots-0.091\ 3+0.379\ 1+0.859\ 5+5.662\ 9=0.087\ 3$$

4.3　威布尔分布的分析法

威布尔分布的概率密度

$$f\ (x)\ =\frac{k}{b}\left(\frac{x-a}{b}\right)^{k-1}e^{-\left(\frac{x-a}{b}\right)^k}\qquad(6.6\text{-}183)$$

当位置参数 $a=0$ 时，则为较简单的两参数威布尔分布

$$f\ (x)\ =\frac{k}{b}\left(\frac{x}{b}\right)^{k-1}e^{-\left(\frac{x}{b}\right)^k}\qquad(6.6\text{-}184)$$

4.3.1　拟合性检验

样本大小为 n，截尾寿命试验得 $t_1\leqslant t_2\leqslant\cdots\leqslant t_r$。检验统计量观测值

$$W=\frac{r_0\sum\limits_{i=r_0+1}^{r-1}l_i}{(r-r_0-1)\sum\limits_{i=1}^{r_0}l_i}\qquad(6.6\text{-}185)$$

$$l_i=\frac{\ln t_{i+1}-\ln t_i}{E\ (z_{i+1})\ -E\ (z_i)},\ i=1,\ 2,\ \cdots,\ r-1$$

式中，$r_0=\left[\frac{r}{2}\right]$，即取括号内整数部分；$E$（$z_i$）查表6.6-33、表6.6-34中的 E（$z_{i,n}$），$i=1,\ 2,\ \cdots,\ r-1$。

满足下式条件则接受两参数威布尔分布的假设，否则拒绝两参数威布尔分布的假设：

$$F_{1-\frac{\alpha}{2}}[2(r-r_0-1),2r_0]\leqslant W\leqslant F_{\frac{\alpha}{2}}[2(r-r_0-1),2r_0]\quad(6.6\text{-}186)$$

式中，α 为显著性水平；F（ν_1，ν_2）为自由度为 $\nu_1=2\ (r-r_0-1)$，$\nu_2=2\ r_0$ 的 F 分布的分位数，查附表4。

$$F_{1-\alpha}\ (\nu_1,\ \nu_2)\ =\frac{1}{F_\alpha\ (\nu_2,\ \nu_1)}$$

若拒绝了两参数威布尔分布，仍有可能服从三参数威布尔分布。对三参数威布尔分布，仍可用上式检验，但此时应先估计出位置参数 a，将式中的 t_i 换成 $\tau_i=t_i-a$ 即可。

例26　抽取某产品15个进行寿命试验，其寿命由小到大为8 300，15 800，22 200，27 500，31 200，37 600，45 000，46 800，53 600，61 000，68 100，78 200，85 000，95 500，124 500 h。检验该产品寿命是否服从威布尔分布。

解　假设该产品寿命服从两参数威布尔分布。用式（6.6-186）计算统计量的观测值，先列表计算式中的 l_i，见表6.6-30。表中 E（z_i）值查表6.6-34。

表6.6-30　例26中 l_i 的列表计算

i	t_i	$\ln t_i$	$\ln t_{i+1}-\ln t_i$	$E\ (z_i)$	$E\ (z_{i+1})\ -E\ (z_i)$	$l_i=\frac{\ln t_{i+1}-\ln t_i}{E\ (z_{i+1})\ -E\ (z_i)}$
1	8 300	9.024	0.644	−3.285 3	1.034 9	0.622 3
2	15 800	9.668	0.340	−2.250 4	0.537 1	0.633 0
3	22 200	10.008	0.214	−1.713 3	0.372 9	0.573 9
4	27 500	10.222	0.126	−1.340 4	0.292 6	0.430 6
5	31 200	10.348	0.187	−1.047 8	0.245 9	0.760 5
6	37 600	10.535	0.179	−0.801 9	0.216 7	0.826 0
7	45 000	10.714	0.040	−0.585 2	0.197 9	0.202 1
8	46 800	10.754	0.135	−0.387 3	0.186 3	0.724 6
9	53 600	10.889	0.130	−0.201 0	0.180 4	0.720 6
10	61 000	11.019	0.110	−0.020 6	0.180 1	0.610 8
11	68 100	11.129	0.138	0.159 5	0.186 3	0.740 7
12	78 200	11.267	0.083	0.345 8	0.202 7	0.409 5
13	85 000	11.350	0.117	0.548 5	0.239 9	0.487 7
14	95 500	11.467	0.265	0.788 4	0.344 3	0.769 7
15	124 500	11.732		1.132 7		

按式 (6.6-185)，式中，$r_0=\left(\frac{r}{2}\right)=\left(\frac{15}{2}\right)=7$

$$W=\frac{7\sum_{i=8}^{14}l_i}{7\sum_{i=1}^{7}l_i}$$
$$=\frac{0.7246+0.7206+0.6108+0.7407+0.4095+0.4877+0.7697}{0.6223+0.6330+0.5739+0.4306+0.7605+0.8260+0.2021}$$
$$=1.1026$$

取显著性水平 $a=0.10$，查附表4得

$F_{\frac{a}{2}}\ (2\ (r-r_0-1),\ 2r_0)\ =F_{0.05}\ (14,\ 14)\ =2.48$

$F_{1-\frac{a}{2}}(2(r-r_0-1),2r_0)=F_{0.95}(14,14)$

$$=\frac{1}{F_{0.05}(14,14)}=\frac{1}{2.48}=0.4032$$

由于 $0.403<W=1.1026<2.48$，满足式 (6.6-186) 的条件，故接受该产品的寿命服从两参数的威布尔分布的假设。

4.3.2 参数估计

1) 矩法估计 样本均值

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \tag{6.6-187}$$

样本标准差

$$s_x=\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{\frac{1}{2}} \tag{6.6-188}$$

样本偏态系数

$$k_k=\frac{n}{(n-1)\ (n-2)\ s_x^3}\sum_{i=1}^{n}(x_i-\bar{x})^3 \tag{6.6-189}$$

根据 K_k 值由表6.6-31查得形状参数 k 的点估计 $\hat{k}$，同时可查得系数 k_a、k_b，故尺度参数 b 的点估计

$$\hat{b}=\frac{s_x}{k_b} \tag{6.6-190}$$

位置参数 a 的点估计

$$\hat{a}=\begin{cases}\bar{x}-\hat{b}k_a, & 若\ \bar{x}-\hat{b}k_a\leqslant\min\{x_i\}\\ \min\{x_i\}, & 若\ \bar{x}-\hat{b}k_a>\min\{x_i\}\end{cases} \tag{6.6-191}$$

若位置参数 a 已知（例如两参数威布尔分布 $a=0$），则求

$$k_c=\frac{s_x}{\bar{x}-a} \tag{6.6-192}$$

根据 k_c 值由表6.6-31查得形状参数 k 的点估计 $\hat{k}$，同时查得 k_b，再由式 (6.6-190) 求尺度参数 b 的点估计 $\hat{b}$。

2) 极大似然估计 两参数威布尔分布的极大似然估计，对截尾寿命试验，若样本大小为 n，截尾时间为 x_0，失效数为 r（对完全样本只需取失效数 $r=n$），$x_1\leqslant x_2\leqslant\cdots\leqslant x_r\leqslant x_0$，则

$$\frac{\sum_{i=1}^{r}x_i^k\ln x_i+(n-r)x_0^k\ln x_0}{\sum_{i=1}^{r}x_i^k+(n-r)x_0^k}-\frac{1}{k}=\frac{1}{r}\sum_{i=1}^{r}\ln x_i \tag{6.6-193}$$

$$b=\left\{\frac{1}{r}\left[\sum_{i=1}^{r}x_i^k+(n-r)\ x_0^k\right]\right\}^{\frac{1}{k}} \tag{6.6-194}$$

用迭代法解超越方程式 (6.6-193)，求得 k 代入式 (6.6-194) 算得 b，就是形状参数 k 和尺度参数 b 的极大似然估计 $\hat{k}$ 和 $\hat{b}$。

3) 最佳线性无偏估计和简单线性无偏估计 利用相应的数表可很方便地求得精度较高的两参数威布尔参数的最佳线性无偏估计（BLUE）和简单线性无偏估计（GLUE）。这时应先做如下变换：

原分布函数

$$F\ (t)\ =1-e^{-\left(\frac{t}{b}\right)^k} \tag{6.6-195}$$

令：

$$x=\ln t \quad 即 \quad t=e^x \tag{6.6-196}$$

$$\sigma=\frac{1}{k} \quad 即 \quad k=\frac{1}{\sigma} \tag{6.6-197}$$

$$\mu=\ln b \quad 即 \quad b=e^{\mu} \tag{6.6-198}$$

则原分布函数变为

$$F\ (x)\ =1-e^{-e^{\frac{x-\mu}{\sigma}}} \tag{6.6-199}$$

式中，参数 μ 和 σ 的最佳线性无偏估计

$$\hat{\mu}=\sum_{j=1}^{r}D\ (n,\ r,\ j)\ x_j \tag{6.6-200}$$

$$\hat{\sigma}=\sum_{j=1}^{r}C\ (n,\ r,\ j)\ x_j \tag{6.6-201}$$

式中，n 为样本大小；r 为截尾失效数；j 为寿命由小到大的排列序号；x_j 为第 j 个对数寿命值；$D(n,r,j)$ 为 μ 的最佳线性无偏估计系数，查表6.6-32；$C(n,r,j)$ 为 σ 的最佳线性无偏估计系数，查表6.6-32。

当样本容量较大时，μ 和 σ 用简单线性无偏估计。μ 和 σ 的简单线性无偏估计

$$\hat{\mu}=x_{s,n}-E\ (z_{r,n})\ \hat{\sigma} \tag{6.6-202}$$

$$\hat{\sigma}=\frac{1}{nk_{r,n}}\left[(2s-r)x_{s,n}-\sum_{j=1}^{s}x_j+\sum_{j=s+1}^{r}x_j\right] \tag{6.6-203}$$

式中，n、r、j、x_j 同式 (6.6-200)；s、$nk_{r,n}$、$E\ (z_{r,n})$ 查表6.6-33。

形状参数 k 的点估计按式 (6.6-197)，为得到 k 的无偏估计再加以修正得

$$\hat{k}=\frac{g_{r,n}}{\hat{\sigma}} \tag{6.6-204}$$

式中，$g_{r,n}$ 为修偏系数，查表6.6-33、表6.6-34。

$$\hat{b}=e^{\hat{\mu}} \tag{6.6-205}$$

4.3.3 可靠度和可靠寿命的估计

指定寿命 t 时可靠度 R 的点估计

$$\hat{R}\ (t)\ =\begin{cases}e^{-\left(\frac{t-\hat{a}}{\hat{b}}\right)^{\hat{k}}}, & 当\ t>\hat{a}\\ 1, & 当\ t\leqslant\hat{a}\end{cases} \tag{6.6-206}$$

指定可靠度 R 时可靠寿命 t (R) 的点估计

$$\hat{t}\ (R)\ =\hat{a}+\hat{b}\left(\ln\frac{1}{R}\right)^{\frac{1}{\hat{k}}} \tag{6.6-207}$$

指定置信水平、可靠度的置信下限，对于全数试验的两参数威布尔分布，可按可靠度的极大似然点估计 $\hat{R}\ (t)$ 查表6.6-35、表6.6-36；对于截尾寿命试验，其可靠度的置信下限可查表6.6-37～表6.6-39。

可靠寿命的单侧置信下限

$$t_L\ (R)\ =e^{\hat{\mu}-\frac{\hat{\sigma}(B_{r,n}+V_{R\gamma})}{2-g_{r,n}}} \tag{6.6-208}$$

式中，$g_{r,n}$、$B_{r,n}$、$V_{R\gamma}$ 查表6.6-34。

表 6.6-31 威布尔分布形状参数和各参数点估计系数

$\hat{k}$	k_k	k_b	k_a	k_c	$\hat{k}$	k_k	k_b	k_a	k_c
0.60	4.593	2.645	1.505	1.758	3.10	0.136	0.413	0.895	0.353
0.70	3.498	1.851	0.266	1.462	3.20	0.106	0.307	0.896	0.343
0.80	2.815	1.428	1.133	1.260	3.30	0.078	0.298	0.897	0.333
0.90	2.345	1.199	1.073	1.113	3.40	0.051	0.292	0.898	0.325
1.00	2.000	1.000	1.000	1.000	3.50	0.025	0.290	0.898	0.316
1.10	1.734	0.878	0.965	0.910	3.60	0.001	0.277	0.899	0.308
1.20	1.521	0.787	0.940	0.837	3.70	-0.023	0.276	0.901	0.301
1.30	1.346	0.716	0.923	0.776	3.80	-0.045	0.265	0.904	0.294
1.40	1.198	0.660	0.911	0.724	3.90	-0.067	0.260	0.905	0.287
1.50	1.072	0.613	0.903	0.679	4.00	-0.087	0.254	0.906	0.280
1.60	0.962	0.574	0.897	0.640	4.10	-0.107	0.249	0.908	0.274
1.70	0.865	0.540	0.892	0.605	4.20	-0.126	0.244	0.909	0.268
1.80	0.779	0.511	0.889	0.575	4.30	-0.144	0.239	0.910	0.263
1.90	0.701	0.486	0.888	0.547	4.40	-0.161	0.234	0.911	0.257
2.00	0.631	0.463	0.886	0.523	4.50	-0.178	0.230	0.913	0.252
2.10	0.567	0.443	0.886	0.500	4.60	-0.195	0.225	0.914	0.247
2.20	0.509	0.425	0.886	0.480	4.70	-0.210	0.221	0.915	0.242
2.30	0.455	0.408	0.886	0.461	4.80	-0.225	0.217	0.916	0.238
2.40	0.405	0.393	0.886	0.444	4.90	-0.240	0.241	0.917	0.233
2.50	0.358	0.380	0.887	0.428	5.00	-0.254	0.210	0.918	0.229
2.60	0.315	0.367	0.888	0.413	5.10	-0.268	0.207	0.919	0.225
2.70	0.275	0.355	0.889	0.399	5.20	-0.281	0.203	0.920	0.221
2.80	0.237	0.344	0.890	0.387	5.30	-0.294	0.199	0.921	0.217
2.90	0.202	0.333	0.891	0.375	5.40	-0.306	0.197	0.922	0.213
3.00	0.168	0.325	0.893	0.363	5.50	-0.318	0.194	0.923	0.210

表 6.6-32 最佳线性无偏估计系数（威布尔分布）

n	r	j	$C(n, r, j)$	$D(n, r, j)$	n	r	j	$C(n, r, j)$	$D(n, r, j)$
5	2	1	-0.896 3	-0.959 9	6	4	4	0.737 2	1.049 6
5	2	2	0.896 3	1.959 9					
					6	5	1	-0.201 5	0.005 7
5	3	1	-0.434 3	-0.210 1	6	5	2	-0.197 3	0.046 6
5	3	2	-0.364 2	-0.086 0	6	5	3	-0.153 6	0.100 2
5	3	3	0.798 6	1.296 1	6	5	4	-0.064 6	0.172 3
					6	5	5	0.617 0	0.675 2
5	4	1	-0.273 0	-0.015 4					
5	4	2	-0.249 9	0.052 0	6	6	1	-0.145 8	0.048 9
5	4	3	-0.149 1	0.152 1	6	6	2	-0.145 9	0.083 5
5	4	4	0.672 1	0.811 3	6	6	3	-0.126 7	0.121 1
					6	6	4	-0.073 2	0.165 6
5	5	1	-0.184 5	0.058 4	6	6	5	0.036 0	0.225 5
5	5	2	-0.181 7	0.108 8	6	6	6	0.459 3	0.355 4
5	5	3	-0.130 5	0.167 6					
5	5	4	-0.006 5	0.246 3	7	2	1	-0.926 7	-1.338 3
5	5	5	0.503 1	0.418 9	7	2	2	0.926 7	2.338 3
6	2	1	-0.914 1	-0.165 6	7	3	1	-0.455 0	-0.403 6
6	2	2	0.914 1	2.165 6	7	3	2	-0.405 6	-0.301 2
					7	3	3	0.860 5	1.704 8
6	3	1	-0.446 6	-0.315 4					
6	3	2	-0.388 6	-0.203 4	7	4	1	-0.294 0	-0.146 3
6	3	3	0.835 3	1.518 8	7	4	2	-0.276 0	-0.094 1
					7	4	3	-0.210 2	0.007 1
6	4	1	-0.285 9	-0.086 5	7	4	4	0.780 2	1.247 5
6	4	2	-0.265 5	-0.028 1					
6	4	3	-0.185 9	0.064 9	7	5	1	-0.211 0	-0.039 3

续表 6.6-32

n	r	j	C (n, r, j)	D (n, r, j)
7	5	2	-0.206 5	-0.004 4
7	5	3	-0.169 1	0.045 8
7	5	4	-0.099 2	0.113 4
7	5	5	0.685 8	0.884 4
7	6	1	-0.158 7	0.013 7
7	6	2	-0.160 9	0.041 8
7	6	3	-0.139 6	0.075 7
7	6	4	-0.095 1	0.117 6
7	6	5	-0.017 6	0.172 1
7	6	6	0.571 9	0.579 1
7	7	1	-0.120 1	0.041 8
7	7	2	-0.125 9	0.067 3
7	7	3	-0.114 9	0.093 7
7	7	4	-0.087 3	0.123 2
7	7	5	-0.036 2	0.158 6
7	7	6	0.060 7	0.206 3
7	7	7	0.423 7	0.309 0
8	2	1	-0.936 1	-1.486 9
8	2	2	0.936 1	2.486 9
8	3	1	-0.461 0	-0.479 4
8	3	2	-0.418 0	-0.384 8
8	3	3	0.879 0	1.864 2
8	4	1	-0.299 8	-0.197 7
8	4	2	-0.283 7	-0.150 2
8	4	3	-0.227 5	-0.068 5
8	4	4	0.810 9	1.416 4
8	5	1	-0.217 2	-0.078 1
8	5	2	-0.212 8	-0.047 4
8	5	3	-0.180 3	-0.000 1
8	5	4	-0.122 5	0.063 7
8	5	5	0.732 8	1.061 9
8	6	1	-0.166 1	-0.017 2
8	6	2	-0.167 5	0.006 5
8	6	3	-0.148 3	0.038 0
8	6	4	-0.110 5	0.078 0
8	6	5	-0.050 0	0.129 2
8	6	6	0.642 4	0.765 5
8	7	1	-0.130 3	0.016 8
8	7	2	-0.134 8	0.037 6
8	7	3	-0.123 8	0.061 2
8	7	4	-0.099 1	0.088 8
8	7	5	-0.057 1	0.122 5
8	7	6	0.010 9	0.165 5
8	7	7	0.534 3	0.507 6
8	8	1	-0.101 9	0.036 5
8	8	2	-0.108 1	0.056 1
8	8	3	-0.102 7	0.075 9
8	8	4	-0.087 2	0.097 1
8	8	5	-0.058 9	0.121 2
8	8	6	-0.011 1	0.150 2
8	8	7	0.075 8	0.189 4
8	8	8	0.394 2	0.273 5
9	2	1	-0.943 4	-1.617 3
9	2	2	0.943 4	2.617 3
9	3	1	-0.465 6	-0.545 8
9	3	2	-0.427 5	-0.457 7
9	3	3	0.893 2	2.003 5
9	4	1	-0.304 0	-0.242 7
9	4	2	-0.289 5	-0.199 0
9	4	3	-0.240 5	-0.121 9
9	4	4	0.834 0	1.563 6
9	5	1	-0.221 7	-0.112 3
9	5	2	-0.217 4	-0.084 7
9	5	3	-0.188 7	-0.039 8
9	5	4	-0.139 4	0.020 6
9	5	5	0.767 3	1.216 1
9	6	1	-0.171 2	-0.044 6
9	6	2	-0.172 0	-0.023 9
9	6	3	-0.154 7	0.005 7
9	6	4	-0.122 0	0.044 0
9	6	5	-0.072 1	0.092 5
9	6	6	0.692 0	0.926 4
9	7	1	-0.136 4	-0.005 8
9	7	2	-0.140 0	0.011 8
9	7	3	-0.129 7	0.033 6
9	7	4	-0.107 6	0.060 0
9	7	5	-0.072 3	0.092 2
9	7	6	-0.019 4	0.132 5
9	7	7	0.605 5	0.675 7
9	8	1	-0.110 2	0.017 8
9	8	2	-0.115 4	0.034 0
9	8	3	-0.109 7	0.051 6
9	8	4	-0.095 0	0.071 4
9	8	5	-0.070 0	0.094 3
9	8	6	-0.031 2	0.121 8
9	8	7	0.029 2	0.156 9
9	8	8	0.502 4	0.452 3
9	9	1	-0.088 4	0.032 3
9	9	2	-0.094 4	0.048 0
9	9	3	-0.092 0	0.063 4
9	9	4	-0.082 7	0.079 6
9	9	5	-0.065 6	0.097 2
9	9	6	-0.038 0	0.117 4
9	9	7	0.006 5	0.141 8
9	9	8	0.085 2	0.174 9

续表 6.6-32

n	r	j	$C(n, r, j)$	$D(n, r, j)$
9	9	9	0.369 2	0.245 5
10	2	1	-0.949 1	-1.733 3
10	2	2	0.949 1	2.733 3
10	3	1	-0.469 3	-0.604 7
10	3	2	-0.435 1	-0.522 3
10	3	3	0.904 4	2.127 0
10	4	1	-0.307 3	-0.282 7
10	4	2	-0.294 1	-0.242 1
10	4	3	-0.250 6	-0.169 1
10	4	4	0.852 0	1.693 8
10	5	1	-0.225 1	-0.142 6
10	5	2	-0.221 0	-0.117 5
10	5	3	-0.195 3	-0.074 8
10	5	4	-0.152 3	-0.017 4
10	5	5	0.793 7	1.352 3
10	6	1	-0.174 8	-0.069 0
10	6	2	-0.175 4	-0.050 6
10	6	3	-0.159 6	-0.022 6
10	6	4	-0.130 8	0.014 1
10	6	5	-0.088 3	0.060 2
10	6	6	0.729 0	1.068 0
10	7	1	-0.140 6	-0.026 1
10	7	2	-0.143 7	-0.010 9
10	7	3	-0.134 0	0.009 5
10	7	4	-0.114 3	0.034 8
10	7	5	-0.083 8	0.065 9
10	7	6	-0.040 2	0.104 1
10	7	7	0.656 5	0.822 7
10	8	1	-0.115 3	0.000 6
10	8	2	-0.119 8	0.014 3
10	8	3	-0.114 2	0.030 5
10	8	4	-0.100 6	0.049 3
10	8	5	-0.078 5	0.071 4
10	8	6	-0.046 0	0.097 9
10	8	7	0.000 9	0.130 7
10	8	8	0.573 5	0.605 4
10	9	1	-0.095 3	0.017 8
10	9	2	-0.100 5	0.030 9
10	9	3	-0.097 8	0.044 6
10	9	4	-0.088 6	0.059 6
10	9	5	-0.072 7	0.076 3
10	9	6	-0.048 6	0.095 7
10	9	7	-0.012 9	0.118 8
10	9	8	0.041 5	0.148 1
10	9	9	0.474 9	0.408 2
10	10	1	-0.077 9	0.028 9
10	10	2	-0.083 6	0.041 7
10	10	3	-0.082 8	0.054 2
10	10	4	-0.077 0	0.067 0
10	10	5	-0.066 1	0.080 6
10	10	6	-0.048 7	0.095 6
10	10	7	-0.022 2	0.112 9
10	10	8	0.019 2	0.133 8
10	10	9	0.091 2	0.162 3
10	10	10	0.347 8	0.222 9
11	2	1	-0.953 8	-1.837 7
11	2	2	0.953 8	2.837 7
11	3	1	-0.472 2	-0.657 8
11	3	2	-0.441 2	-0.580 2
11	3	3	0.913 4	2.237 9
11	4	1	-0.309 9	-0.318 6
11	4	2	-0.297 9	-0.280 6
11	4	3	-0.258 8	-0.211 3
11	4	4	0.866 5	1.810 5
11	5	1	-0.227 7	-0.169 9
11	5	2	-0.223 9	-0.146 7
11	5	3	-0.200 6	-0.106 1
11	5	4	-0.162 5	-0.051 4
11	5	5	0.814 7	1.474 1
11	6	1	-0.177 7	-0.091 0
11	6	2	-0.178 0	-0.074 5
11	6	3	-0.163 6	-0.047 7
11	6	4	-0.137 9	-0.012 6
11	6	5	-0.100 7	0.031 4
11	6	6	0.757 8	1.194 4
11	7	1	-0.143 7	-0.044 5
11	7	2	-0.146 4	-0.031 1
11	7	3	-0.137 4	-0.011 8
11	7	4	-0.119 5	0.012 6
11	7	5	-0.092 7	0.042 4
11	7	6	-0.055 6	0.078 8
11	7	7	0.695 2	0.953 5
11	8	1	-0.118 9	-0.015 0
11	8	2	-0.122 8	-0.003 2
11	8	3	-0.117 5	0.011 8
11	8	4	-0.104 9	0.029 8
11	8	5	-0.085 3	0.051 2
11	8	6	-0.057 3	0.076 8
11	8	7	-0.018 7	0.107 8
11	8	8	0.625 4	0.740 7
11	9	1	-0.099 6	0.004 4
11	9	2	-0.104 3	0.015 4
11	9	3	-0.101 4	0.028 0
11	9	4	-0.092 7	0.042 1
11	9	5	-0.078 2	0.058 3

续表 6.6-32

n	r	j	$C(n, r, j)$	$D(n, r, j)$
11	9	6	-0.057 0	0.077 1
11	9	7	-0.027 2	0.099 3
11	9	8	0.014 9	0.126 6
11	9	9	0.545 6	0.548 8
11	10	1	-0.083 8	0.017 4
11	10	2	-0.088 9	0.028 2
11	10	3	-0.087 7	0.039 3
11	10	4	-0.081 9	0.051 1
11	10	5	-0.071 4	0.064 0
11	10	6	-0.055 4	0.078 4
11	10	7	-0.032 5	0.095 0
11	10	8	0.000 6	0.114 8
11	10	9	0.050 0	0.139 8
11	10	10	0.450 9	0.372 1
11	11	1	-0.069 6	0.026 2
11	11	2	-0.074 8	0.036 9
11	11	3	-0.075 0	0.047 2
11	11	4	-0.071 4	0.057 6
11	11	5	-0.064 1	0.068 5
11	11	6	-0.052 5	0.080 2
11	11	7	-0.035 3	0.093 2
11	11	8	-0.010 0	0.108 2
11	11	9	0.028 6	0.126 5
11	11	10	0.094 9	0.151 4
11	11	11	0.329 2	0.204 1
12	2	1	-0.957 7	-1.932 7
12	2	2	0.957 7	2.932 7
12	3	1	-0.474 6	-0.706 0
12	3	2	-0.446 3	-0.632 6
12	3	3	0.920 9	2.338 5
12	4	1	-0.312 0	-0.351 2
12	4	2	-0.300 9	-0.315 5
12	4	3	-0.265 4	-0.249 4
12	4	4	0.878 4	1.916 1
12	5	1	-0.229 8	-0.194 7
12	5	2	-0.226 2	-0.173 2
12	5	3	-0.205 0	-0.134 3
12	5	4	-0.170 7	-0.082 0
12	5	5	0.831 7	1.584 2
12	6	1	-0.179 9	-0.111 0
12	6	2	-0.180 1	-0.095 9
12	6	3	-0.166 8	-0.070 4
12	6	4	-0.143 6	-0.036 6
12	6	5	-0.110 6	0.005 5
12	6	6	0.781 0	1.308 5
12	7	1	-0.146 1	-0.061 2
12	7	2	-0.148 5	-0.049 2
12	7	3	-0.140 1	-0.030 9
12	7	4	-0.123 8	-0.007 5
12	7	5	-0.099 9	0.021 3
12	7	6	-0.067 5	0.056 1
12	7	7	0.725 9	1.071 3
12	8	1	-0.121 6	-0.029 3
12	8	2	-0.125 1	-0.019 0
12	8	3	-0.120 0	-0.004 9
12	8	4	-0.108 5	0.012 5
12	8	5	-0.090 7	0.033 2
12	8	6	-0.066 1	0.057 9
12	8	7	-0.033 3	0.087 5
12	8	8	0.665 3	0.862 2
12	9	1	-0.102 7	-0.008 1
12	9	2	-0.107 0	0.001 5
12	9	3	-0.104 0	0.013 1
12	9	4	-0.095 9	0.026 6
12	9	5	-0.082 6	0.042 3
12	9	6	-0.063 8	0.060 5
12	9	7	-0.038 2	0.082 0
12	9	8	-0.003 7	0.107 8
12	9	9	0.597 8	0.674 2
12	10	1	-0.087 5	0.006 6
12	10	2	-0.092 1	0.015 8
12	10	3	-0.090 8	0.025 9
12	10	4	-0.085 1	0.037 0
12	10	5	-0.075 3	0.049 4
12	10	6	-0.060 8	0.063 4
12	10	7	-0.040 8	0.079 5
12	10	8	-0.013 3	0.098 5
12	10	9	0.024 8	0.121 7
12	10	10	0.520 9	0.502 2
12	11	1	-0.074 6	0.016 8
12	11	2	-0.079 4	0.025 9
12	11	3	-0.079 3	0.035 1
12	11	4	-0.075 5	0.044 6
12	11	5	-0.068 3	0.054 9
12	11	6	-0.057 3	0.066 2
12	11	7	-0.041 7	0.078 7
12	11	8	-0.020 0	0.093 2
12	11	9	0.010 6	0.110 4
12	11	10	0.055 9	0.132 0
12	11	11	0.429 7	0.342 1
12	12	1	-0.062 9	0.023 9
12	12	2	-0.067 7	0.033 0
12	12	3	-0.068 4	0.041 6
12	12	4	-0.066 1	0.050 3
12	12	5	-0.061 1	0.059 3
12	12	6	-0.053 1	0.068 7
12	12	7	-0.041 3	0.079 0
12	12	8	-0.024 5	0.090 5
12	12	9	-0.000 5	0.103 7

续表 6.6-32

n	r	j	C (n, r, j)	D (n, r, j)
12	12	10	0.035 7	0.119 8
12	12	11	0.097 1	0.141 8
12	12	12	0.312 8	0.188 4
13	2	1	-0.961 0	-2.019 7
13	2	2	0.961 0	3.019 7
13	3	1	-0.476 7	-0.750 1
13	3	2	-0.450 5	-0.680 4
13	3	3	0.927 2	2.430 5
13	4	1	-0.313 8	-0.381 1
13	4	2	-0.303 5	-0.347 3
13	4	3	-0.271 0	-0.284 2
13	4	4	0.888 3	2.012 6
13	5	1	-0.231 6	-0.217 3
13	5	2	-0.228 2	-0.197 2
13	5	3	-0.208 7	-0.160 0
13	5	4	-0.177 5	-0.109 9
13	5	5	0.845 9	1.684 5
13	6	1	-0.181 7	-0.129 3
13	6	2	-0.181 8	-0.115 4
13	6	3	-0.169 6	-0.091 0
13	6	4	-0.148 3	-0.058 5
13	6	5	-0.118 7	-0.018 1
13	6	6	0.800 1	1.412 4
13	7	1	-0.148 1	-0.076 5
13	7	2	-0.150 1	-0.065 7
13	7	3	-0.142 3	-0.048 3
13	7	4	-0.127 4	-0.025 6
13	7	5	-0.105 8	0.002 2
13	7	6	-0.077 0	0.035 5
13	7	7	0.750 7	1.178 5
13	8	1	-0.123 7	-0.042 5
13	8	2	-0.126 9	-0.033 3
13	8	3	-0.122 0	-0.020 0
13	8	4	-0.111 3	-0.003 2
13	8	5	-0.095 2	0.016 9
13	8	6	-0.073 3	0.040 7
13	8	7	-0.044 7	0.069 0
13	8	8	0.697 2	0.972 4
13	9	1	-0.105 1	-0.019 5
13	9	2	-0.108 9	-0.011 2
13	9	3	-0.106 1	-0.000 3
13	9	4	-0.098 4	0.012 7
13	9	5	-0.086 3	0.027 9
13	9	6	-0.069 3	0.045 6
13	9	7	-0.046 8	0.066 4
13	9	8	-0.017 5	0.091 0
13	9	9	0.638 5	0.787 5

表 6.6-33　简单线性无偏估计系数（威布尔分布）

n	r	s	$E(z_{r,n})$	$nk_{r,n}$	$g_{r,n}$
26	5	5	-1.668 7	4.211 8	0.759 2
	10	10	-0.798 9	10.092 1	0.900 0
	15	15	-0.216 4	17.029 2	0.939 8
	20	20	0.310 3	25.990 2	0.959 3
	25	24	1.013 0	37.302 2	0.967 5
	26	24	1.299 3	37.776 7	0.968 7
30	5	5	-1.823 7	4.181 0	0.761 2
	10	10	-0.974 6	9.912 8	0.899 2
	15	15	-0.425 3	16.444 5	0.938 2
	20	20	0.036 4	24.280 0	0.957 1
	25	25	0.512 3	34.778 4	0.967 9
	30	27	0.744 4	41.758 9	0.972 9
35	5	5	-1.988 7	4.153 2	0.759 1
	10	10	-1.157 4	9.757 9	0.897 6
	15	15	-0.634 1	15.974 0	0.936 5
	20	20	-0.214 7	23.080 1	0.956 1
	25	25	0.174 5	31.640 0	0.966 6
	30	30	0.600 5	43.179 3	0.973 7
	35	32	0.815 2	50.730 8	0.977 2
40	5	5	-2.130 2	4.132 9	0.756 5
	10	10	-1.311 5	9.648 1	0.896 4
	15	15	-0.805 2	15.657 4	0.935 7
	20	20	-0.410 6	22.337 0	0.954 6
40	25	25	-0.062 1	29.993 6	0.965 6
	30	30	0.280 5	39.249 5	0.973 0
	35	35	0.671 2	51.792 3	0.977 6
	40	36	0.766 1	56.492 3	0.979 8
45	5	5	-2.254 1	4.117 3	0.757 5
	10	10	-1.444 7	9.566 2	0.896 2
	15	15	-0.950 5	15.429 2	0.935 0
	20	20	-0.572 3	21.827 7	0.953 9
	25	25	-0.247 8	28.953 6	0.965 2
	30	30	0.055 6	37.141 1	0.972 2
	35	35	0.365 5	47.069 5	0.977 3
	40	40	0.729 8	60.584 9	0.980 7
	45	41	0.819 5	65.498 2	0.982 4
50	5	5	-2.364 3	4.105 0	0.755 7
	10	10	-1.562 1	9.502 7	0.895 2
	15	15	-1.076 9	15.256 6	0.934 7
	20	20	-0.710 5	21.455 3	0.953 1
	25	25	-0.401 8	28.230 0	0.964 2
	30	30	-0.121 6	35.787 8	0.971 5
	35	35	0.150 3	44.489 7	0.976 3
	40	40	0.436 0	55.071 0	0.980 1
	45	45	0.779 4	69.532 6	0.982 6
	50	45	0.779 4	71.236 1	0.983 8

表 6.6-34 最佳线性无偏估计和置信下限系数（威布尔分布）

n	r	$E(z_{r,n})$	$g_{r,n}$	$B_{r,n}$	$V_{0.90\gamma}$		$V_{0.95\gamma}$		$V_{0.99\gamma}$	
					0.90	0.95	0.90	0.95	0.90	0.95
5	1	−2.186 7								
	2	−1.070 9	0.105 0	1.011 6						
	3	−0.425 6	0.583 2	0.235 4	8.78	12.58	12.07	17.36	19.73	28.71
	4	0.106 9	0.746 2	0.038 6	6.49	8.48	8.56	11.14	13.31	17.41
	5	0.690 2	0.833 4	−0.034 0	5.48	6.73	7.06	8.68	10.75	13.23
6	1	−2.369 0								
	2	−1.275 0	0.086 7	−1.208 2						
	3	−0.662 7	0.567 9	0.333 2	8.24	11.74	11.53	16.66	19.28	28.02
	4	−0.188 4	0.730 3	0.102 0	6.33	8.18	8.47	10.95	13.49	17.54
	5	0.254 5	0.813 9	0.010 5	5.42	6.73	7.08	8.82	10.96	13.67
	6	0.777 3	0.868 0	−0.031 4	4.86	5.83	6.27	7.53	9.49	11.41
7	1	−2.523 1								
	2	−1.444 1	0.073 9	1.374 6						
	3	−0.852 5	0.557 4	0.416 7	7.80	11.12	11.20	16.07	19.27	27.76
	4	−0.409 7	0.719 8	0.157 0	6.16	7.89	8.39	10.80	13.53	17.47
	5	−0.022 4	0.822 4	0.050 4	5.36	6.68	7.12	8.84	11.20	13.91
	6	0.365 3	0.853 7	−0.001 5	4.86	5.82	6.33	7.61	9.75	11.74
	7	0.846 0	0.890 4	−0.028 6	4.46	5.25	5.76	6.73	8.75	10.20
8	1	−2.656 7								
	2	−1.588 4	0.064 4	1.518 6						
	3	−1.011 1	0.549 7	0.489 2	7.51	10.67	11.02	15.76	19.24	27.28
	4	−0.588 2	0.712 4	0.205 1	5.96	7.79	8.19	10.74	13.42	17.64
	5	−0.231 2	0.791 6	0.085 9	5.28	6.50	7.07	8.78	11.12	13.92
	6	0.102 9	0.845 0	0.026 0	4.83	5.83	6.35	7.67	9.82	11.87
	7	0.452 8	0.879 8	−0.007 2	4.49	5.31	5.83	6.91	8.92	10.52
	8	0.902 1	0.907 1	−0.026 1	4.21	4.90	5.44	6.29	8.27	9.52
9	1	−2.774 4								
	2	−1.714 4	0.057 0	1.645 3						
	3	−1.147 5	0.543 9	0.553 3	7.14	10.21	10.71	15.33	19.00	27.91
	4	−0.738 3	0.706 8	0.247 8	5.77	7.39	8.02	10.40	13.28	17.57
	5	−0.400 5	0.788 9	0.117 7	5.13	6.34	6.90	8.59	10.98	13.78
	6	−0.095 8	0.838 8	0.050 8	4.74	5.67	6.27	7.51	9.82	11.78
	7	0.202 7	0.872 9	0.012 8	4.48	5.28	5.86	6.91	9.06	10.66
	8	0.524 4	0.898 2	−0.010 0	4.26	4.95	5.53	6.39	8.44	9.75
	9	0.949 3	0.919 1	−0.023 9	4.04	4.66	5.32	6.00	7.90	9.04
10	1	−2.879 8								
	2	−1.826 2	0.051 1	1.758 5						
	3	−1.267 2	0.539 3	0.610 5	6.75	9.36	10.24	14.50	18.61	27.05
	4	−0.868 1	0.702 5	0.286 1	5.56	7.17	7.81	10.12	13.12	17.17
	5	−0.543 6	0.784 5	0.146 3	5.00	6.13	6.87	8.39	11.05	13.70
	6	−0.257 4	0.834 2	0.073 4	4.67	5.59	6.24	7.50	9.81	11.86
	7	0.012 0	0.867 9	0.031 3	4.41	5.18	5.79	6.83	8.99	10.66
	8	0.283 7	0.892 5	0.005 3	4.22	4.91	5.52	6.40	8.49	9.88
	9	0.584 6	0.911 9	−0.011 4	4.03	4.63	5.23	6.01	7.97	9.17
	10	0.989 9	0.928 4	−0.022 0	3.86	4.41	4.98	5.67	7.57	8.57
11	1	−2.975 1								
	2	−1.926 7	0.046 4	1.860 6						
	3	−1.373 9	0.535 5	0.662 2	6.41	9.11	9.89	14.11	18.19	26.45
	4	−0.982 5	0.699 0	0.320 8	5.46	7.04	7.71	10.03	13.02	17.25
	5	−0.667 8	0.781 0	0.172 2	4.90	6.07	6.72	8.34	10.99	13.66
	6	−0.394 6	0.830 6	0.094 1	4.58	5.52	6.16	7.42	9.79	11.86
	7	−0.143 2	0.864 1	0.048 2	4.36	4.96	5.79	6.83	9.03	10.72
	8	0.100 7	0.888 3	0.019 5	4.15	4.87	5.46	6.38	8.43	9.86
	9	0.352 3	0.907 1	0.000 6	4.01	4.63	5.23	6.04	8.02	9.24
	10	0.636 2	0.922 3	−0.012 0	3.87	4.44	5.03	5.75	7.67	8.73

续表 6.6-34

n	r	$E\ (z_{r,n})$	$g_{r,n}$	$B_{r,n}$	$V_{0.90\gamma}$		$V_{0.95\gamma}$		$V_{0.99\gamma}$	
					0.90	0.95	0.90	0.95	0.90	0.95
11	11	1.025 2	0.935 8	-0.020 3	3.76	4.26	4.85	5.49	7.36	8.31
12	1	-3.062 1								
	2	-2.018 0	0.042 5	1.953 6						
	3	-1.470 3	0.532 4	0.709 4	6.00	8.40	9.41	13.40	17.59	25.73
	4	-1.084 9	0.696 1	0.352 4	5.17	6.60	7.42	9.56	12.72	16.60
	5	-0.777 7	0.778 2	0.196 0	4.72	5.79	6.54	8.08	10.68	13.37
	6	-0.514 0	0.827 7	0.113 0	4.41	5.31	5.97	7.22	9.61	11.60
	7	-0.275 2	0.861 0	0.063 9	4.21	4.98	5.63	6.66	8.85	10.55
	8	-0.048 9	0.885 1	0.032 7	4.06	4.75	5.36	6.27	8.34	9.75
	9	0.175 6	0.903 5	0.011 9	3.94	4.53	5.16	5.95	7.98	9.18
	10	0.411 2	0.918 2	-0.002 3	3.87	4.37	4.99	5.67	7.66	8.68
	11	0.681 2	0.930 6	-0.012 1	3.72	4.23	4.84	5.47	7.37	8.31
	12	1.056 5	0.941 8	-0.018 9	3.62	4.07	4.68	5.26	7.09	7.96
13	1	-3.142 2								
	2	-2.101 6	0.039 1	2.039 0						
	3	-1.558 1	0.529 9	0.752 7	5.88	8.16	9.23	13.11	17.58	25.37
	4	-1.177 6	0.693 7	0.381 5	5.10	6.45	7.38	9.47	12.81	16.56
	5	-0.876 3	0.775 9	0.217 9	4.71	5.75	6.57	8.04	10.89	13.47
	6	-0.619 9	0.825 4	0.130 5	4.43	5.30	6.03	7.24	9.71	11.72
	7	-0.390 4	0.858 6	0.078 4	4.23	4.96	5.65	6.68	8.97	10.61
	8	-0.176 4	0.882 5	0.044 9	4.06	4.73	5.40	6.29	8.46	9.88
	9	0.030 8	0.900 7	0.022 5	3.94	4.55	5.17	6.00	8.02	9.27
	10	0.239 9	0.915 2	0.006 9	3.83	4.37	5.01	5.70	7.72	8.80
	11	0.462 6	0.927 1	-0.004 2	3.74	4.23	4.87	5.50	7.46	8.42
	12	0.720 9	0.936 3	-0.012 1	3.65	4.09	4.73	5.30	7.21	8.03
	13	1.084 5	0.946 8	-0.017 6	3.57	3.97	4.61	5.12	6.99	7.75
14	1	-3.215 3								
	2	-2.178 8	0.036 3	2.117 9						
	3	-1.638 7	0.527 7	0.792 7	5.56	7.63	8.84	12.73	17.23	25.09
	4	-1.262 6	0.691 7	0.408 4	4.93	6.17	7.18	9.10	12.48	16.19
	5	-0.965 9	0.773 9	0.238 2	4.58	5.54	6.38	7.82	10.64	13.22
	6	-0.715 2	0.823 7	0.146 7	4.33	5.12	5.91	7.07	9.59	11.55
	7	-0.492 8	0.856 5	0.091 8	4.15	4.82	5.58	6.53	8.88	10.47
	8	-0.287 9	0.880 4	0.056 4	4.03	4.61	5.35	6.16	8.40	9.71
	9	-0.092 8	0.898 5	0.032 4	3.90	4.45	5.14	5.88	8.00	9.17
	10	0.099 4	0.912 8	0.015 6	3.78	4.30	4.98	5.65	7.71	8.72
	11	0.296 1	0.924 4	0.003 4	3.71	4.20	4.86	5.48	7.46	8.42
	12	0.508 0	0.934 3	-0.005 4	3.64	4.09	4.73	5.31	7.25	8.11
	13	0.756 4	0.942 9	-0.011 9	3.55	3.98	4.61	5.14	7.03	7.82
	14	1.109 7	0.951 1	-0.016 5	3.46	3.85	4.48	4.97	9.81	7.49
15	1	-3.285 3								
	2	-2.250 4	0.033 8	2.191 2						
	3	-1.713 3	0.525 8	0.829 9	5.39	7.23	8.75	12.22	17.15	24.54
	4	-1.340 4	0.689 9	0.433 4	4.78	5.95	7.00	8.90	12.41	15.96
	5	-1.478	0.772 2	0.257 0	4.43	5.36	6.25	7.64	10.56	12.94
	6	-0.801 9	0.821 7	0.161 8	4.22	4.97	5.79	6.91	9.45	11.38
	7	-0.585 2	0.854 8	0.104 4	4.08	4.72	5.50	6.41	8.81	10.31
	8	-0.387 3	0.878 6	0.067 1	3.95	4.57	5.29	6.10	8.36	9.64
	9	-0.201 0	0.896 6	0.041 7	3.85	4.40	5.11	5.81	8.00	9.11
	10	-0.020 6	0.910 6	0.023 7	3.76	4.26	4.96	5.61	7.71	8.72
	11	0.159 5	0.922 3	0.010 7	3.69	4.15	4.84	5.43	7.48	8.37
	12	0.345 8	0.931 9	0.001 0	3.62	4.08	4.73	5.31	7.29	8.17
	13	0.548 5	0.940 2	-0.006 2	3.55	3.98	4.63	5.17	7.09	7.89
	14	0.788 4	0.947 6	-0.011 6	3.49	3.89	4.53	5.02	6.92	7.63
	15	1.132 7	0.954 7	-0.015 6	3.41	3.77	4.43	4.88	6.70	7.37

表 6.6-35　R (t) 的 90% 下置信限（威布尔分布）

$\hat{R}$ \ n	8	10	12	15	18	20	25	30	40	50
0.60	0.397	0.419	0.435	0.455	0.469	0.477	0.490	0.500	0.515	0.524
0.62	0.414	0.437	0.453	0.473	0.488	0.496	0.510	0.520	0.535	0.544
0.64	0.432	0.455	0.472	0.492	0.507	0.516	0.529	0.540	0.555	0.564
0.66	0.450	0.474	0.491	0.512	0.526	0.535	0.549	0.560	0.575	0.584
0.68	0.468	0.493	0.511	0.532	0.546	0.555	0.569	0.580	0.596	0.605
0.70	0.486	0.512	0.530	0.552	0.566	0.575	0.589	0.601	0.616	0.626
0.72	0.504	0.532	0.550	0.573	0.586	0.596	0.610	0.622	0.537	0.646
0.74	0.524	0.552	0.571	0.593	0.607	0.617	0.631	0.643	0.658	0.668
0.76	0.544	0.573	0.592	0.615	0.628	0.638	0.653	0.665	0.680	0.690
0.78	0.566	0.595	0.613	0.637	0.651	0.660	0.675	0.687	0.702	0.711
0.80	0.588	0.618	0.635	0.660	0.674	0.683	0.698	0.709	0.724	0.733
0.82	0.611	0.641	0.659	0.683	0.697	0.706	0.721	0.732	0.746	0.756
0.84	0.636	0.666	0.683	0.707	0.722	0.730	0.745	0.755	0.769	0.778
0.86	0.662	0.692	0.709	0.732	0.747	0.755	0.769	0.780	0.793	0.802
0.88	0.689	0.719	0.736	0.759	0.773	0.781	0.795	0.805	0.818	0.825
0.90	0.719	0.748	0.765	0.787	0.800	0.808	0.821	0.831	0.843	0.851
0.92	0.751	0.780	0.796	0.817	0.829	0.837	0.849	0.859	0.869	0.876
0.94	0.787	0.815	0.831	0.849	0.861	0.867	0.879	0.887	0.897	0.903
0.96	0.829	0.855	0.870	0.887	0.896	0.901	0.911	0.918	0.926	0.931
0.98	0.885	0.906	0.917	0.930	0.937	0.941	0.948	0.953	0.959	0.962

表 6.6-36　R (t) 的 95% 下置信限（威布尔分布）

$\hat{R}$ \ n	8	10	12	15	18	20	25	30	40	50
0.60	0.347	0.372	0.393	0.416	0.432	0.443	0.459	0.473	0.490	0.500
0.62	0.363	0.389	0.411	0.434	0.450	0.462	0.478	0.493	0.510	0.519
0.64	0.380	0.406	0.428	0.452	0.469	0.480	0.497	0.521	0.530	0.539
0.66	0.396	0.424	0.445	0.471	0.488	0.499	0.517	0.532	0.550	0.559
0.68	0.414	0.443	0.464	0.490	0.507	0.519	0.536	0.552	0.570	0.580
0.70	0.432	0.461	0.483	0.510	0.527	0.538	0.557	0.573	0.591	0.601
0.72	0.450	0.481	0.502	0.530	0.547	0.559	0.577	0.594	0.612	0.622
0.74	0.469	0.500	0.523	0.550	0.568	0.580	0.598	0.616	0.633	0.644
0.76	0.489	0.520	0.544	0.572	0.590	0.602	0.620	0.638	0.654	0.666
0.78	0.509	0.542	0.567	0.594	0.612	0.625	0.643	0.661	0.676	0.688
0.80	0.529	0.564	0.590	0.617	0.636	0.648	0.666	0.683	0.700	0.711
0.82	0.552	0.587	0.614	0.641	0.660	0.672	0.689	0.706	0.724	0.734
0.84	0.576	0.611	0.638	0.667	0.685	0.697	0.714	0.730	0.748	0.758
0.86	0.602	0.638	0.664	0.693	0.710	0.723	0.740	0.755	0.772	0.783
0.88	0.629	0.666	0.692	0.721	0.737	0.750	0.767	0.781	0.798	0.808
0.90	0.661	0.696	0.722	0.751	0.766	0.780	0.795	0.809	0.824	0.834
0.92	0.695	0.729	0.755	0.782	0.798	0.811	0.825	0.838	0.853	0.862
0.94	0.735	0.767	0.792	0.817	0.832	0.845	0.858	0.869	0.882	0.890
0.96	0.782	0.812	0.835	0.857	0.872	0.882	0.893	0.903	0.915	0.921
0.98	0.844	0.869	0.890	0.907	0.918	0.926	0.935	0.943	0.950	0.955

表 6.6-37　R (t) 的 90% 下置信限（威布尔分布）（截尾试验）

$\hat{R}$	$\frac{r}{n}=0.75$				$\frac{r}{n}=0.50$			
	$n=40$	$n=60$	$n=80$	$n=100$	$n=40$	$n=60$	$n=80$	$n=100$
0.70	0.623	0.638	0.641	0.650	0.616	0.639	0.644	0.652
0.72	0.641	0.657	0.661	0.669	0.635	0.658	0.663	0.672
0.74	0.659	0.676	0.681	0.690	0.653	0.677	0.683	0.691
0.76	0.678	0.696	0.702	0.710	0.674	0.696	0.703	0.711
0.78	0.698	0.716	0.723	0.731	0.694	0.716	0.723	0.732
0.80	0.718	0.737	0.744	0.752	0.715	0.736	0.744	0.752

续表 6.6-37

$\hat{R}$	$\frac{r}{n}=0.75$				$\frac{r}{n}=0.50$			
	$n=40$	$n=60$	$n=80$	$n=100$	$n=40$	$n=60$	$n=80$	$n=100$
0.82	0.739	0.758	0.766	0.774	0.737	0.757	0.765	0.773
0.84	0.761	0.780	0.789	0.796	0.759	0.779	0.787	0.795
0.86	0.783	0.802	0.810	0.818	0.783	0.801	0.810	0.817
0.88	0.807	0.826	0.833	0.841	0.807	0.824	0.832	0.839
0.90	0.832	0.850	0.857	0.864	0.832	0.847	0.855	0.862
0.92	0.858	0.875	0.882	0.888	0.858	0.872	0.879	0.886

续表 6.6-37

$\hat{R}$	$\frac{r}{n}=0.75$				$\frac{r}{n}=0.50$			
	$n=40$	$n=60$	$n=80$	$n=100$	$n=40$	$n=60$	$n=80$	$n=100$
0.94	0.886	0.901	0.907	0.913	0.886	0.898	0.904	0.910
0.95	0.901	0.915	0.920	0.925	0.901	0.911	0.917	0.922
0.96	0.917	0.929	0.934	0.939	0.917	0.925	0.930	0.935
0.97	0.938	0.943	0.947	0.952	0.933	0.940	0.944	0.949
0.98	0.951	0.959	0.963	0.966	0.951	0.956	0.959	0.964
0.99	0.971	0.976	0.979	0.981	0.971	0.974	0.977	0.979
0.995	0.983	0.987	0.989	0.990	0.983	0.985	0.987	0.988
0.996	0.986	0.989	0.990	0.992	0.986	0.987	0.989	0.990
0.997	0.989	0.992	0.993	0.994	0.989	0.989	0.991	0.992
0.998	0.992	0.994	0.995	0.995	0.992	0.992	0.994	0.995
0.999	0.994	0.996	0.997	0.998	0.994	0.995	0.996	0.997

表 6.6-38 $R(t)$ 的 95% 下置信限（威布尔分布）（截尾试验）

$\hat{R}$	$\frac{r}{n}=0.75$				$\frac{r}{n}=0.50$			
	$n=40$	$n=60$	$n=80$	$n=100$	$n=40$	$n=60$	$n=80$	$n=100$
0.72	0.613	0.634	0.643	0.647	0.614	0.641	0.647	0.659
0.74	0.632	0.662	0.664	0.669	0.632	0.660	0.667	0.678
0.76	0.651	0.680	0.684	0.691	0.651	0.679	0.686	0.698
0.78	0.671	0.699	0.705	0.713	0.671	0.699	0.707	0.719
0.80	0.692	0.719	0.726	0.736	0.691	0.719	0.727	0.741
0.82	0.714	0.740	0.748	0.759	0.712	0.740	0.749	0.761
0.84	0.737	0.761	0.771	0.782	0.734	0.761	0.771	0.782
0.86	0.760	0.784	0.795	0.806	0.757	0.784	0.793	0.805
0.88	0.785	0.808	0.819	0.830	0.781	0.807	0.817	0.827
0.90	0.811	0.833	0.844	0.854	0.807	0.831	0.841	0.851
0.92	0.839	0.860	0.870	0.879	0.834	0.857	0.866	0.876
0.94	0.869	0.888	0.897	0.904	0.963	0.883	0.892	0.902
0.95	0.885	0.903	0.911	0.917	0.878	0.899	0.906	0.915
0.96	0.902	0.919	0.926	0.932	0.894	0.913	0.920	0.929
0.97	0.920	0.935	0.941	0.946	0.913	0.929	0.936	0.946
0.98	0.940	0.953	0.957	0.962	0.933	0.947	0.952	0.960
0.99	0.964	0.972	0.976	0.978	0.957	0.968	0.971	0.975
0.995	0.978	0.984	0.986	0.988	0.973	0.980	0.983	0.985
0.996	0.981	0.986	0.988	0.990	0.976	0.983	0.985	0.988
0.997	0.985	0.989	0.991	0.992	0.980	0.986	0.988	0.990
0.998	0.988	0.992	0.993	0.994	0.985	0.990	0.992	0.993
0.999	0.993	0.995	0.996	0.997	0.990	0.994	0.995	0.996

表 6.6-39 $R(t)$ 的 99% 下置信限（威布尔分布）（截尾试验）

$\hat{R}$	$\frac{r}{n}=0.75$				$\frac{r}{n}=0.50$			
	$n=40$	$n=60$	$n=80$	$n=100$	$n=40$	$n=60$	$n=80$	$n=100$
0.72	0.574	0.603	0.620	0.636	0.582	0.607	0.626	0.633
0.74	0.592	0.622	0.638	0.655	0.599	0.624	0.643	0.652
0.76	0.612	0.642	0.658	0.674	0.617	0.643	0.661	0.672
0.78	0.632	0.662	0.678	0.694	0.636	0.662	0.679	0.693
0.80	0.652	0.684	0.699	0.715	0.655	0.681	0.698	0.714
0.82	0.674	0.705	0.720	0.736	0.675	0.702	0.718	0.735
0.84	0.697	0.728	0.743	0.759	0.696	0.723	0.739	0.758
0.86	0.722	0.752	0.766	0.782	0.718	0.746	0.761	0.780
0.88	0.747	0.777	0.791	0.806	0.742	0.770	0.784	0.804
0.90	0.775	0.804	0.816	0.831	0.768	0.796	0.809	0.829
0.92	0.805	0.832	0.844	0.858	0.795	0.823	0.836	0.854
0.94	0.838	0.863	0.873	0.886	0.826	0.853	0.865	0.881

续表 6.6-39

$\hat{R}$	$\frac{r}{n}=0.75$				$\frac{r}{n}=0.50$			
	$n=40$	$n=60$	$n=80$	$n=100$	$n=40$	$n=60$	$n=80$	$n=100$
0.95	0.855	0.879	0.889	0.901	0.843	0.869	0.881	0.895
0.96	0.874	0.896	0.906	0.916	0.861	0.886	0.897	0.910
0.97	0.895	0.915	0.823	0.933	0.881	0.905	0.915	0.927
0.98	0.918	0.936	0.943	0.951	0.904	0.926	0.935	0.945
0.99	0.947	0.960	0.966	0.971	0.931	0.953	0.959	0.965
0.995	0.966	0.975	0.980	0.984	0.954	0.969	0.977	0.978
0.996	0.970	0.979	0.983	0.986	0.959	0.973	0.978	0.981
0.997	0.975	0.983	0.986	0.989	0.966	0.977	0.982	0.985
0.998	0.981	0.987	0.990	0.992	0.974	0.982	0.986	0.988
0.999	0.987	0.992	0.994	0.995	0.983	0.988	0.991	0.993

例 27 某金属材料的疲劳寿命服从威布尔分布，15 个试件的疲劳寿命分别为 283 00，358 00，422 00，475 00，512 00，576 00，650 00，668 00，736 00，810 00，880 00，982 00，105 000，115 500，144 500 次。估计分布参数、$N=4\times10^4$ 次的可靠度及 $R=0.09$ 时的可靠寿命。

解 用矩法估计分布参数。按式（6.6-187）~式（6.6-189）求样本均值、样本标准差和样本偏态系数

$$\overline{N}=\frac{1}{n}\sum_{i=1}^{n}N_i=\frac{1}{15}(283\,00+358\,00+\cdots+144\,500)=73\,346.67 \text{ 次}$$

$$s_N=\left[\frac{1}{n-1}\sum_{i=1}^{n}(N_i-\overline{N})^2\right]^{\frac{1}{2}}=\left\{\frac{1}{15-1}[(283\,00-73\,346.67)^2+(358\,00-73\,346.67)^2+\cdots+(144\,500-73346.67)^2]\right\}^{\frac{1}{2}}=32\,360.42$$

$$k_k=\frac{n}{(n-1)(n-2)s_N^3}\sum_{i=1}^{n}(N_i-\overline{N})^3=\frac{15}{(15-1)(15-2)\times32\,360.42^3}\times[(283\,00-73\,346.67)^3+(358\,00-73\,346.67)^3+\cdots+(144\,500-73\,346.67)^3]=0.603\,46$$

按 k_k 由表 6.6-31 查得 $\hat{k}=2.04$，$k_b=0.454\,4$，$k_a=0.886$。按式（6.6-190），尺度参数 b 的点估计

$$\hat{b}=\frac{s_N}{k_b}=\frac{32\,360.42}{0.454\,4}=71\,215.71 \text{ 次}$$

按式（6.6-191），因

$$\overline{N}-\hat{b}k_a=73\,346.67-71\,215.71\times0.886=102\,49.55<N_1=283\,00$$

故位置参数 a 的点估计

$$\hat{a}=\overline{N}-\hat{b}k_a=10\,249.55 \text{ 次}$$

按式（6.6-206），$N=4\times10^4$ 时可靠度 R 的点估计

$$\hat{R}(4\times10^4)=e^{-\left(\frac{4\times10^4-10\,249.55}{71\,215.71}\right)^{2.04}}=0.844\,9$$

按式（6.6-207），$R=0.09$ 时可靠寿命 $N(R)$ 的点估计

$$\hat{N}(0.90)=\hat{a}+\hat{b}\left(\ln\frac{1}{R}\right)^{\frac{1}{\hat{k}}}=10\,249.55+71\,215.71\left(\ln\frac{1}{0.90}\right)^{\frac{1}{2.04}}=33\,881.32 \text{ 次}$$

例 28 某产品寿命服从两参数威布尔分布。抽取 30 件做寿命试验，当失效 15 件停止试验。失效时间分别为 1 230，2 100,2 750，3 300，3 800，4 250，4 750，5 300，5 600，6 000，6 600,7 000，7 600，7 950，8 400 h。求可靠度 $R=0.90$ 时可靠寿命 $t(R)$ 的点估计。

解 首先估计分布参数。因样本大小 $n>26$，故用简单线性无偏估计。将失效时间取对数，并由小到大为 7.115，7.650，7.919，8.102，8.243，8.355，8.466，8.575，8.631，8.700，8.795，8.854，8.936，8.981，9.036。按式（6.6-

202)、式 (6.6-203)，式中系数查表 6.6-33，$n=30$，$r=15$，则 $s=15$，$E(z_{r,n})=-0.4253$，$nk_{r,n}=16.4445$，故

$$\hat{\sigma}=\frac{1}{nk_{r,n}}\left[(2s-r)x_{s,n}-\sum_{j=1}^{s}x_j+\sum_{j=s+1}^{r}x_j\right]=\frac{1}{16.4445}[(2\times15-15)\times9.036-(7.115+7.650+\cdots+9.036)+0]=0.558$$

$$\hat{\mu}=x_{s,n}-E(z_{r,n})\hat{\sigma}=9.036+0.4253\times0.558=9.273$$

按式 (6.6-204)、式 (6.6-205)，式中 $g_{r,n}$ 查表 6.6-33 得 $g_{r,n}=0.9382$，故

$$\hat{k}=\frac{g_{r,n}}{\hat{\sigma}}=\frac{0.9328}{0.558}=1.68$$

$$\hat{b}=e^{\hat{\mu}}=e^{9.273}=10646.64\ \text{h}$$

按式 (6.6-207)，式中 $\hat{a}=0$，故可靠寿命的点估计

$$\hat{t}(0.90)=\hat{b}\left(\ln\frac{1}{R}\right)^{\frac{1}{\hat{k}}}=10646.64\left(\ln\frac{1}{0.90}\right)^{\frac{1}{1.68}}=2789.16\ \text{h}$$

5 静强度概率设计方法

5.1 静强度概率设计的特点、内容与方法

传统强度设计采用确定性的材料性能参数、工作许用应力和安全系数，不考虑参数的分散性。概率设计则把设计指标及有关参数作为随机变量处理，在给定的载荷条件下，应用统计与概率分析法，设计出安全、合理的零件或结构。概率设计的基本特点是将应力和强度作为随机变量处理，应用概率和统计方法进行分析、计算，定量地评价产品的失效率和失效概率。

概率设计的基本内容及步骤为：

1) 提出设计任务，规定详细的失效概率指标；
2) 确定有关的设计变量及参数；
3) 确定零件的应力分布；
4) 确定零件强度计算公式；
5) 确定零件的强度分布；
6) 应用应力－强度干涉模型，采用合适的计算方法，计算零件的失效概率或失效概率；
7) 确定零件失效概率或失效概率的置信度区间。

5.2 应力－强度干涉模型与失效概率计算

令 $f_1(x_1)$ 为应力分布的概率密度函数，$f_s(x_s)$ 为强度分布的概率密度函数，两者一般存在干涉区，如图 6.6-27 所示。

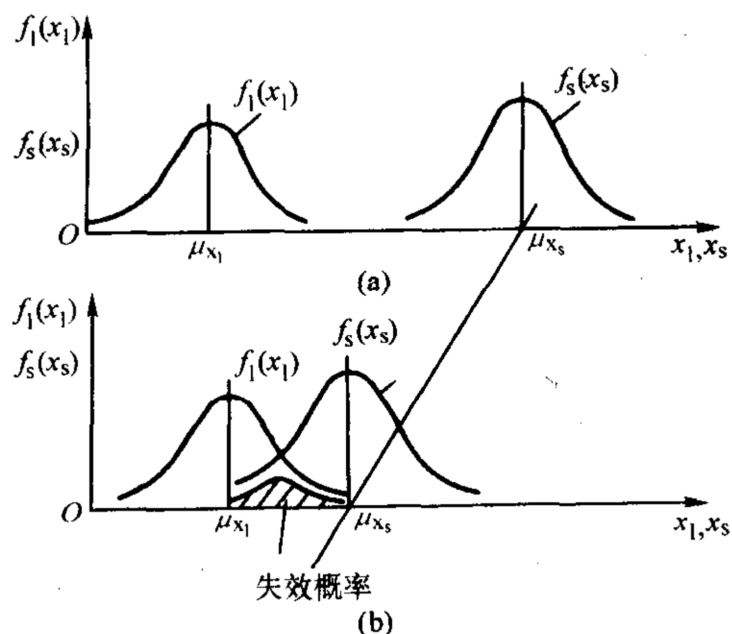


图 6.6-27　应力－强度干涉模型

$$P=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{x_1}[f_s(x_s)\mathrm{d}x_s\right]f_1(x_1)\mathrm{d}x_1 \qquad (6.6\text{-}209)$$

计算失效概率需要首先确定应力分布和强度分布。确定应力分布包括：①确定零件的失效模式及其判据；②对零件进行应力分析；③确定失效应力、应力修正系数及有关设计参数的分布；④确定可靠性计算用应力分布。

5.2.1 尺寸误差

由于每个零件的尺寸在批量生产中会各有不同，在公差范围内合格的零件，其尺寸必须作为随机变量来处理。再加上载荷也是随机变量，危险点的应力自然也是随机变量，一般假设它服从正态分布。

通常，机械加工中的容许尺寸偏差是用公差来表示的。如果现有的与尺寸的变异性有关的惟一数据是容许偏差 $\pm\Delta x$，常常可以利用它来估计标准差 σ。当预期的数据能够集中在 $\bar{x}\pm\Delta x$ 的界限内，这个界限便可用来确定一个大子样的幅度。通常，标准差的近似值为

$$s_x\approx\frac{(\bar{x}+\Delta x)-(\bar{x}-\Delta x)}{6}=\frac{\Delta x}{3} \qquad (6.6\text{-}210)$$

一般认为尺寸服从正态分布。s_x 与 Δx 的关系如图 6.6-28 所示。

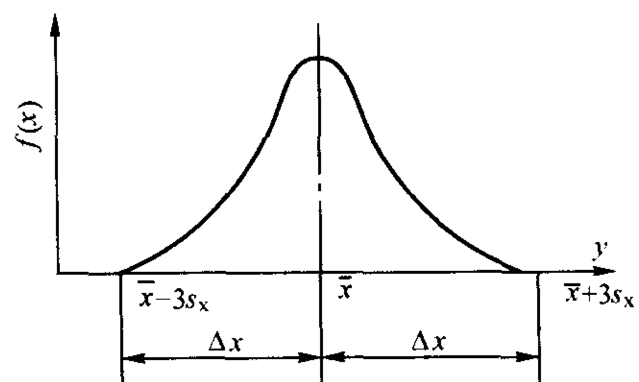


图 6.6-28　s_x 与 Δx 的关系

例 29　已知宽度为 1 000 ~ 1 200 mm 的冷轧碳钢板的名义厚度为 $t=4.60\sim4.80$ mm，容许偏差为 ±0.250 mm，问钢板厚度的标准差为多大？

解　由式 (6.6-210) 可得

$$\sigma_t\approx\frac{\Delta t}{3}=\frac{0.250}{3}=0.083\ \text{mm}$$

式 (6.6-210) 不仅可以用于确定几何尺寸的标准差，也可用于确定载荷的标准差。

显然，当误差对称于公称尺寸时，可取公称尺寸为均值 $\bar{x}$，取 $\frac{\Delta x}{3}$ 为标准差 s_x。若误差不对称于公称尺寸，可根据公称尺寸和误差先求出最大值 x_{max} 和最小值 x_{min}，然后将均值和标准差分别取为

$$\bar{x}=\frac{x_{max}+x_{min}}{2} \qquad (6.6\text{-}211)$$

$$s_x=\frac{x_{max}-x_{min}}{6} \qquad (6.6\text{-}212)$$

5.2.2 应力分布参数的代数运算

如果影响零件工作应力 s 的参数 $X_1\sim(\mu_1,\sigma_1^2)$，$X_2\sim(\mu_2,\sigma_2^2)$，…，$X_n\sim(\mu_3,\sigma_3^2)$ 均为正态随机变量，则可以根据这些参数与应力的函数关系，把它们综合为仅含单一随机变量 Z 的应力函数 $S(Z)=f(X_1,X_2,\cdots,X_n)$，并确定其分布。

如果各随机变量的变异系数都小于 0.1，即 $\sigma_i/\mu_i<0.1$，且满足随机变量的多重性要求，则由中心极限定理可知，这个应力函数近似于正态分布。正态分布随机变量的代数运算公式列于表 6.6-40。

5.2.3 强度分布

确定强度分布包括：①确定强度判据；②确定名义强度分布；③修正名义强度分布。

表 6.6-40　正态分布随机变量的代数运算公式

序号	z 的运算	均值 μ_z	标准离差 σ_z
1	$z=c$	c	0
2	$z=cx$	$c\mu_x$	$c\sigma_x$
3	$z=cx\pm d$	$c\mu_x\pm d$	$c\sigma_x$
4	$z=x+y$	$\mu_x+\mu_y$	$(\sigma_x^2+\sigma_y^2)^{\frac{1}{2}}$ 或 $(\sigma_x^2+\sigma_y^2+2\rho\sigma_x\sigma_y)^{\frac{1}{2}}$
5	$z=x-y$	$\mu_x-\mu_y$	$(\sigma_x^2+\sigma_y^2)^{\frac{1}{2}}$ 或 $(\sigma_x^2+\sigma_y^2-2\rho\sigma_x\sigma_y)^{\frac{1}{2}}$
6	$z=xy$	$\mu_x\mu_y$ 或 $\mu_x\mu_y+\rho\sigma_x\sigma_y$	$(\mu_x^2\sigma_y^2+\mu_y^2\sigma_x^2+2\rho\mu_x\mu_y\sigma_x\sigma_y)^{\frac{1}{2}}$ 或 $[(\mu_x^2\sigma_y^2+\mu_y^2\sigma_x^2+\sigma_x^2\sigma_y^2)(1+\rho^2)]^{\frac{1}{2}}$
7	$z=\dfrac{x}{y}$	μ_x/μ_y 或 $\dfrac{\mu_x}{\mu_y}+\dfrac{\mu_x\sigma_y^2}{\mu_y^3}\left(\dfrac{\sigma_y}{\mu_y}-\rho\dfrac{\sigma_x}{\mu_x}\right)$	$\dfrac{1}{\mu_y}\left(\dfrac{\mu_x^2\sigma_y^2+\mu_y^2\sigma_x^2}{\mu_y^2+\sigma_y^2}\right)^{\frac{1}{2}}$ 或 $\dfrac{1}{\mu_y{}^2}(\mu_x^2\sigma_y^2+\mu_y^2\sigma_x^2)^{\frac{1}{2}}$ 或 $\dfrac{\mu_x}{\mu_y}\left(\dfrac{\sigma_x^2}{\mu_x^2}+\dfrac{\sigma_y^2}{\mu_y^2}-2\rho\dfrac{\sigma_x\sigma_y}{\mu_x\mu_y}\right)^{\frac{1}{2}}$
8	$z=x^2$	μ_x^2 或 $\mu_x^2+\sigma_x^2$	$2\mu_x\sigma_x$ 或 $(4\mu_x^2\sigma_x^2+2\sigma_x^4)^{\frac{1}{2}}$
9	$z=x^3$	μ_x^3 或 $\mu_x^3+3\sigma_x^2\mu_x$	$3\mu_x^2\sigma_x$ 或 $(3\sigma_x^6+8\sigma_x^4\mu_x^2+5\sigma_x^2\mu_x^4)$
10	$z=x^n$	μ_x^n	$n\mu_x^{n-1}\sigma_x$
11	$z=x^{\frac{1}{2}}$	$\left(\dfrac{1}{2}\sqrt{4\mu_x^2-2\sigma_x^2}\right)^{\frac{1}{2}}$	$\left(\mu_x-\dfrac{1}{2}\sqrt{4\mu_x^2-2\sigma_x^2}\right)^{\frac{1}{2}}$
12	$z=(x^2+y^2)^{\frac{1}{2}}$	$(\mu_x^2+\mu_y^2)^{\frac{1}{2}}$	$\left(\dfrac{\mu_x^2\sigma_y^2+\mu_y^2\sigma_x^2}{\mu_x^2+\mu_y^2}\right)^{\frac{1}{2}}$

需要注意的是，材料性能手册给出的材料数据，多是用标准试样得到的材料的名义强度。如果将名义强度转化成零件的实际强度，必须进行修正。由此，引入了应力集中系数、尺寸系数和表面系数，这些系数本身也是随机变量，一般服从正态分布。

在材料性能手册中，可以查到各种材料在不同载荷类型作用下的应力集中系数。

5.2.4　失效概率计算方法

例 30　钢制拉杆，工作应力 $x_1=\sigma\sim N\ (400,\ 25^2)$ MPa，屈服强度 $x_s=\sigma_s\sim N\ (500,\ 50^2)$ MPa，求不发生屈服失效的失效概率。

解　按可靠性联结方程

$$z_R=\frac{\overline{x}_s-\overline{x}_1}{(s_1^2+s_s^2)^{\frac{1}{2}}}=\frac{500-400}{(25^2+50^2)^{\frac{1}{2}}}=1.789$$

查附表 1 或计算可得

$$P=1-R=\Phi\ (z_R)\ =1-\Phi\ (1.789)$$
$$=1-0.963\ 18=0.036\ 82$$

5.3　蒙特卡罗方法

蒙特卡罗方法是一种标准的数值方法，可以用来计算失效概率。

5.3.1　蒙特卡罗法求解失效概率的原理

蒙特卡罗法是通过随机变量的统计试验或随机模拟，求解工程技术问题近似解的方法。因此，也称为统计试验法或随机模拟法。

应用蒙特卡罗法求解失效概率，必须知道应力和强度分布函数的类型及参数。这种方法的实质是从应力分布中随机抽取一个应力值，再与从强度分布随机抽取一个强度值进行比较，如果应力大于强度，则零件失效。反之，零件安全。由此可见，每一次随机模拟相当于对一个随机抽取的零件进行一次试验。

设模拟次数为 N，失效数为 F，则零件的失效概率近似值等于 F/N，失效概率 $R=1-\dfrac{F}{N}$。模拟次数 N 越大，则模拟精度越高。要获得可靠的模拟结果，往往需要进行千次以上的模拟。由于模拟次数多，所以，用蒙特卡罗法进行模拟，一般由计算机完成。

在计算机上进行零件失效概率的模拟，其大致过程如下。

1）确定随机模拟次数 N 及计算失效概率的公式：$P\ (x_s-x_1)\ >0$。

2）输入原始资料　原始资料包括：零件工作应力 x_1 及强度 x_s 的概率密度函数，影响应力和强度的各独立随机变量的参数等。

3）给出各独立随机变量的分布规律。

4）产生符合上述给定分布规律的随机数组，每个随机数组中各有 N 个随机数。

5）计算零件的工作应力　在影响零件工作应力的随机数组中，抽取一组随机数代入应力公式，计算零件的工作应力 x_1。

6）计算零件的强度　在影响零件强度的随机数组中，抽取一组随机数代入强度公式，计算零件的强度值 x_s。

7）比较应力及强度，若 $x_1-x_s>$，则零件失效；反之，零件安全。如此重复进行 N 次，可以得到 N 次模拟时零件的失效数 F。

8）计算零件失效概率的模拟值 $R=1-\frac{F}{N}$ 模拟零件失效概率的计算框图如图 6.6-29。

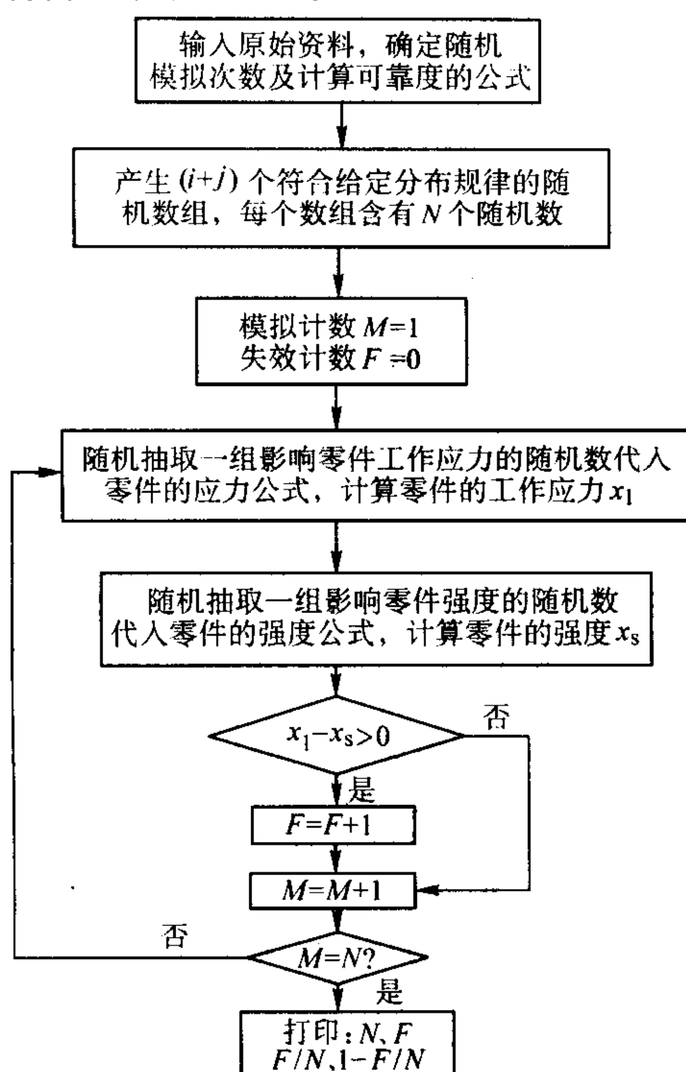


图 6.6-29　蒙特卡罗法失效概率的计算框图

从零件失效概率随机模拟的过程可以看出，蒙特卡罗法的主要优点是可以确定复杂随机变量的概率分布和数字特征。此外，对于用干涉理论解析法难以处理的分布，也可用蒙特卡罗法求解。

5.3.2　随机数的产生方法

蒙特卡罗法中重要的一步是生成随机数。产生随机变量的基础是产生 [0，1] 区间上均匀分布的随机数（即随机数发生器）。其他各类分布如正态分布、对数正态分布、泊松分布等，都是通过 [0，1] 区间上的均匀分布进行变换得到的，所以随机数的产生至关重要。

目前计算机仿真中采用的随机数发生器都是按一定的数学函数的递推算法得到的，因此所产生的数严格讲，不是概率意义下真正的随机数，被称为伪随机数。但如果算法选择合适，所得到的伪随机数序列周期足够长，而且能满足均匀性、独立性要求，将这种伪随机数用于仿真是可行的。产生随机数的常用方法有同余法、混合同余法、组合法等。

1）运用乘同余法产生均匀分布随机数的迭代公式为：

$$x_i = ax_{i-1}\ (\bmod m)\qquad (i=1,\ 2,\ 3,\ \cdots)\tag{6.6-213}$$

式中，x_i 为第 i 个随机数；a 为乘子，m 为模，两者均为正整数；(mod$m$) 表示除以 m 后取余数，当 x_0 给定时用上式递推可以得到一组数的序列，然后将该序列的每个值除以模 m，便可以得到 [0，1] 间上均匀分布的随机数。一般来说以 2^k 为模的乘同余随机数发生器统计性质不好，很少使用。为了提高乘同余法的可用性，较为理想的方法是，m 取素数（只能被 1 和自身整除的数）。a 选取条件是：选择 a 使 $a^{m-1}-1$ 为能被 m 整除的最小整数，此时得到的随机数具有最大周期。

2）混合同余法产生随机数的迭代公式为：

$$x_i = (ax_{i-1}+c)\ (\bmod m)\qquad (i=1,\ 2,\ 3,\ \cdots)$$

$$r_i = \frac{x_i}{m}\tag{6.6-214}$$

式中，x_i 为第 i 个随机数；a 为乘子；c 为增量；m 为模。x_0 称为随机数源或种子。如果 m 足够大，且发生器具有满周期，可以用 r_i 代替均匀分布的随机变量。加大 m 可以提高伪随机数的均匀性。在实际应用中，一般取 $m=2^k$，c 为奇数，$a-1$ 能被 4 整除，将得到满周期随机数序列，其中 x_0 可以去任意整数，如果需要多次运行模拟程序，而每次运行需采用不同的均匀随机数时，x_0 可以取当前时间值。

3）组合发生器的基本思想是将两组独立的随机数序列组合起来，形成新的随机数序列以提高线性同余发生器的性能、减少随机数序列的自相关，一般来说两个性能不太好的随机数发生器组合后能得到统计特性较好的随机数序列。乘同余法和混合同余法统称为线性同余法，方法简单，程序容易实现，因而应用广泛。

5.3.3　伪随机数序列的检验

由于数学方法产生的随机数有时并非均匀分布，而且序列间具有一定的相关性，在使用前应进行随机数发生器的均匀性和独立性检验。

1）均匀性检验　最常用的均匀性检验方法是频率检验法，即将随机数的取值范围 [0，1] 分成 K 个互不重叠的等长的子区间，计算由随机数发生器产生的 N 个随机数落在一个子区间的概率 $P=n_j/N$（n_j 为落在第 j 个子区间的随机数的个数），均匀分布的理论概率为 $P=1/K$，实际频率与理论频率的偏差为：

$$\chi^2 = \sum_{j=1}^{K}\left(\frac{n_j}{N}-\frac{1}{K}\right)^2\tag{6.6-215}$$

χ^2 的大小反映了随机数的均匀程度。当 K 足够大时，χ^2 近似服从自由度为 $K-1$ 的 χ^2 分布。因此可以用第 4.1.1 节中的 χ^2 检验法进行随机数的均匀性检验。

2）独立性检验　独立性检验的常用的方法是计算序列相邻一定间隔的随机数之间的相关系数，判断其相关程度。对于由随机数发生器产生的 N 个随机数 r_i，前后相隔 j 个数的两个随机数 r_i 和 r_{i+j} 的相关系数计算公式为：

$$\hat{\rho}_j = \frac{1}{S^2}\left[\frac{1}{N-j}\times\sum_{i=1}^{N-j}r_ir_{i+j}-\bar{r}^2\right]\tag{6.6-216}$$

式中，S^2 为随机数的方差。取统计量 $\mu=\hat{\rho}\sqrt{N-j}$，当 N 充分大时，μ 渐进服从标准正态分布 N（0，1）。因此通过 μ 来检验随机数序列的独立性。

5.3.4　抽样方法

在含有随机变量的模型中，随机变量是通过随机抽样的方法进行模拟的。常用的产生给定分布的随机变量的方法有三种：反变换法、组合法、取舍法。其中组合法适合于无法得到反分布函数，且可以将分布函数分解的情况；取舍法用于概率密度函数复杂又难以求得其累计分布函数的情况。

1）反变换法　适合于可以得到反分布函数的模型，具有直观、快速的特点。因此广泛采用反变换法进行随机抽样。大量试验结果表明，常用的概率分布有正态分布、对数正态分布和威布尔分布。它们变换公式为：

标准正态分布的反变换公式：

$$\left.\begin{aligned}Z_1 &= \sqrt{-2\ln X_1}\cos\ (2\pi X_2)\\ Z_2 &= \sqrt{-2\ln X_2}\cos\ (2\pi X_1)\end{aligned}\right\}\tag{6.6-217}$$

式中，X_1、X_2 是 (0，1) 区间上两个独立的均匀分布的随机变量。计算得到的 Z_1、Z_2 也是相互独立的随机变量。为了得到服从（μ，σ^2）分布的随机变量，需进行如下变换：

$$Z_i\times\sigma^2+\mu\qquad (i=1,\ 2)\tag{6.6-218}$$

威布尔分布的反变换公式：

$$Z = b\ (-\ln X)^{\frac{1}{k}}+a\tag{6.6-219}$$

式中，α 为位置参数；b 为尺度参数；k 为形状参数；X 为

(0, 1) 区间均匀分布的随机变量。

对数正态分布的反变换公式：

$$Z=\exp(\sigma t+\mu) \tag{6.6-220}$$

式中，t 为标准正态分布的随机变量；σ 为对数标准差；μ 为对数均值。

2) 取舍法　对于任何一个分布函数 $g(y)$，$y\in(0,1)$，如果想选出一个 Y，可先求出 g_0，g_0 是 $g(y)$ 在 y (0, 1)中的最小上限，然后产生两个平均分布于 (0, 1) 间的随机数 ξ_1 和 ξ_2，若有

$$\xi_2<\frac{g(\xi_1)}{g_0}$$

则 $Y=\xi_1$

否则，重新产生两个随机数 ξ_3 和 ξ_4，重复上述试验，直到满足

$$\xi_{i+1}<\frac{g(\xi_i)}{g_0}$$

则　$Y=\xi_i$

且 Y 具有 $g(y)$ 的分布。

例 31　已知　$g(y)=\frac{4}{\pi}\times\frac{1}{1+y^2}$，$y\in(0,1)$，试依据 $g(y)$ 求出 $Y(\xi)$。

解　因为 $y\geqslant0$，所以 $\frac{4}{\pi}$ 是 $g(y)$ 在 (0, 1) 区间内的最大值。对于两个连续的 ξ，不妨设为 ξ_i 和 ξ_{i+1}，若

$$\xi_{i+1}\leqslant\frac{\frac{4}{\pi}\times\frac{1}{1+\xi_i^2}}{\frac{4}{\pi}}$$

即 $\xi_{i+1}(1+\xi_i^2)\leqslant1$，则 $Y=\xi_i$

否则，重新产生两个连续的随机数 ξ_{i+2} 和 ξ_{i+3}，重复上面的试验，求得的 Y 具有 $\frac{4}{\pi}\times\frac{1}{1+y^2}$ 的分布形式。

3) 组合法　若 $g(y)$ 为一概率分布，而 $g(y)$ 可以分解为一些分布函数之和，写为

$$g(y)=\int f_z(y)h(z)\mathrm{d}z$$

或　$$g(y)=\sum f_n(y)p(n)$$

先根据 $h(z)$ 或 $p(n)$ 找出 z 或 n，然后再根据 f_z 或 f_n 去求出 y。

例 32　应用组合法对正态分布进行抽样。

解　引入分布

$$f_1(y)=y\exp(-y^2/2),\qquad 0\leqslant y$$

$$f_2(y)=\frac{1}{\pi}\sqrt{1-x^2},\qquad -1\leqslant x\leqslant1$$

则正态分布可以表示为如下的组合分布形式：

$$f(x)=\int_{-\infty}^{\infty}\frac{1}{|y|}f_2\left(\frac{x}{y}\right)f_1(y)\mathrm{d}y$$

于是有如下正态分布的组合抽样法：

$$X_f=\sqrt{-2\ln\xi_1}\cos2\pi\xi_2$$

例 33　已知一零件的应力分布和强度分布都为正态分布，其数据为：$\bar{x}_1=94.1$ MPa，$s_1=20.7$ MPa；$\bar{x}_s=188.2$ MPa，$s_s=15.2$ MPa。试分别用联结方程及蒙特卡罗法计算其失效概率，并加以比较。

解　1) 两个正态分布函数，用下面联结方程求失效概率

$$z=\frac{\bar{x}_s-\bar{x}_1}{\sqrt{s_s^2+s_1^2}}=\frac{188.2-94.1}{\sqrt{20.7^2+15.2^2}}=3.66$$

由附表1查得失效概率 $R=0.999\ 87$。

2) 由图 6.6-29 的计算机框图编制计算机程序，并得出下列打印结果：

模拟次数	失效概率
1 000	0.999 0
5 000	0.999 0
10 000	0.999 8
50 000	0.999 78

可见两种计算方法所得的结果比较接近。随着模拟试验次数的增加，模拟结果精度也随之提高。

5.4 有多种失效模式的失效概率

前面讨论零件的失效概率时只考虑一种失效模式，因此只有一种工作能力的强度判据，但是在实际中常常存在多种失效模式的零件。例如，对于传动螺旋轴，轴承受的载荷过高，使螺杠截面发生断裂、塑性变形，同时还会丧失稳定性而发生纵向弯曲三种失效模式；螺纹牙的失效可能有弯曲、剪切和磨损三种失效模式，因此讨论多种失效模式失效概率具有现实的意义。

假设零件有两种失效模式，第一种失效的失效概率为 F_1，第二种失效的失效概率为 F_2，两种同时失效的失效概率为 F_{12}，则该零件的失效概率的一般表达式为

$$R=1-F_1-F_2+F_{12} \tag{6.6-221}$$

如果两种失效模式是独立的，则失效概率为

$$R=1-F_1-F_2+F_1F_2 \tag{6.6-222}$$

若零件具有三种失效模式，设第 i 种失效的失效概率为 F_i ($i=1,2,3$)；第 i 种失效与第 j 种失效同时发生的失效概率为 F_{ij} (i、$j=1,2,3$；$i\neq j$)，第一、二、三种同时失效的概率为 F_{123}，则失效概率为

$$R=1-F_1-F_2-F_3+F_{12}+F_{23}+F_{13}-F_{123} \tag{6.6-223}$$

5.5 可靠度的置信度和置信区间

5.5.1 置信度

置信区间估计，是估计母体的待估计参量（例如母体均值 μ）以某一给定概率落在多大的区间之内。设母体待估计的参数为 θ，给定的概率为 $1-\gamma$，区间的下限、上限分别为 $L(x)$、$U(x)$，则根据上述定义可表示为：

$$P[L(x)\leqslant\theta\leqslant U(x)]=1-\gamma \tag{6.6-224}$$

式 (6.6-224) 表示母体待估计的参量 θ 的真值落在区间 $[L(x), U(z)]$ 内的概率为 $1-\gamma$。区间 $[L(x), U(x)]$ 称为置信区间，$1-\gamma$ 称为置信度或置信水平或写成 $C=1-\gamma$。

5.5.2 单侧置信区间下限和最低可靠度

可靠度的置信区间有三种（参见图 6.6-30），即：真实的可靠度值落在双侧置信区间 $[R_{L2}, R_{U2}]$ 内的概率；真实的可靠度值小于单侧置信区间上限 R_{U1} 的概率；和真实的可靠度值大于单侧置信区间下限 R_{L1} 的概率。在可靠性设计中，常用的是第三种情况，因为就机械零件的强度和寿命而言，关心的是保证零件的最低可靠度。

图 6.6-30 所示三种情况的置信度 $C=1-\gamma$ 的表达式分别为

$$C=P(R_{L2}\leqslant R\leqslant R_{U2})=1-\gamma \tag{6.6-225}$$

$$C=P(R\leqslant R_{U1})=1-\gamma \tag{6.6-226}$$

$$C=P(R\leqslant R_{L1})=1-\gamma \tag{6.6-227}$$

式中，C 为置信度；R_{L2} 为可靠度的双侧置信区间的下限；R_{U2} 为可靠度的双侧置信区间的上限；R_{U1} 为可靠度的单侧置信区间的上限；R_{L1} 为可靠度的单侧置信区间的下限。

下面只介绍置信区间下限的可靠度。

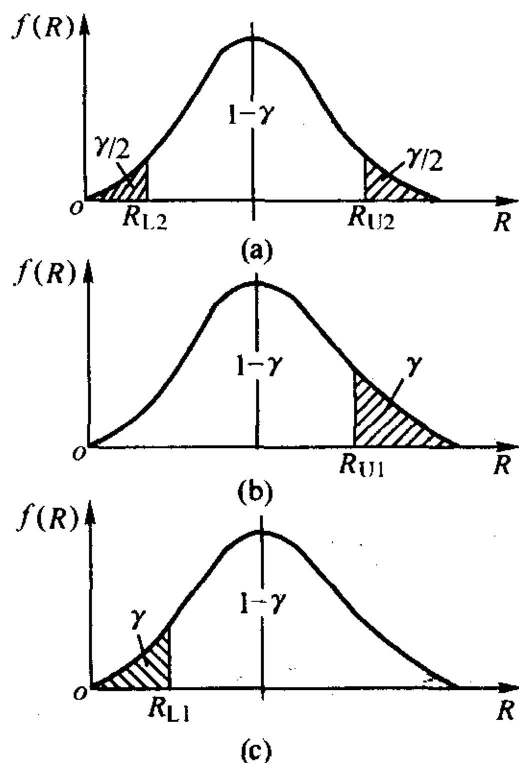


图 6.6-30 可靠度的置信区间

如前所述，可靠度是 $\delta=(x_s-x_1)>0$ 的概率，可写为

$$R=P(\delta>0)=\int_{-\infty}^{\infty}f(\delta)\mathrm{d}\delta$$

如图 6.6-31 所示。

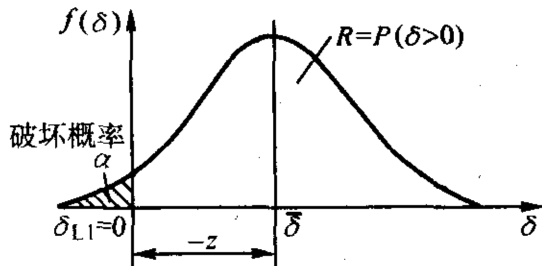


图 6.6-31 强度与应力差值（$\delta=x_s-x_1$）的分布

当 $f(\delta)$ 为正态分布，而且已知其母体分布参量 μ_δ 和 σ_δ，则随机变量 δ 分布的至少 100（$1-\alpha$）%将大于界限值 δ_{L1}，表示为

$$\delta_{L1}=\mu_\delta-k_\alpha\sigma_\delta \qquad (6.6\text{-}228)$$

但实际上很难知道母体的均值 μ_δ 及标准差 σ_δ 的真值，而只知道他们的估计值，即根据强度子样均值 $\bar{y}$、应力子样均值 $\bar{x}$，强度子样标准差 s_y 和应力子样标准差 s_x 计算的。为此，需要确定一个常数 k_a 使得在一系列来自正态分布母体的关于 δ 的子样中，区间 $[\mu_\delta-k_a\sigma_\delta,\ \infty]$ 能以一个固定的概率（$1-\gamma$）包含 δ 分布的 100（$1-\alpha$）%或更多。

由于可以接受的最小的 $\delta_{L1}=0$（图 6.6-31），式（6.6-228）中当 $\delta_{L1}=0$ 时，有

$$k_a=\frac{\mu_\delta}{\sigma_\delta} \qquad (6.6\text{-}229)$$

上式与图 6.6-31 相比较，有

$$k_a=\frac{\mu_\delta}{\sigma_s}=-z \qquad (6.6\text{-}230)$$

对于应力－强度干涉模型，应综合考虑应力和强度分布参量得到的有效子样容量 n_e

$$n_e=\frac{(s_s^2/n_s+s_1^2/n_1)}{\dfrac{(s_s^2/n_s)^2}{n_s-1}+\dfrac{(s_1^2/n_1)^2}{n_1-1}}+1 \qquad (6.6\text{-}231)$$

式中，s_s 为正态强度分布的子样标准差；s_1 为正态应力分布的子样标准差；n_s 为强度分布依据的子样容量；n_1 为应力分布依据的子样容量。

当置信度 C 给定时，k_a（用 z 表示）、n_e 与 R_{L1} 的关系图表，如图 6.6-32～图 6.6-39。为了图上注字的简化，以 R' 代替 R_{L1}。

例 34 已知一零件的应力分布和强度分布都是正态分布，其参数分别为应力 $\bar{x}_1=920$ MPa，$s_1=81.2$ MPa；强度 $\bar{x}_s=121\,3$ MPa，$s_s=41.2$ MPa。又知确定上述参数的子样容量分别为 $n_1=24$，$n_s=40$。求置信度分别为 50%，80%，90%和 99%时的单侧置信区间下限 R'。

解 由联结方程

$$z=\frac{\bar{x}_s-\bar{x}_1}{\sqrt{s_s^2+s_1^2}}=\frac{121\,3-920}{\sqrt{41.2^2+81.2^2}}=3.218$$

由式（6.6-231）

$$n_e=\frac{(41.2^2/40+81.2^2/24)^2}{\dfrac{(41.2^2/40)^2}{40-1}+\dfrac{(81.2^2/24)^2}{24-1}}+1=31.23$$

由图 6.6-32～图 6.6-39，得与所要求的置信度相对应的 $\bar{R}$ 及 R' 如下：

置信度 C/%	可靠度均值 $\bar{R}$	可靠度的单侧置信下限 R'
50	0.999 32	0.999 32
80	0.999 32	0.996 5
90	0.999 32	0.994 5
99	0.999 32	0.985

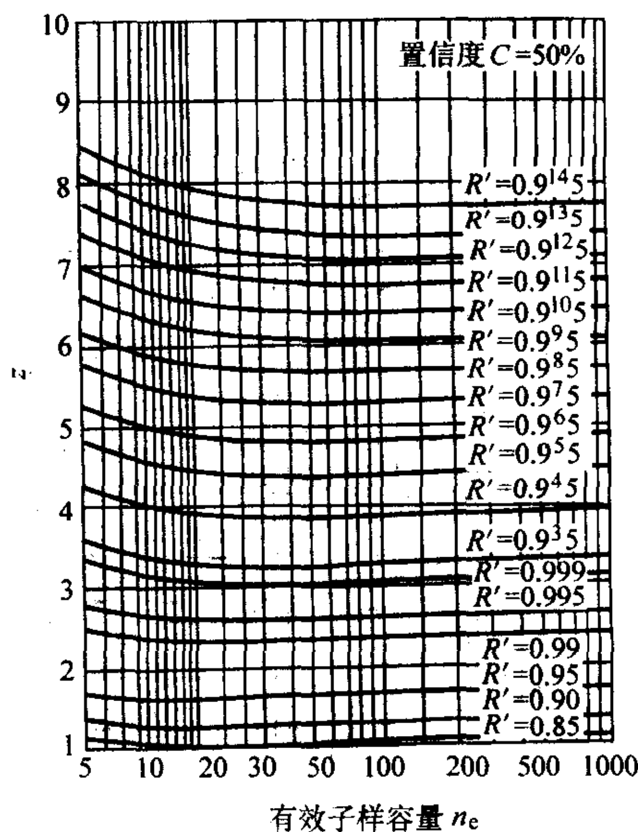


图 6.6-32 50%置信度下的可靠度

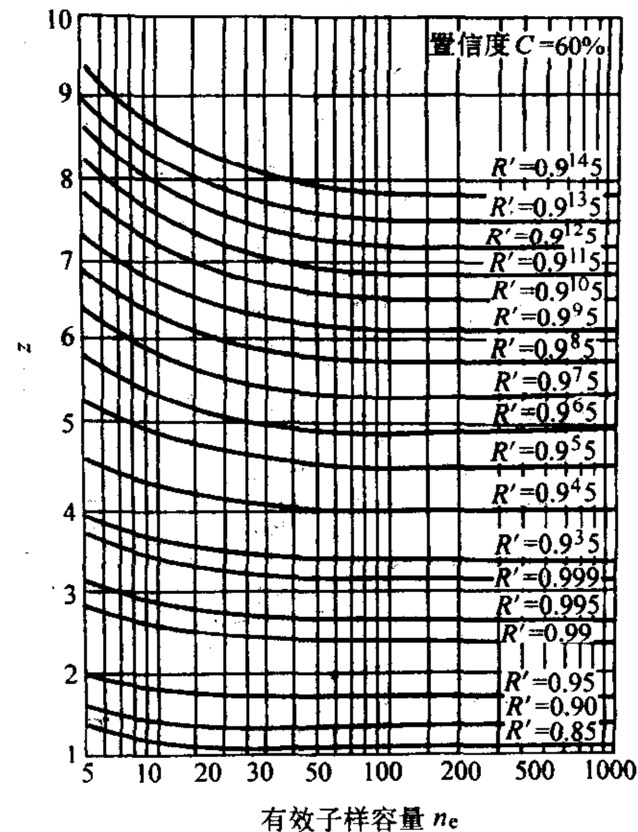


图 6.6-33 60%置信度下的可靠度

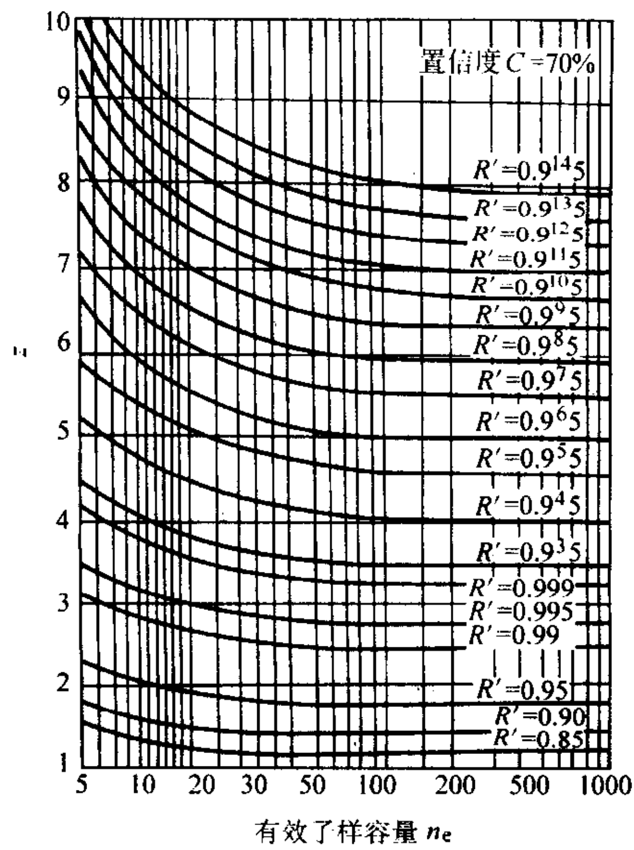


图 6.6-34 70%置信度下的可靠度

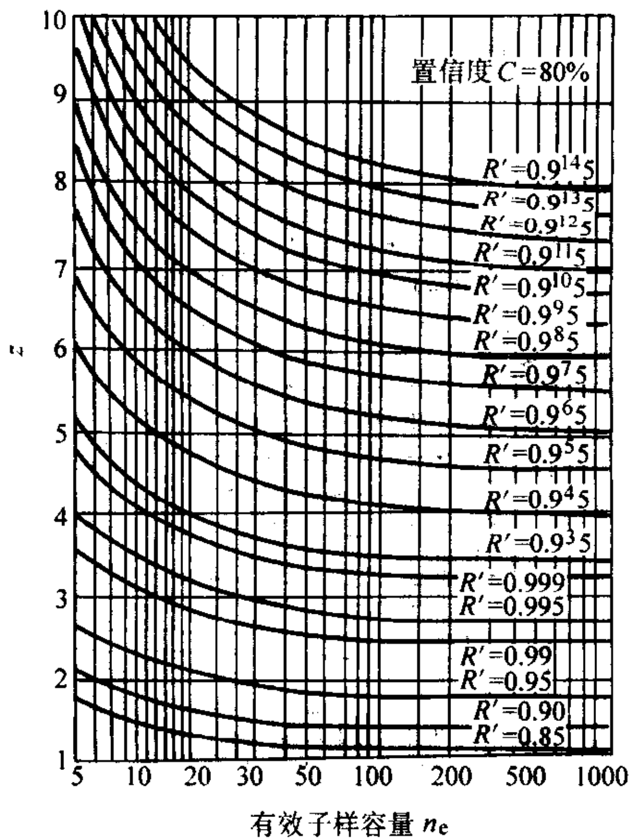


图 6.6-35 80%置信度下的可靠度

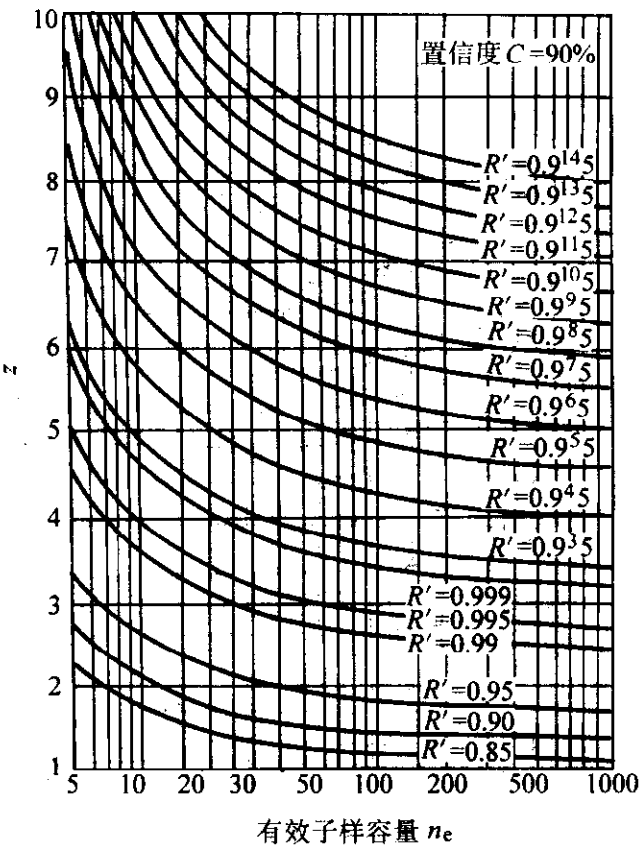


图 6.6-36 90%置信度下的可靠度

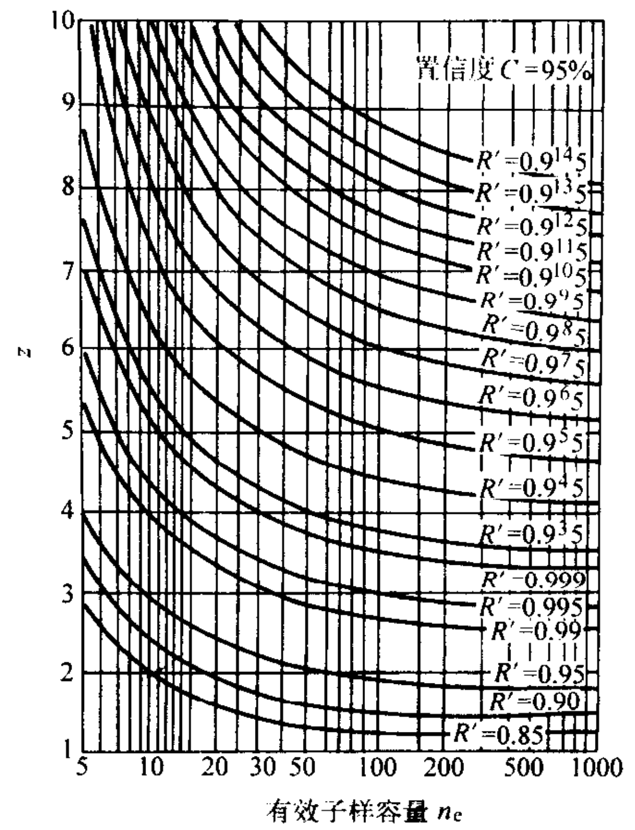


图 6.6-37 95%置信度下的可靠度

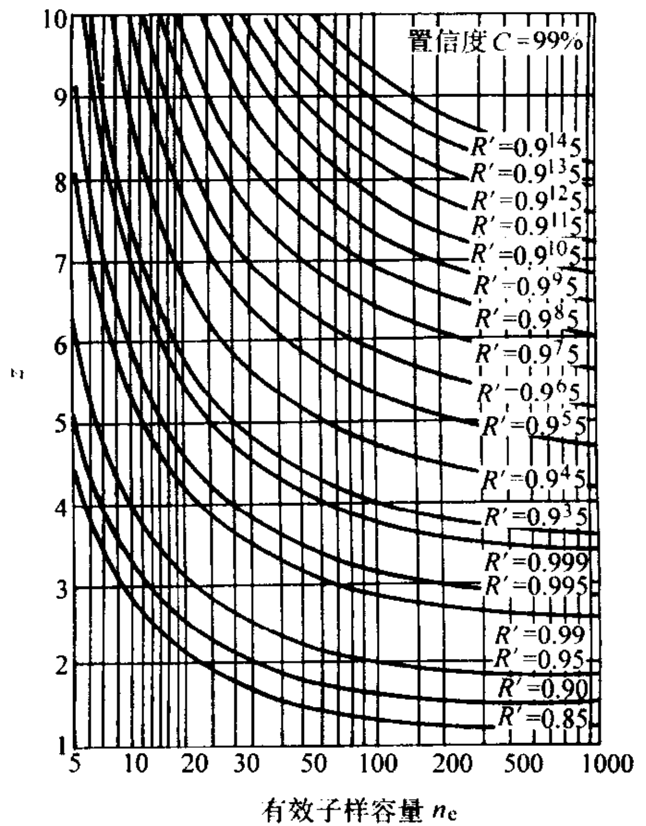


图 6.6-38 99%置信度下的可靠度

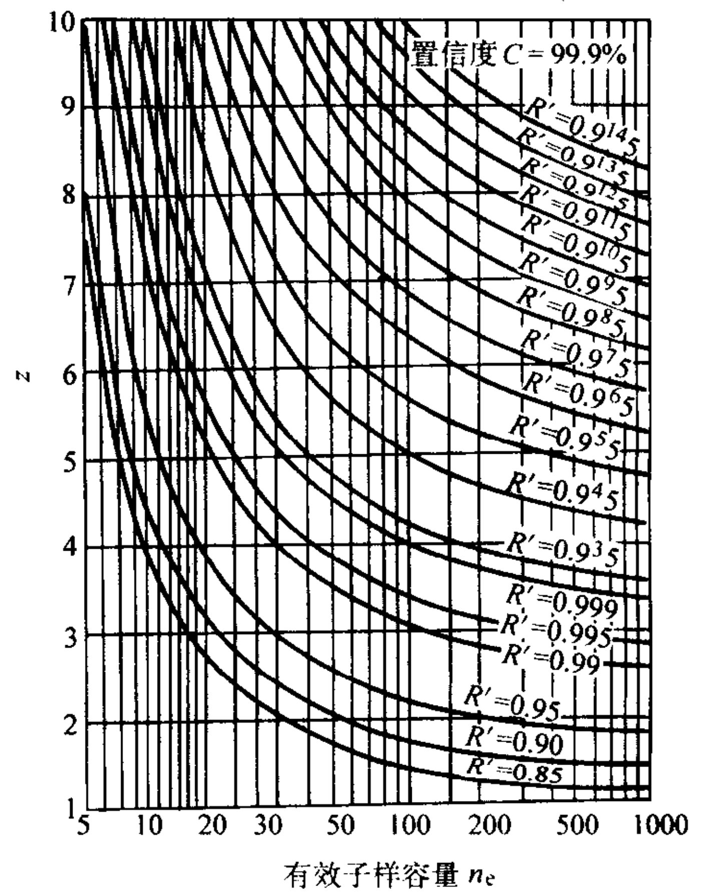
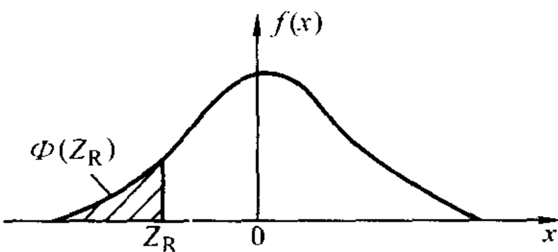


图 6.6-39 99.9%置信度下的可靠度

附表 1 标准正态分布表

$$R=\Phi(Z_R)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Z_R} e^{-\frac{x^2}{2}}dx \quad (Z_R\leqslant 0)$$

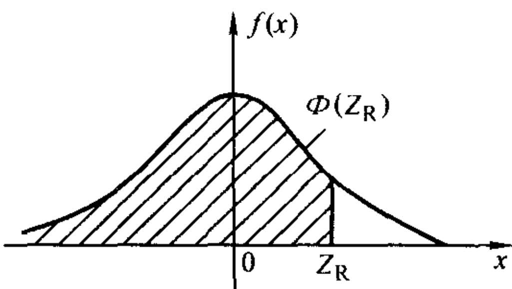


a)

Z_R	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0.0	0.500 0	0.496 0	0.492 0	0.488 0	0.484 0	0.480 1	0.476 1	0.472 1	0.468 1	0.464 1
-0.1	0.460 2	0.456 2	0.452 2	0.448 3	0.444 3	0.440 4	0.436 4	0.432 5	0.428 6	0.424 7
-0.2	0.420 7	0.416 8	0.412 9	0.409 0	0.405 2	0.401 3	0.397 4	0.393 6	0.389 7	0.385 9
-0.3	0.382 1	0.378 3	0.374 5	0.370 7	0.366 9	0.363 2	0.359 4	0.355 7	0.352 0	0.348 3
-0.4	0.344 6	0.340 9	0.337 2	0.333 6	0.330 0	0.326 4	0.322 8	0.319 2	0.315 6	0.312 1
-0.5	0.308 5	0.305 0	0.301 5	0.298 1	0.294 6	0.291 2	0.287 7	0.284 3	0.281 0	0.277 6
-0.6	0.274 3	0.270 9	0.267 6	0.264 3	0.261 1	0.257 8	0.254 6	0.251 4	0.248 3	0.245 1
-0.7	0.242 0	0.238 9	0.235 8	0.232 7	0.229 7	0.226 6	0.223 6	0.220 6	0.217 7	0.214 8
-0.8	0.211 9	0.209 0	0.206 1	0.203 3	0.200 5	0.197 7	0.194 9	0.192 2	0.189 4	0.186 7
-0.9	0.184 1	0.181 4	0.178 8	0.176 2	0.173 6	0.171 1	0.168 5	0.166 0	0.163 5	0.161 1
-1.0	0.158 7	0.156 2	0.153 9	0.151 5	0.149 2	0.146 9	0.144 6	0.142 3	0.140 1	0.137 9
-1.1	0.135 7	0.133 5	0.131 4	0.129 2	0.127 1	0.125 1	0.123 0	0.121 0	0.119 0	0.117 0
-1.2	0.115 1	0.113 1	0.111 2	0.109 3	0.107 5	0.105 6	0.103 8	0.102 0	0.100 3	0.098 53
-1.3	0.096 80	0.095 10	0.093 42	0.091 76	0.090 12	0.088 51	0.086 91	0.085 34	0.083 79	0.082 26
-1.4	0.080 76	0.079 27	0.077 80	0.076 36	0.074 93	0.073 53	0.072 15	0.070 78	0.069 44	0.068 11
-1.5	0.066 81	0.065 52	0.064 26	0.063 01	0.061 78	0.060 57	0.059 38	0.058 21	0.057 05	0.055 92
-1.6	0.054 80	0.053 70	0.052 62	0.051 55	0.050 50	0.049 47	0.048 46	0.047 46	0.046 48	0.045 51
-1.7	0.044 57	0.043 63	0.042 72	0.041 82	0.040 93	0.040 06	0.039 20	0.038 36	0.037 54	0.036 73
-1.8	0.035 93	0.035 15	0.034 38	0.033 62	0.032 88	0.032 16	0.031 44	0.030 74	0.030 05	0.029 38
-1.9	0.028 72	0.028 07	0.027 43	0.026 80	0.026 19	0.025 59	0.025 00	0.024 42	0.023 85	0.023 30
-2.0	0.022 75	0.022 22	0.021 69	0.021 18	0.020 68	0.020 18	0.019 70	0.019 23	0.018 76	0.018 31
-2.1	0.017 86	0.017 43	0.017 00	0.016 59	0.016 18	0.015 78	0.015 39	0.015 00	0.014 63	0.014 26
-2.2	0.013 90	0.013 55	0.013 21	0.012 87	0.012 55	0.012 22	0.011 91	0.011 60	0.011 30	0.011 01
-2.3	0.010 72	0.010 44	0.010 17	$0.0^{2}99\,03$	$0.0^{2}96\,42$	$0.0^{2}93\,87$	$0.0^{2}91\,37$	$0.0^{2}88\,94$	$0.0^{2}86\,56$	$0.0^{2}84\,24$
-2.4	$0.0^{2}81\,98$	$0.0^{2}79\,76$	$0.0^{2}77\,60$	$0.0^{2}75\,49$	$0.0^{2}73\,44$	$0.0^{2}71\,43$	$0.0^{2}69\,47$	$0.0^{2}67\,56$	$0.0^{2}65\,69$	$0.0^{2}63\,87$
-2.5	$0.0^{2}62\,10$	$0.0^{2}60\,37$	$0.0^{2}58\,68$	$0.0^{2}57\,03$	$0.0^{2}55\,43$	$0.0^{2}53\,86$	$0.0^{2}52\,34$	$0.0^{2}50\,85$	$0.0^{2}49\,40$	$0.0^{2}47\,99$
-2.6	$0.0^{2}46\,61$	$0.0^{2}45\,27$	$0.0^{2}43\,96$	$0.0^{2}42\,69$	$0.0^{2}41\,45$	$0.0^{2}40\,25$	$0.0^{2}39\,07$	$0.0^{2}87\,93$	$0.0^{2}86\,81$	$0.0^{2}35\,73$
-2.7	$0.0^{2}34\,67$	$0.0^{2}33\,64$	$0.0^{2}32\,64$	$0.0^{2}31\,67$	$0.0^{2}30\,72$	$0.0^{2}29\,80$	$0.0^{2}28\,90$	$0.0^{2}28\,03$	$0.0^{2}27\,18$	$0.0^{2}26\,35$
-2.8	$0.0^{2}25\,55$	$0.0^{2}24\,77$	$0.0^{2}24\,01$	$0.0^{2}23\,27$	$0.0^{2}22\,56$	$0.0^{2}21\,86$	$0.0^{2}21\,18$	$0.0^{2}20\,52$	$0.0^{2}19\,88$	$0.0^{2}19\,26$
-2.9	$0.0^{2}18\,66$	$0.0^{2}18\,07$	$0.0^{2}17\,50$	$0.0^{2}16\,95$	$0.0^{2}16\,41$	$0.0^{2}15\,89$	$0.0^{2}15\,38$	$0.0^{2}14\,89$	$0.0^{2}14\,41$	$0.0^{2}13\,95$
-3.0	$0.0^{2}13\,50$	$0.0^{2}13\,06$	$0.0^{2}12\,64$	$0.0^{2}12\,23$	$0.0^{2}11\,83$	$0.0^{2}11\,44$	$0.0^{2}11\,07$	$0.0^{2}10\,70$	$0.0^{2}10\,35$	$0.0^{2}10\,01$
-3.1	$0.0^{3}96\,76$	$0.0^{3}93\,54$	$0.0^{3}90\,43$	$0.0^{3}87\,40$	$0.0^{3}84\,47$	$0.0^{3}81\,64$	$0.0^{3}78\,88$	$0.0^{3}76\,22$	$0.0^{3}73\,64$	$0.0^{3}71\,14$
-3.2	$0.0^{3}68\,71$	$0.0^{3}66\,37$	$0.0^{3}64\,10$	$0.0^{3}61\,90$	$0.0^{3}59\,76$	$0.0^{3}57\,70$	$0.0^{3}55\,71$	$0.0^{3}53\,77$	$0.0^{3}51\,90$	$0.0^{3}50\,09$
-3.3	$0.0^{3}48\,34$	$0.0^{3}46\,65$	$0.0^{3}45\,01$	$0.0^{3}43\,42$	$0.0^{3}41\,89$	$0.0^{3}40\,41$	$0.0^{3}38\,97$	$0.0^{3}37\,58$	$0.0^{3}36\,24$	$0.0^{3}34\,95$
-3.4	$0.0^{3}33\,69$	$0.0^{3}32\,48$	$0.0^{3}31\,31$	$0.0^{3}30\,18$	$0.0^{3}29\,09$	$0.0^{3}28\,03$	$0.0^{3}27\,01$	$0.0^{3}26\,02$	$0.0^{3}25\,07$	$0.0^{3}24\,15$
-3.5	$0.0^{3}23\,26$	$0.0^{3}22\,41$	$0.0^{3}21\,58$	$0.0^{3}20\,78$	$0.0^{3}20\,01$	$0.0^{3}19\,26$	$0.0^{3}18\,54$	$0.0^{3}17\,85$	$0.0^{3}17\,18$	$0.0^{3}16\,53$
-3.6	$0.0^{3}15\,91$	$0.0^{3}15\,31$	$0.0^{3}14\,73$	$0.0^{3}14\,17$	$0.0^{3}13\,63$	$0.0^{3}13\,11$	$0.0^{3}12\,61$	$0.0^{3}12\,13$	$0.0^{3}11\,66$	$0.0^{3}11\,21$
-3.7	$0.0^{3}10\,78$	$0.0^{3}10\,36$	$0.0^{4}91\,61$	$0.0^{4}95\,74$	$0.0^{4}92\,10$	$0.0^{4}88\,42$	$0.0^{4}84\,96$	$0.0^{4}81\,62$	$0.0^{4}78\,41$	$0.0^{4}75\,32$
-3.8	$0.0^{4}72\,35$	$0.0^{4}69\,48$	$0.0^{4}66\,73$	$0.0^{4}64\,07$	$0.0^{4}61\,52$	$0.0^{4}59\,06$	$0.0^{4}56\,69$	$0.0^{4}54\,42$	$0.0^{4}52\,23$	$0.0^{4}50\,12$
-3.9	$0.0^{4}48\,10$	$0.0^{4}46\,15$	$0.0^{4}44\,27$	$0.0^{4}42\,47$	$0.0^{4}40\,74$	$0.0^{4}39\,08$	$0.0^{4}37\,47$	$0.0^{4}35\,94$	$0.0^{4}34\,46$	$0.0^{4}33\,04$
-4.0	$0.0^{4}31\,67$	$0.0^{4}30\,36$	$0.0^{4}29\,10$	$0.0^{4}27\,89$	$0.0^{4}26\,73$	$0.0^{4}25\,61$	$0.0^{4}24\,54$	$0.0^{4}23\,51$	$0.0^{4}22\,52$	$0.0^{4}21\,57$
-4.1	$0.0^{4}20\,66$	$0.0^{4}19\,78$	$0.0^{4}18\,94$	$0.0^{4}18\,14$	$0.0^{4}17\,37$	$0.0^{4}16\,62$	$0.0^{4}15\,91$	$0.0^{4}15\,23$	$0.0^{4}14\,58$	$0.0^{4}13\,95$
-4.2	$0.0^{4}13\,35$	$0.0^{4}12\,77$	$0.0^{4}12\,22$	$0.0^{4}11\,68$	$0.0^{4}11\,18$	$0.0^{4}10\,69$	$0.0^{4}10\,22$	$0.0^{4}97\,74$	$0.0^{4}93\,45$	$0.0^{4}89\,34$
-4.3	$0.0^{5}85\,40$	$0.0^{5}81\,63$	$0.0^{5}78\,01$	$0.0^{5}74\,55$	$0.0^{5}71\,24$	$0.0^{5}68\,07$	$0.0^{5}65\,03$	$0.0^{5}62\,12$	$0.0^{5}59\,34$	$0.0^{5}56\,68$
-4.4	$0.0^{5}54\,13$	$0.0^{5}51\,69$	$0.0^{5}49\,35$	$0.0^{5}47\,12$	$0.0^{5}44\,98$	$0.0^{5}42\,94$	$0.0^{5}40\,98$	$0.0^{5}39\,11$	$0.0^{5}37\,32$	$0.0^{5}35\,61$
-4.5	$0.0^{5}33\,98$	$0.0^{5}32\,41$	$0.0^{5}30\,92$	$0.0^{5}29\,49$	$0.0^{5}28\,13$	$0.0^{5}26\,82$	$0.0^{5}25\,58$	$0.0^{5}24\,39$	$0.0^{5}23\,25$	$0.0^{5}22\,16$
-4.6	$0.0^{5}21\,12$	$0.0^{5}20\,13$	$0.0^{5}19\,19$	$0.0^{5}18\,28$	$0.0^{5}17\,42$	$0.0^{5}16\,60$	$0.0^{5}15\,81$	$0.0^{5}15\,06$	$0.0^{5}14\,34$	$0.0^{5}16\,66$
-4.7	$0.0^{5}13\,01$	$0.0^{5}12\,39$	$0.0^{5}11\,79$	$0.0^{5}11\,23$	$0.0^{5}10\,69$	$0.0^{5}10\,17$	$0.0^{5}96\,80$	$0.0^{5}92\,11$	$0.0^{5}87\,65$	$0.0^{5}83\,39$
-4.8	$0.0^{6}79\,33$	$0.0^{6}75\,47$	$0.0^{6}71\,78$	$0.0^{6}68\,27$	$0.0^{6}64\,92$	$0.0^{6}61\,73$	$0.0^{6}53\,69$	$0.0^{6}55\,80$	$0.0^{6}53\,04$	$0.0^{6}50\,42$
-4.9	$0.0^{6}47\,92$	$0.0^{6}45\,54$	$0.0^{6}43\,27$	$0.0^{6}41\,11$	$0.0^{6}39\,96$	$0.0^{6}37\,11$	$0.0^{6}35\,25$	$0.0^{6}33\,48$	$0.0^{6}31\,79$	$0.0^{6}30\,19$

续附表 1

$$R = \Phi(Z_R) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Z_R} e^{-\frac{x^2}{2}}dx \quad (Z_R \geqslant 0)$$



b)

Z_R	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500 0	0.504 0	0.508 0	0.512 0	0.516 0	0.519 9	0.523 9	0.527 9	0.531 9	0.535 9
0.1	0.539 8	0.543 8	0.547 8	0.551 7	0.555 7	0.559 6	0.563 6	0.567 5	0.571 4	0.575 3
0.2	0.579 3	0.583 2	0.587 1	0.591 0	0.594 8	0.598 7	0.602 6	0.606 4	0.610 3	0.614 1
0.3	0.617 9	0.621 7	0.625 5	0.629 3	0.633 1	0.636 8	0.640 6	0.644 3	0.648 0	0.651 7
0.4	0.655 4	0.659 1	0.662 8	0.666 4	0.670 0	0.673 6	0.677 2	0.680 8	0.684 4	0.687 9
0.5	0.691 5	0.695 0	0.698 5	0.701 9	0.705 4	0.708 8	0.712 3	0.715 7	0.719 0	0.722 4
0.6	0.725 7	0.729 1	0.732 4	0.735 7	0.738 9	0.742 2	0.745 4	0.748 6	0.751 7	0.754 9
0.7	0.758 0	0.761 1	0.764 2	0.767 3	0.770 3	0.773 4	0.776 4	0.779 4	0.782 3	0.785 2
0.8	0.788 1	0.791 0	0.793 9	0.796 7	0.799 5	0.802 3	0.805 1	0.807 8	0.810 6	0.813 3
0.9	0.815 9	0.818 6	0.821 2	0.823 8	0.826 4	0.828 9	0.831 5	0.834 0	0.836 5	0.838 9
1.0	0.841 3	0.843 8	0.846 1	0.848 5	0.850 8	0.853 1	0.855 4	0.857 7	0.859 9	0.862 1
1.1	0.864 3	0.866 5	0.868 6	0.870 8	0.872 9	0.874 9	0.877 0	0.879 0	0.881 0	0.883 0
1.2	0.884 9	0.886 9	0.888 8	0.890 7	0.892 5	0.894 4	0.896 2	0.898 0	0.899 7	0.901 47
1.3	0.903 20	0.904 90	0.906 58	0.908 24	0.909 88	0.911 49	0.913 09	0.914 66	0.916 21	0.917 74
1.4	0.919 24	0.920 73	0.922 20	0.923 64	0.925 07	0.926 47	0.927 85	0.929 22	0.930 56	0.931 89
1.5	0.933 19	0.934 48	0.935 74	0.936 99	0.938 22	0.939 43	0.940 62	0.941 79	0.942 95	0.944 08
1.6	0.945 20	0.946 30	0.947 38	0.948 45	0.949 50	0.950 53	0.951 54	0.952 54	0.953 52	0.954 49
1.7	0.955 43	0.956 37	0.957 28	0.958 18	0.959 07	0.959 94	0.960 80	0.961 64	0.962 46	0.963 27
1.8	0.964 07	0.964 85	0.965 62	0.966 38	0.967 12	0.967 84	0.968 56	0.969 26	0.969 95	0.970 62
1.9	0.971 28	0.971 93	0.972 57	0.973 20	0.973 81	0.974 41	0.975 00	0.975 58	0.976 15	0.976 70
2.0	0.977 25	0.977 78	0.978 31	0.978 82	0.979 32	0.979 82	0.980 30	0.980 77	0.981 24	0.981 69
2.1	0.982 14	0.982 57	0.983 00	0.933 41	0.983 82	0.984 22	0.984 61	0.985 00	0.985 37	0.985 74
2.2	0.986 10	0.986 45	0.986 79	0.987 13	0.987 45	0.987 78	0.988 09	0.988 40	0.988 70	0.988 99
2.3	0.989 28	0.989 56	0.989 83	0.9^{2}00 97	0.9^{2}03 58	0.9^{2}06 13	0.9^{2}08 63	0.9^{2}11 06	0.9^{2}13 44	0.9^{2}15 76
2.4	0.9^{2}18 02	0.9^{2}20 24	0.9^{2}22 40	0.9^{2}24 51	0.9^{2}26 56	0.9^{2}28 57	0.9^{2}30 53	0.9^{2}32 44	0.9^{2}34 31	0.9^{2}36 13
2.5	0.9^{2}37 90	0.9^{2}39 63	0.9^{2}41 32	0.9^{2}42 97	0.9^{2}44 57	0.9^{2}46 14	0.9^{2}47 66	0.9^{2}49 15	0.9^{2}50 60	0.9^{2}52 01
2.6	0.9^{2}53 39	0.9^{2}54 73	0.9^{2}56 04	0.9^{2}57 31	0.9^{2}58 55	0.9^{2}59 75	0.9^{2}60 93	0.9^{2}62 07	0.9^{2}63 19	0.9^{2}64 27
2.7	0.9^{2}65 33	0.9^{2}66 36	0.9^{2}67 36	0.9^{2}68 33	0.9^{2}69 28	0.9^{2}70 20	0.9^{2}71 10	0.9^{2}71 97	0.9^{2}72 82	0.9^{2}73 65
2.8	0.9^{2}74 45	0.9^{2}75 23	0.9^{2}75 99	0.9^{2}76 73	0.9^{2}77 44	0.9^{2}78 14	0.9^{2}78 82	0.9^{2}79 48	0.9^{2}80 12	0.9^{2}80 74
2.9	0.9^{2}81 34	0.9^{2}81 93	0.9^{2}82 50	0.9^{2}83 05	0.9^{2}83 59	0.9^{2}84 11	0.9^{2}84 62	0.9^{2}85 11	0.9^{2}85 59	0.9^{2}86 06
3.0	0.9^{2}86 50	0.9^{2}86 94	0.9^{2}87 36	0.9^{2}87 77	0.9^{2}88 17	0.9^{2}88 56	0.9^{2}88 93	0.9^{2}89 30	0.9^{2}89 65	0.9^{2}89 99
3.1	0.9^{3}03 24	0.9^{3}06 46	0.9^{3}09 57	0.9^{3}12 60	0.9^{3}15 53	0.9^{3}18 36	0.9^{3}21 12	0.9^{3}23 78	0.9^{3}26 36	0.9^{3}28 86
3.2	0.9^{3}31 29	0.9^{3}33 63	0.9^{3}35 90	0.9^{3}38 10	0.9^{3}40 24	0.9^{3}42 30	0.9^{3}44 29	0.9^{3}46 23	0.9^{3}48 10	0.9^{3}49 91
3.3	0.9^{3}51 66	0.9^{3}53 35	0.9^{3}54 99	0.9^{3}56 58	0.9^{3}58 11	0.9^{3}59 59	0.9^{3}51 03	0.9^{3}62 42	0.9^{3}63 76	0.9^{3}65 05
3.4	0.9^{3}66 31	0.9^{3}67 52	0.9^{3}68 69	0.9^{3}69 82	0.9^{3}70 91	0.9^{3}71 97	0.9^{3}72 99	0.9^{3}73 98	0.9^{3}74 93	0.9^{3}75 85
3.5	0.9^{3}76 74	0.9^{3}77 59	0.9^{3}78 42	0.9^{3}79 22	0.9^{3}79 91	0.9^{3}80 74	0.9^{3}81 46	0.9^{3}82 15	0.9^{3}82 82	0.9^{3}83 47
3.6	0.9^{3}84 09	0.9^{3}84 69	0.9^{3}85 27	0.9^{3}85 83	0.9^{3}86 37	0.9^{3}86 89	0.9^{3}87 39	0.9^{3}87 87	0.9^{3}88 34	0.9^{3}88 79
3.7	0.9^{3}89 22	0.9^{3}89 64	0.9^{4}00 39	0.9^{4}04 26	0.9^{4}07 99	0.9^{4}11 58	0.9^{4}15 04	0.9^{4}18 38	0.9^{4}21 59	0.9^{4}24 68
3.8	0.9^{4}27 65	0.9^{4}30 52	0.9^{4}33 27	0.9^{4}35 93	0.9^{4}38 48	0.9^{4}40 94	0.9^{4}43 31	0.9^{4}45 58	0.9^{4}47 77	0.9^{4}49 88
3.9	0.9^{4}51 90	0.9^{4}53 85	0.9^{4}55 73	0.9^{4}57 53	0.9^{4}59 26	0.9^{4}60 92	0.9^{4}62 53	0.9^{4}64 06	0.9^{4}65 54	0.9^{4}66 96
4.0	0.9^{4}68 33	0.9^{4}69 64	0.9^{4}70 90	0.9^{4}72 11	0.9^{4}73 27	0.9^{4}74 39	0.9^{4}75 46	0.9^{4}76 49	0.9^{4}77 48	0.9^{4}78 43
4.1	0.9^{4}79 34	0.9^{4}80 22	0.9^{4}81 06	0.9^{4}81 86	0.9^{4}82 63	0.9^{4}83 38	0.9^{4}84 09	0.9^{4}84 77	0.9^{4}85 42	0.9^{4}86 05
4.2	0.9^{4}86 65	0.9^{4}87 23	0.9^{4}87 78	0.9^{4}88 32	0.9^{4}88 82	0.9^{4}89 31	0.9^{4}89 78	0.9^{5}02 26	0.9^{5}06 55	0.9^{5}10 66
4.3	0.9^{5}14 60	0.9^{5}18 37	0.9^{5}21 99	0.9^{5}25 45	0.9^{5}28 76	0.9^{5}31 93	0.9^{5}34 97	0.9^{5}37 88	0.9^{5}40 66	0.9^{5}43 32
4.4	0.9^{5}45 87	0.9^{5}48 31	0.9^{5}50 65	0.9^{5}52 88	0.9^{5}55 02	0.9^{5}57 06	0.9^{5}59 02	0.9^{5}60 89	0.9^{5}62 68	0.9^{5}64 39
4.5	0.9^{5}66 02	0.9^{5}67 59	0.9^{5}69 08	0.9^{5}70 51	0.9^{5}71 87	0.9^{5}73 18	0.9^{5}74 42	0.9^{5}75 61	0.9^{5}76 75	0.9^{5}77 84
4.6	0.9^{5}78 88	0.9^{5}79 87	0.9^{5}80 81	0.9^{5}81 72	0.9^{5}82 58	0.9^{5}83 40	0.9^{5}84 19	0.9^{5}84 94	0.9^{5}85 66	0.9^{5}86 34
4.7	0.9^{5}86 99	0.9^{5}87 61	0.9^{5}88 21	0.9^{5}88 77	0.9^{5}89 31	0.9^{5}89 83	0.9^{6}03 20	0.9^{6}07 89	0.9^{6}12 35	0.9^{6}16 61
4.8	0.9^{6}20 67	0.9^{6}24 53	0.9^{6}28 22	0.9^{6}31 73	0.9^{6}35 08	0.9^{6}38 27	0.9^{6}41 31	0.9^{6}44 20	0.9^{6}46 96	0.9^{6}49 58
4.9	0.9^{6}52 08	0.9^{6}54 46	0.9^{6}56 73	0.9^{6}58 89	0.9^{6}60 94	0.9^{6}62 89	0.9^{6}64 75	0.9^{6}66 52	0.9^{6}68 21	0.9^{6}69 81

注：1. $0.9^{3}0 = 0.9990$，其余类似。

2. $0.0^{3}1 = 0.0001$，其余类似。

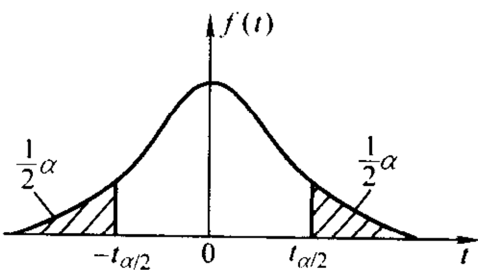
附表2　x^2 分布表

$$P\ (x^2 \geqslant x_\alpha^2)\ = \alpha$$

v \ α	0.995	0.99	0.98	0.80	0.70	0.30	0.20	0.10	0.05	0.01
1	0.000 16	0.003 93	0.015 8	0.064 2	0.148	1.074	1.642	2.706	3.841	6.635
2	0.020 1	0.102 6	0.211	0.446	0.713	2.408	3.219	4.605	5.991	9.210
3	0.115	0.352	0.584	1.005	0.424	3.665	4.642	6.251	7.815	11.341
4	0.297	0.711	1.064	1.649	2.195	4.878	5.989	7.779	9.448	13.277
5	0.554	1.145	1.610	2.343	3.000	6.064	7.289	9.236	11.070	15.086
6	0.782	1.635	2.204	3.070	3.828	7.231	8.588	10.645	12.592	16.812
7	1.239	2.167	2.833	3.822	4.671	8.383	9.803	12.017	14.067	18.475
8	1.646	2.733	3.490	4.594	5.527	9.524	11.030	13.362	15.507	20.090
9	2.088	3.325	4.168	5.380	6.393	10.656	12.242	14.684	16.919	21.666
10	2.558	3.940	4.865	6.179	7.267	11.781	13.442	15.987	18.307	23.209
11	3.053	4.575	5.578	6.980	8.148	12.899	14.631	17.275	19.675	24.725
12	3.571	5.226	6.304	7.807	9.034	14.011	15.812	18.549	21.026	26.217
13	4.107	5.892	7.042	8.634	9.926	15.119	16.985	19.812	22.362	27.688
14	4.660	6.571	7.790	9.467	10.821	16.222	18.151	21.064	23.685	29.141
15	5.229	7.261	8.547	10.307	11.721	17.322	19.311	22.307	24.996	30.578
16	5.812	7.962	9.312	11.152	12.642	18.418	20.465	23.542	26.296	32.000
17	6.408	8.672	10.085	12.002	13.531	19.511	21.615	24.769	27.587	33.409
18	7.015	9.390	10.865	12.857	14.440	20.601	22.760	25.989	28.869	34.805
19	7.633	10.117	11.651	13.716	15.352	21.689	23.900	27.204	30.144	36.191
20	8.260	10.851	12.443	14.578	16.266	22.775	25.038	28.412	31.410	37.566
21	8.897	11.591	13.240	15.445	17.182	23.858	26.171	29.615	32.671	38.932
22	9.542	12.338	14.042	16.314	18.101	24.939	27.301	30.813	33.924	40.289
23	10.196	13.091	14.848	17.187	19.021	26.018	28.428	32.007	35.172	41.638
24	10.856	13.848	15.659	18.062	19.943	27.096	29.558	33.196	36.415	42.980
25	11.524	14.611	16.473	18.940	20.867	28.172	30.675	34.382	37.6542	44.314
26	12.198	15.379	17.292	19.820	21.792	29.246	31.795	35.566	38.885	45.642
27	12.879	16.151	18.114	20.703	22.719	30.319	32.912	36.741	40.113	46.933
28	13.565	16.928	18.939	21.588	23.647	31.391	34.027	37.916	41.337	48.278
29	14.257	17.708	19.768	22.475	24.577	32.461	35.139	39.087	42.557	49.588
30	14.954	18.493	20.599	23.364	25.508	33.530	36.250	40.256	43.773	50.892
40	22.164	26.509	29.051	32.352	34.876	44.163	37.263	51.805	55.758	63.691
60	37.485	43.188	46.459	50.647	53.815	65.225	38.969	74.397	79.082	88.379
80	53.540	60.391	64.278	69.213	72.920	86.122	90.403	96.578	101.879	112.329
100	70.065	77.929	82.358	87.950	92.137	106.908	111.667	118.498	123.342	135.807
200	156.432	168.279	174.835	183.006	189.052	209.997	216.618	226.021	233.994	249.445

附表3　t 分布的双侧分位数（t）表

$P\left(|t| > t_{\frac{\alpha}{2}}\right) = \alpha$

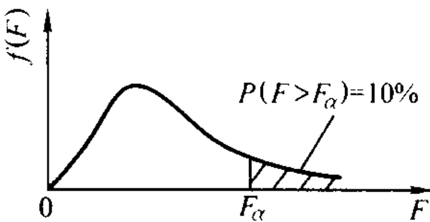


（　）中 α 值是单侧分位数（t_α）

v \ α	0.9 (0.45)	0.8 (0.4)	0.7 (0.35)	0.6 (0.3)	0.5 (0.25)	0.4 (0.2)	0.3 (0.15)	0.2 (0.1)	0.1 (0.05)	0.05 (0.12)	0.02 (0.01)	0.01 (0.005)	0.001 (0.0005)	α / v
1	0.158	0.325	0.510	0.727	1.000	1.376	1.963	3.708	6.314	12.706	31.821	63.657	636.619	1
2	0.142	0.289	0.445	0.617	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598	2
3	0.137	0.277	0.424	0.584	0.765	0.978	1.250	1.633	2.353	3.182	4.541	5.841	12.924	3
4	0.137	0.271	0.414	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610	4
5	0.132	0.267	0.408	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859	5
6	0.131	0.265	0.404	0.553	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959	6
7	0.130	0.268	0.402	0.540	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.405	7
8	0.130	0.262	0.399	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041	8
9	0.129	0.261	0.398	0.543	0.703	0.833	1.100	1.383	1.833	2.262	2.821	3.250	4.781	9
10	0.129	0.260	0.397	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587	10
11	0.129	0.260	0.396	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437	11
12	0.128	0.259	0.395	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.631	3.055	4.318	12
13	0.128	0.259	0.394	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221	13
14	0.128	0.258	0.393	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140	14
15	0.128	0.258	0.393	0.536	0.691	0.866	1.074	1.341	1.753	2.161	2.602	2.947	4.073	15
16	0.128	0.258	0.392	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015	16
17	0.128	0.257	0.392	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965	17
18	0.127	0.257	0.392	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922	18
19	0.127	0.257	0.391	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.833	19
20	0.127	0.257	0.391	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850	20
21	0.127	0.257	0.391	0.532	0.636	0.859	1.063	1.323	1.721	2.030	2.518	2.831	3.819	21
22	0.127	0.256	0.390	0.532	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792	22
23	0.127	0.256	0.390	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767	23
24	0.127	0.256	0.390	0.531	0.685	0.857	1.059	1.313	1.711	2.064	2.492	2.797	3.745	24
25	0.127	0.256	0.390	0.531	0.681	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725	25
26	0.127	0.256	0.390	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707	26
27	0.127	0.256	0.389	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690	27
28	0.127	0.256	0.389	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674	28
29	0.127	0.256	0.389	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659	29
30	0.127	0.256	0.389	0.530	0.683	0.854	1.055	0.310	1.697	2.042	2.457	2.750	3.646	30
40	0.126	0.255	0.388	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551	40
60	0.126	0.254	0.387	0.527	0.670	0.848	1.046	1.293	1.671	2.000	2.390	2.660	3.460	60
120	0.126	0.254	0.386	0.526	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373	120
∞	0.126	0.253	0.385	0.524	0.674	0.842	1.036	1.232	1.645	1.960	2.326	2.576	3.291	∞

附表 4　*F* 分布表

$P(F > F_\alpha) = \alpha$



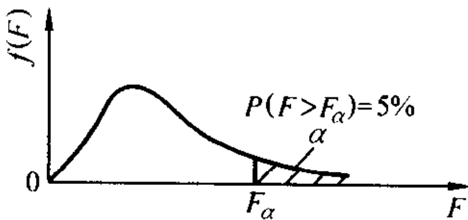
$(\alpha = 0.10)$

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	61.22	61.74	62.26	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44	9.46	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18	5.17	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84	3.82	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.24	3.21	3.17	3.14	3.12	3.13
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.87	2.84	2.80	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.63	2.59	2.56	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.46	2.42	2.38	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.34	2.30	2.25	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.24	2.20	2.16	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.17	2.12	2.08	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.10	2.06	2.01	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.05	2.01	1.96	1.90	1.88	1.85
14	3.10	2.78	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.01	1.96	1.91	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	1.97	1.92	1.87	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.94	1.89	1.84	1.78	1.75	1.72
17	3.03	2.65	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.91	1.86	1.81	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.89	1.84	1.78	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.86	1.81	1.76	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.84	1.79	1.74	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.83	1.78	1.72	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.81	1.76	1.70	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.80	1.74	1.69	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.78	1.73	1.67	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.77	1.72	1.66	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.76	1.71	1.65	1.58	1.54	1.50
27	2.90	2.21	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.75	1.70	1.64	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.74	1.69	1.63	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.73	1.68	1.62	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.72	1.67	1.61	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.66	1.61	1.54	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.76	1.71	1.60	1.54	1.48	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.55	1.48	1.41	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.49	1.42	1.34	1.24	1.17	1.00

续附表 4

$P(F > F_\alpha) = \alpha$

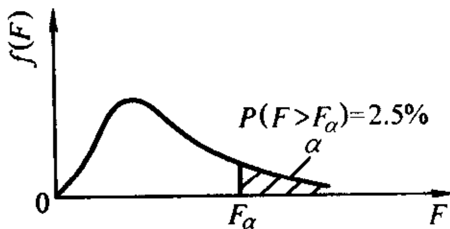
($\alpha = 0.05$)



v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	246	248	250	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66	8.62	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80	5.75	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.56	4.50	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.87	3.81	3.74	3.70	3.67
7	5.99	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.35	3.51	3.44	3.38	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.64	3.22	3.15	3.08	3.00	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.28	3.18	3.14	3.01	2.94	2.86	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.84	2.77	2.70	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.72	2.65	2.57	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.62	2.54	2.47	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.53	2.46	2.38	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.46	2.39	2.31	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.40	2.33	2.25	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.35	2.28	2.19	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.31	2.23	2.15	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.27	2.19	2.11	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.23	2.16	2.07	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.20	2.12	2.04	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.18	2.10	2.01	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.15	2.07	1.98	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.13	2.05	1.96	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.11	2.03	1.94	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.09	2.01	1.92	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.07	1.99	1.90	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.64	2.37	2.31	2.25	2.20	2.06	1.97	1.88	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.04	1.96	1.87	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.03	1.94	1.85	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.01	1.93	1.84	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	1.92	1.84	1.74	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.84	1.75	1.65	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.75	1.66	1.55	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.67	1.57	1.46	1.32	1.22	1.00

续附表 4

$P(F>F_{\alpha})=\alpha$

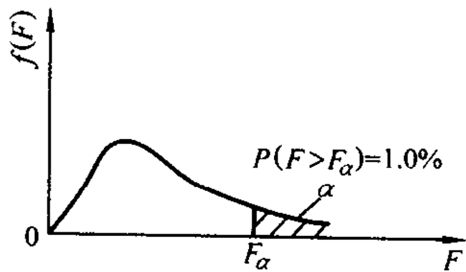


$(\alpha=0.025)$

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	∞
1	648	800	864	900	922	937	948	957	963	969	985	993	1 001	1 010	1 014	1 018
2	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5
3	17.4	16.0	15.4	15.1	14.9	14.7	14.4	14.5	14.5	14.4	14.3	14.2	14.1	14.0	13.9	13.9
4	12.2	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.66	8.56	8.46	8.36	8.31	8.26
5	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.43	6.33	6.23	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.27	5.17	5.07	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.57	4.47	4.36	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.10	4.00	3.89	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.77	3.67	3.56	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.52	3.42	3.31	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.33	3.23	3.12	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.18	3.07	2.96	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.05	2.95	2.84	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	2.95	2.84	2.73	2.61	2.55	2.49
15	6.20	4.76	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.86	2.76	2.64	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.79	2.68	2.57	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.96	2.72	2.62	2.50	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.67	2.56	2.44	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.62	2.51	2.39	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.57	2.46	2.35	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.53	2.42	2.31	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.50	2.39	2.27	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.47	2.36	2.24	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.44	2.33	2.21	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.88	2.61	2.41	2.30	2.18	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.85	2.59	2.39	2.28	2.16	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.83	2.57	2.36	2.25	2.13	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.81	2.55	2.34	2.23	2.11	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.32	2.21	2.09	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.31	2.20	2.07	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.18	2.07	1.94	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.06	1.94	1.82	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	1.94	1.82	1.69	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.83	1.71	1.57	1.39	1.27	1.00

续附表 4

$P(F>F_\alpha)=\alpha$

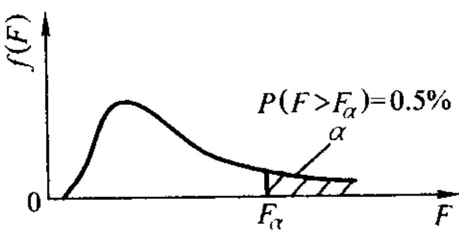


($\alpha=0.01$)

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	∞
1	4 052	5 000	5 403	5 625	5 764	5 859	5 928	5 982	6 002	6 056	6 157	6 209	6 261	6 313	6 339	6 366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	26.9	26.7	26.5	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.2	14.0	13.8	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.72	9.55	9.38	9.20	9.11	9.08
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.56	7.40	7.23	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.31	6.16	5.99	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.52	5.36	5.20	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	4.96	4.81	4.65	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.56	4.41	4.25	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.25	4.10	3.94	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.01	3.86	3.70	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.82	3.66	3.51	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.66	3.51	3.35	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.52	3.37	3.21	3.05	3.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.41	3.26	3.10	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.31	3.16	3.00	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.23	3.08	2.92	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.15	3.00	2.84	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.09	2.94	2.78	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.03	2.88	2.72	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	2.98	2.83	2.67	2.50	2.40	2.31
23	7.88	5.66	4.76	4.21	3.94	3.71	3.54	3.41	3.30	3.21	2.93	2.78	2.62	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	2.89	2.74	2.58	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.85	2.70	2.54	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	3.82	2.66	2.50	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.78	2.63	2.47	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.75	2.60	2.44	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.73	2.57	2.41	2.23	2.14	2.03
30	7.56	3.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.70	2.55	2.39	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.52	2.37	2.20	2.02	1.92	2.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.35	2.20	2.03	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	3.96	2.79	2.66	2.56	2.47	2.19	2.03	1.86	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.04	1.88	1.70	1.47	1.32	1.00

续附表 4

$P(F > F_\alpha) = \alpha$



($\alpha = 0.005$)

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	∞
1	16 211	20 000	21 615	22 500	23 056	23 437	23 715	23 925	24 091	24 224	24 630	24 836	25 044	25 253	25 359	25 465
2	198	199	199	199	199	199	199	199	199	199	199	199	199	199	199	200
3	55.6	49.8	47.5	46.2	45.4	44.8	44.4	44.1	43.9	43.7	43.1	42.8	42.5	42.2	42.0	41.8
4	31.3	26.3	24.3	23.2	22.5	22.0	21.6	21.4	21.1	21.0	20.4	20.2	19.9	19.6	19.5	19.3
5	22.8	18.3	16.5	15.6	14.9	14.5	14.2	14.0	13.8	13.6	13.1	12.9	12.7	12.4	12.3	12.1
6	18.6	14.5	12.9	12.0	11.5	11.1	10.8	10.6	10.4	10.2	9.81	9.59	9.36	9.12	9.00	8.88
7	16.2	12.4	10.9	10.0	9.52	9.16	8.89	8.68	8.51	8.38	7.97	7.75	7.53	7.31	7.19	7.08
8	14.7	11.0	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	6.81	6.61	6.40	6.18	6.06	5.95
9	13.6	10.1	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.03	5.83	5.62	5.41	5.30	5.19
10	12.8	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.47	5.27	5.07	4.86	4.75	4.64
11	12.2	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54	5.42	5.05	4.86	4.65	4.44	4.34	4.23
12	11.8	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.73	4.53	4.34	4.12	4.01	3.90
13	11.4	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94	4.82	4.46	4.27	4.07	3.87	3.76	3.65
14	11.1	7.92	6.68	6.00	5.65	5.26	5.03	4.86	4.72	4.60	4.25	4.06	3.86	3.66	3.55	3.44
15	10.8	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.07	3.88	3.69	3.48	3.37	3.26
16	10.6	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38	4.27	3.92	3.73	3.54	3.33	3.22	3.11
17	10.4	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25	4.14	3.79	3.61	3.41	3.21	3.10	2.98
18	10.2	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14	4.03	3.68	3.50	3.30	3.10	2.99	2.87
19	10.1	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04	3.93	3.59	3.40	3.21	3.00	2.89	2.78
20	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.50	3.32	3.12	2.92	2.81	2.69
21	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88	3.77	3.43	3.24	3.05	2.84	2.73	2.61
22	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81	3.70	3.36	3.18	2.98	2.77	2.66	2.55
23	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75	3.64	3.30	3.12	2.92	2.71	2.60	2.48
24	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69	3.59	3.25	3.06	2.87	2.66	2.55	2.43
25	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64	3.54	3.20	3.01	2.82	2.61	2.50	2.38
26	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60	3.49	3.15	2.97	2.77	2.56	2.45	2.33
27	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56	3.45	3.11	2.93	2.73	2.52	2.41	2.29
28	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52	3.41	3.07	2.89	2.69	2.48	2.37	2.25
29	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48	3.38	3.04	2.86	2.66	2.45	2.33	2.21
30	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.01	2.82	2.63	2.42	2.30	2.18
40	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22	3.12	2.78	2.60	2.40	2.18	2.06	1.93
60	8.49	5.80	4.73	4.14	3.76	3.49	3.29	3.32	3.01	2.90	2.57	2.39	2.19	1.96	1.83	1.69
120	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.37	2.19	1.98	1.75	1.61	1.43
∞	7.88	5.30	4.28	3.72	3.35	3.09	3.90	2.74	2.62	2.52	2.19	2.00	1.79	1.53	1.36	1.00

附表5 Γ函数表

x	Γ(x)									
	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
1.00	1.000 0	0.999 4	0.998 8	0.998 3	0.997 7	0.997 1	0.996 6	0.996 0	0.995 4	0.994 9
1.01	0.994 3	0.993 8	0.993 2	0.992 7	0.992 1	0.991 6	0.991 0	0.990 5	0.989 9	0.989 4
1.02	0.988 8	0.988 3	0.987 8	0.987 2	0.986 7	0.986 2	0.985 6	0.985 1	0.984 6	0.984 1
1.03	0.983 5	0.983 0	0.982 5	0.982 0	0.981 5	0.981 0	0.980 5	0.980 0	0.979 4	0.978 9
1.04	0.978 4	0.977 9	0.977 4	0.976 9	0.976 4	0.975 9	0.975 5	0.975 0	0.974 5	0.974 0
1.05	0.973 5	0.973 0	0.972 5	0.972 1	0.971 6	0.971 1	0.970 6	0.970 2	0.969 7	0.969 2
1.06	0.968 7	0.968 3	0.967 8	0.967 3	0.966 9	0.966 4	0.966 0	0.965 5	0.965 1	0.964 6
1.07	0.961 2	0.963 7	0.963 3	0.962 8	0.962 4	0.961 9	0.961 5	0.961 0	0.960 6	0.960 2
1.08	0.959 7	0.959 3	0.958 9	0.958 4	0.958 0	0.957 6	0.957 1	0.956 7	0.956 3	0.955 9
1.09	0.955 5	0.955 0	0.954 6	0.954 2	0.953 8	0.953 4	0.953 0	0.952 6	0.952 2	0.951 3
1.10	0.951 4	0.950 9	0.950 5	0.950 1	0.949 8	0.949 4	0.949 0	0.948 6	0.948 2	0.947 8
1.11	0.947 4	0.947 0	0.946 6	0.946 2	0.945 9	0.945 5	0.945 1	0.944 7	0.944 3	0.944 0
1.12	0.943 6	0.943 2	0.942 8	0.942 5	0.942 1	0.941 7	0.941 4	0.941 0	0.940 7	0.940 3
1.13	0.939 9	0.939 6	0.939 2	0.938 9	0.938 5	0.938 2	0.937 8	0.937 5	0.937 1	0.936 8
1.14	0.936 4	0.936 1	0.935 7	0.935 4	0.935 0	0.934 7	0.934 4	0.934 0	0.933 7	0.933 4
1.15	0.933 0	0.937 2	0.932 4	0.932 1	0.931 7	0.931 4	0.931 1	0.930 8	0.930 4	0.930 1
1.16	0.929 8	0.929 5	0.929 2	0.928 9	0.928 5	0.928 2	0.927 9	0.927 6	0.927 3	0.927 0
1.17	0.926 7	0.926 4	0.926 1	0.925 8	0.925 5	0.925 2	0.924 9	0.924 6	0.924 3	0.924 0
1.18	0.923 7	0.923 4	0.923 1	0.922 9	0.922 3	0.922 3	0.922 0	0.921 7	0.921 4	0.921 2
1.19	0.920 9	0.920 6	0.920 3	0.920 1	0.919 8	0.919 5	0.919 2	0.919 0	0.918 7	0.918 4
1.20	0.918 2	0.917 9	0.917 6	0.917 4	0.917 1	0.916 9	0.916 6	0.916 3	0.916 1	0.915 8
1.21	0.915 6	0.915 3	0.915 1	0.914 8	0.914 6	0.914 3	0.914 1	0.913 8	0.913 6	0.913 3
1.22	0.913 1	0.912 9	0.912 6	0.912 4	0.912 2	0.914 9	0.911 7	0.911 4	0.911 2	0.911 0
1.23	0.910 8	0.910 5	0.910 3	0.910 1	0.909 8	0.909 6	0.909 4	0.909 2	0.909 0	0.908 7
1.24	0.908 5	0.908 3	0.908 1	0.907 9	0.907 7	0.907 4	0.907 2	0.907 0	0.906 8	0.906 6
1.25	0.906 4	0.906 2	0.906 0	0.905 8	0.905 6	0.905 4	0.905 2	0.905 0	0.904 8	0.904 6
1.26	0.904 4	0.904 2	0.904 0	0.903 8	0.903 6	0.903 4	0.903 2	0.903 1	0.902 9	0.902 7
1.27	0.902 5	0.902 3	0.902 1	0.902 0	0.901 8	0.901 6	0.901 4	0.901 2	0.901 1	0.900 9
1.28	0.900 7	0.900 5	0.900 4	0.900 2	0.900 0	0.899 9	0.899 7	0.899 5	0.899 4	0.899 2
1.29	0.899 0	0.898 9	0.898 7	0.898 6	0.898 4	0.898 2	0.898 1	0.897 9	0.897 8	0.897 6
1.30	0.897 5	0.897 3	0.897 2	0.897 0	0.896 9	0.896 7	0.896 5	0.896 4	0.896 3	0.896 1
1.31	0.896 0	0.895 9	0.895 7	0.895 6	0.895 4	0.895 3	0.895 2	0.895 0	0.894 9	0.894 8
1.32	0.894 6	0.894 5	0.894 4	0.894 3	0.894 1	0.894 0	0.893 9	0.893 7	0.893 6	0.893 5
1.33	0.893 4	0.893 3	0.893 1	0.893 0	0.892 9	0.892 8	0.892 7	0.892 6	0.892 4	0.892 3
1.34	0.892 2	0.892 1	0.892 0	0.891 9	0.891 8	0.891 7	0.891 6	0.891 5	0.891 4	0.891 2
1.35	0.891 2	0.891 1	0.891 0	0.890 9	0.890 8	0.890 7	0.890 6	0.890 5	0.890 4	0.890 3
1.36	0.890 2	0.890 1	0.890 0	0.889 9	0.889 8	0.889 7	0.889 7	0.889 6	0.889 5	0.889 4
1.37	0.889 3	0.889 2	0.889 2	0.889 1	0.889 0	0.888 9	0.888 8	0.888 8	0.888 7	0.888 6
1.38	0.888 5	0.888 5	0.888 4	0.888 3	0.888 3	0.888 2	0.888 1	0.888 0	0.888 0	0.887 9
1.39	0.887 9	0.887 8	0.887 7	0.887 7	0.887 6	0.887 5	0.887 5	0.887 4	0.887 4	0.887 3
1.40	0.887 3	0.887 2	0.887 2	0.887 1	0.887 1	0.887 0	0.887 0	0.886 8	0.886 9	0.886 8
1.41	0.886 8	0.886 7	0.886 7	0.886 6	0.886 6	0.886 5	0.886 5	0.886 5	0.886 4	0.886 4
1.42	0.886 4	0.886 3	0.886 3	0.886 3	0.886 2	0.886 2	0.886 2	0.886 1	0.886 1	0.886 1
1.43	0.886 0	0.886 0	0.886 0	0.886 0	0.885 9	0.885 9	0.885 9	0.885 9	0.885 8	0.885 8
1.44	0.885 8	0.885 8	0.885 8	0.885 8	0.885 7	0.885 7	0.885 7	0.885 7	0.885 7	0.885 7
1.45	0.885 7	0.885 7	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6
1.46	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6	0.885 6
1.47	0.885 6	0.885 6	0.885 6	0.885 7	0.885 7	0.885 7	0.885 7	0.885 7	0.885 7	0.885 7
1.48	0.885 7	0.885 8	0.885 8	0.885 8	0.885 8	0.885 8	0.885 9	0.885 9	0.885 9	0.885 9
1.49	0.885 9	0.886 0	0.886 0	0.886 0	0.886 0	0.886 1	0.886 1	0.886 1	0.886 2	0.886 2
1.50	0.886 2	0.886 3	0.886 3	0.886 3	0.886 4	0.886 4	0.886 4	0.886 5	0.886 5	0.886 6
1.51	0.886 6	0.886 6	0.886 7	0.886 7	0.886 8	0.886 8	0.886 9	0.886 9	0.886 9	0.887 0

续附表 5

x	Γ (x)									
	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
1.52	0.887 0	0.887 1	0.887 1	0.887 2	0.887 2	0.887 3	0.887 3	0.887 4	0.887 5	0.887 5
1.53	0.887 6	0.887 6	0.887 7	0.887 7	0.887 8	0.887 9	0.887 9	0.888 0	0.888 0	0.888 1
1.54	0.888 2	0.888 2	0.888 3	0.888 4	0.888 4	0.888 5	0.888 6	0.888 7	0.888 7	0.888 8
1.55	0.888 9	0.888 9	0.889 0	0.889 1	0.889 2	0.889 2	0.889 3	0.889 4	0.889 5	0.889 6
1.56	0.889 6	0.889 7	0.889 8	0.889 9	0.890 0	0.890 1	0.890 1	0.890 2	0.890 3	0.890 4
1.57	0.890 5	0.890 6	0.890 7	0.890 8	0.890 9	0.890 9	0.891 0	0.891 1	0.891 2	0.891 3
1.58	0.891 4	0.891 5	0.891 6	0.891 7	0.891 8	0.891 9	0.892 0	0.892 1	0.892 2	0.892 3
1.59	0.892 4	0.892 5	0.892 6	0.892 7	0.892 9	0.893 0	0.893 1	0.893 2	0.893 3	0.893 4
1.60	0.893 5	0.893 6	0.893 7	0.893 9	0.894 0	0.894 1	0.894 2	0.894 3	0.894 4	0.894 6
1.61	0.894 7	0.894 8	0.894 9	0.895 0	0.895 2	0.895 3	0.895 4	0.895 5	0.895 7	0.895 8
1.62	0.895 9	0.891 6	0.896 2	0.896 3	0.896 4	0.896 6	0.896 7	0.896 8	0.897 0	0.897 1
1.63	0.897 2	0.897 4	0.897 5	0.897 7	0.897 8	0.897 9	0.898 1	0.898 2	0.898 4	0.898 5
1.64	0.898 6	0.898 8	0.898 9	0.899 1	0.899 2	0.899 4	0.899 5	0.899 7	0.899 8	0.900 0
1.65	0.900 1	0.990 3	0.990 4	0.900 6	0.900 7	0.900 9	0.901 0	0.901 2	0.901 4	0.901 5
1.66	0.901 7	0.901 8	0.902 0	0.902 1	0.902 3	0.902 5	0.902 6	0.902 8	0.903 0	0.901 3
1.67	0.903 3	0.903 5	0.903 6	0.903 8	0.904 0	0.904 1	0.904 3	0.904 5	0.904 7	0.904 8
1.68	0.905 0	0.905 2	0.905 4	0.905 5	0.905 7	0.905 9	0.906 1	0.906 2	0.906 4	0.906 6
1.69	0.906 8	0.907 0	0.907 1	0.907 3	0.907 5	0.907 7	0.907 9	0.901 8	0.908 3	0.908 4
1.70	0.908 6	0.908 8	0.909 0	0.909 2	0.909 4	0.909 6	0.909 8	0.910 0	0.910 2	0.910 4
1.71	0.910 6	0.910 8	0.911 0	0.911 2	0.911 4	0.911 6	0.911 8	0.912 0	0.912 2	0.912 4
1.72	0.912 6	0.912 8	0.913 0	0.943 2	0.913 4	0.913 6	0.913 8	0.914 0	0.914 2	0.914 5
1.73	0.914 7	0.914 9	0.915 1	0.915 3	0.915 5	0.915 7	0.916 0	0.914 6	0.916 4	0.916 6
1.74	0.916 8	0.917 0	0.917 3	0.917 5	0.917 7	0.917 9	0.918 2	0.914 8	0.918 6	0.918 8
1.75	0.919 1	0.919 3	0.919 5	0.919 7	0.920 0	0.920 2	0.920 4	0.920 7	0.920 9	0.921 1
1.76	0.921 4	0.921 6	0.921 8	0.922 1	0.922 3	0.922 6	0.922 8	0.923 0	0.923 3	0.923 5
1.77	0.923 8	0.924 0	0.924 2	0.924 5	0.924 7	0.925 0	0.925 2	0.925 5	0.925 7	0.926 0
1.78	0.926 2	0.925 6	0.926 7	0.927 0	0.927 2	0.927 5	0.927 7	0.928 0	0.928 3	0.928 5
1.79	0.928 8	0.929 0	0.929 3	0.929 5	0.929 8	0.930 1	0.930 3	0.930 6	0.930 9	0.931 1
1.80	0.931 4	0.931 6	0.931 9	0.932 2	0.932 5	0.932 7	0.933 0	0.933 3	0.933 5	0.933 8
1.81	0.934 1	0.934 3	0.934 6	0.934 9	0.935 2	0.935 5	0.935 7	0.936 0	0.936 3	0.936 6
1.82	0.936 8	0.931 7	0.937 4	0.937 7	0.938 0	0.938 3	0.938 5	0.938 8	0.939 1	0.939 4
1.83	0.939 7	0.940 0	0.940 3	0.940 6	0.940 8	0.941 1	0.941 4	0.941 7	0.942 0	0.942 3
1.84	0.942 6	0.942 9	0.943 2	0.943 5	0.943 8	0.944 1	0.944 4	0.944 7	0.945 0	0.945 3
1.85	0.945 6	0.945 9	0.946 2	0.946 5	0.946 8	0.947 1	0.947 4	0.947 8	0.948 1	0.948 4
1.86	0.948 7	0.949 0	0.949 3	0.949 6	0.949 9	0.950 3	0.950 6	0.950 9	0.951 2	0.951 5
1.87	0.951 8	0.952 2	0.952 5	0.952 8	0.953 1	0.953 4	0.953 8	0.984 1	0.954 4	0.954 7
1.88	0.955 1	0.955 4	0.955 7	0.956 1	0.956 4	0.956 7	0.957 0	0.957 4	0.957 7	0.958 0
1.89	0.954 8	0.958 7	0.959 1	0.959 4	0.959 7	0.960 4	0.960 4	0.960 7	0.961 1	0.961 4
1.90	0.961 8	0.962 1	0.962 5	0.962 8	0.963 1	0.963 5	0.963 8	0.964 2	0.964 5	0.964 9
1.91	0.965 2	0.965 6	0.965 9	0.966 3	0.966 5	0.967 0	0.967 3	0.967 7	0.968 1	0.968 4
1.92	0.968 8	0.969 1	0.969 5	0.969 9	0.970 6	0.970 6	0.970 9	0.971 3	0.971 7	0.972 0
1.93	0.972 4	0.972 8	0.973 1	0.973 5	0.973 9	0.974 2	0.974 6	0.975 0	0.975 4	0.975 7
1.94	0.976 1	0.976 5	0.976 8	0.977 2	0.977 6	0.978 0	0.978 4	0.978 7	0.979 1	0.979 5
1.95	0.979 9	0.980 3	0.980 6	0.981 0	0.981 4	0.981 8	0.982 2	0.982 6	0.983 0	0.983 4
1.96	0.983 7	0.984 1	0.984 5	0.984 9	0.985 3	0.985 7	0.986 1	0.986 5	0.986 9	0.987 3
1.97	0.987 7	0.988 1	0.988 5	0.988 9	0.989 3	0.989 7	0.990 1	0.990 5	0.990 9	0.991 3
1.98	0.991 7	0.992 1	0.992 5	0.992 9	0.993 3	0.993 8	0.994 2	0.994 6	0.995 0	0.995 4
1.99	0.995 8	0.996 2	0.996 6	0.997 1	0.997 5	0.997 9	0.998 3	0.998 7	0.999 2	0.999 6

附表6　可靠性分析常用的概率分布

名称记号	概率分布及其定义域、参数条件	均值 $E(X)$	方差 $D(X)$	图　形
二项分布 $B(n,\ p)$	$P_B(X=x)=\binom{n}{x}p^x q^{n-x}$ $x=0,\ 1\cdots,\ n\quad 0<p<1,$ $p+q=1$　n 为正整数	np	npq	
泊松分布 $P(\lambda)$	$P_p(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}$ $x=0,\ 1,\ 2,\ \cdots\quad \lambda>0$	λ	λ	
对数正态分布（自然对数）$\ln(\mu,\ \sigma^2)$	$f_{\ln}(x)=\begin{cases}\frac{1}{\sigma x\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} & x>0\\ 0, & x\leqslant 0\end{cases}$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$	
对数正态分布（常用对数）$\lg(\mu,\ \sigma^2)$	$f_{\lg}(x)=\begin{cases}\frac{\lg e}{\sigma\lambda\sqrt{2\pi}}e^{-\frac{(\lg x-\mu)^2}{2\sigma^2}} & x>0\\ 0, & x\leqslant 0\end{cases}$	$10^{\mu+\frac{\sigma^2}{2}\ln10}$	$10^{2\mu+\sigma^2\ln10}\times(10^{\sigma^2\ln10}-1)$	
威布尔分布（Ⅲ型极值分布）$W(k,a,b)$	$f_W(x)=\begin{cases}\frac{k}{b}\left(\frac{x-a}{b}\right)^{k-1}\times e^{-\left(\frac{x-a}{b}\right)^k} & x\geqslant a\\ 0, & x<a\end{cases}$ 形状参数 $k>0$ 尺度参数 $b>0$ 位置参数 a	$b\Gamma\left(1+\frac{1}{k}\right)+a$	$b^2\left[\Gamma\left(1+\frac{2}{k}\right)-\Gamma^2\left(1+\frac{1}{k}\right)\right]$	
指数分布 $e(\lambda)$	$f_e(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}$ $\lambda>0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	
超几何分布 $H(n,M,N)$	$P_H(X=x)=\frac{\binom{N-M}{n-x}\binom{M}{x}}{\binom{N}{n}}$ $x=\max(0,\ n-N+M),\ \cdots,\ \min(n,\ M)$ $N,\ M,\ n$ 为正整数 $0\leqslant M\leqslant N,\ 0\leqslant n\leqslant N$	$\frac{nM}{N}$	$\frac{n(N-n)(N-M)M}{N^2(N-1)}$	

续附表 6

名称记号	概率分布及其定义域、参数条件	均值 $E(X)$	方差 $D(X)$	图　形
均匀分布 $u(a,\ b)$	$f_u(x)=\begin{cases}\frac{1}{b-a}, & a\leqslant x\leqslant b\\ 0, & x<a \text{ 或 } x>b\end{cases}$ $-\infty<a<b<\infty$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	
正态分布 (高斯分布) $N(\mu,\ \sigma^2)$	$f_N(x)=\frac{1}{\sigma\sqrt{2x}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty<x<\infty$ $-\infty<\mu<\infty,\ \sigma>0$	μ	σ^2	
瑞利分布 $R(\mu)$	$f_R(x)=\begin{cases}\frac{x}{\mu^2}e^{-\frac{x^2}{2\mu^2}}, & x\geqslant 0\\ 0, & x<0\end{cases}$ $\mu>0$	$\sqrt{\frac{\pi}{2}}\mu$	$\frac{(4-\pi)\ \mu^2}{2}$	
β 分布 (贝塔分布) $\beta(\alpha,\ \beta)$	$f_\beta(x)=\begin{cases}\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\beta(\alpha,\ \beta)}\\ 0<x<1\\ 0, & x\leqslant 0 \text{ 或 } x\geqslant 1\end{cases}$ 贝塔函数 $\beta(\alpha,\ \beta)=\frac{\Gamma(\alpha)\ \Gamma(\beta)}{\Gamma(\alpha+\beta)}$ $\alpha>0,\ \beta>0$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$	
Γ 分布 (伽马分布) $\Gamma(\alpha,\ \beta)$	$f_\Gamma(x)=\begin{cases}\frac{\beta\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x},\\ x>0\\ 0, & x\leqslant 0\end{cases}$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	
最大Ⅰ型极值分布 (贡贝尔分布) $M(\sigma,\ \mu)$	$f_M(x)=\frac{1}{\sigma}e^{-\frac{x-\mu}{\sigma}}e^{-e^{-\frac{x-\mu}{\sigma}}}$ $-\infty<x<\infty$ $-\infty<\mu<\infty$ $0<\sigma$	$\mu+\gamma\sigma$ 欧拉常数 $\gamma\approx 0.577\ 2$	$\frac{\pi^2}{6}\sigma^2$ $\approx(1.283\sigma)^2$	
最小Ⅰ型极值分布 (贡贝尔分布) $m(\sigma,\ \mu)$	$f_m(x)=\frac{1}{\sigma}e^{\frac{x-\mu}{\sigma}}e^{-e^{-\frac{x-\mu}{\sigma}}}$ $-\infty<x<\infty$ $-\infty<\mu<\infty$ $0<\sigma$	$\mu-\gamma\sigma$ 欧拉常数 $\gamma\approx 0.577\ 2$	$\frac{\pi^2}{6}\sigma^2$ $\approx(1.283\sigma)^2$	
X^2 分布 $X^2(\nu)$	$f_x^2(x)=$ $\begin{cases}\frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)}x^{\frac{\nu}{2}-1}\times e^{-\frac{x}{2}}, & x>0\\ 0, & x\leqslant 0\end{cases}$ 自由度 ν 为正整数	ν	2ν	

续附表6

名称记号	概率分布及其定义域、参数条件	均值 $E(X)$	方差 $D(X)$	图形
t分布 (学生分布) $t(\nu)$	$f_t(x)=\dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\times\left(1+\dfrac{x^2}{2}\right)^{-\frac{\nu+1}{2}}$ 自由度ν为正整数	0 $(\nu>1)$	$\dfrac{\nu}{\nu-2}$ $(\nu>2)$	$f_t(x)$; $\nu\to\infty$即$N(0,1)$; $\nu=10$; O; x
F分布 $F(\nu_1,\nu_2)$	$f_F(x)=\begin{cases}\dfrac{x^{\frac{\nu_1}{2}-1}}{\beta\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)}\left(\dfrac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}\times\left(1+\dfrac{\nu_1x}{\nu_2}\right)^{-\frac{\nu_1+\nu_2}{2}} & x\geqslant0\\0 & x<0\end{cases}$ 贝塔函数 $\beta\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)=\dfrac{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}+\frac{\nu_2}{2}\right)}$ 自由度ν_1，ν_2为正整数	$\dfrac{\nu_2}{\nu_2-2}$ $(\nu_2>2)$	$\dfrac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$ $(\nu_2>4)$	$f_F(x)$; $F(1,5)$; $F(10,10)$; $F(5.4)$; O; x

注：1. $\binom{N}{n}=\binom{N}{N-n}=\dfrac{N!}{(N-n)!\ n!}$；

2. $\Gamma(\cdot)$ —Γ函数，数值查附表5。

编写：谢里阳（东 北 大 学）
李 佳（东 北 大 学）
王永岩（青岛科技大学）

第7章 材料的环境强度设计

1 材料的高温强度

1.1 高温短时拉伸强度

高温短时拉伸是指在恒定的温度和规定的拉伸时间或拉伸速度下，用单向拉伸载荷把试样拉断的试验方法。这是一种测定材料高温强度和塑性指标的最基本手段，由此得到的一些性能指标是高温受力构件强度设计中常用的基本指标。

1.1.1 短时拉伸曲线

以拉伸载荷 P 作纵坐标，试样两标点间的伸长量 Δl 作横坐标画出的载荷－伸长量（$P-\Delta l$）曲线称为拉伸曲线。图 6.7-1 是退火低碳钢在室温下的典型的短时拉伸曲线。曲线中各点对应的载荷为：P_p 为比例极限载荷；P_e 为弹性极限载荷；P_s 为屈服极限载荷；P_b 为强度极限载荷；P_k 为断裂载荷。

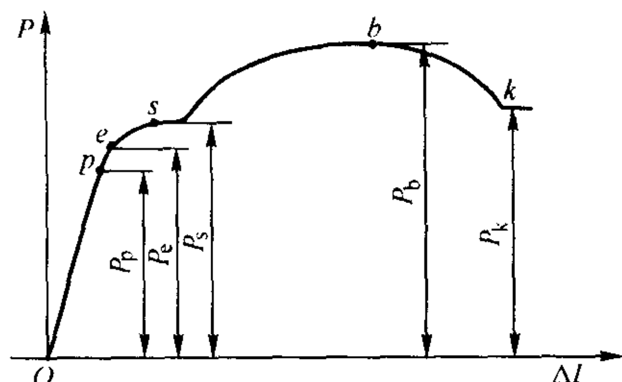


图 6.7-1 退火低碳钢室温下的拉伸曲线

图 6.7-2 为低碳钢及 1Cr18Ni9Ti 不锈钢在各种温度下的短时拉伸曲线。

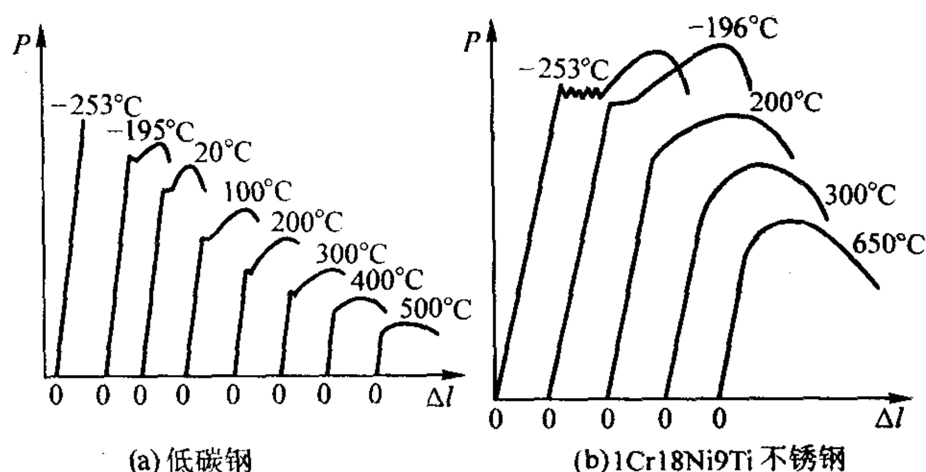


图 6.7-2 低碳钢及 1Cr18Ni9Ti 不锈钢在各种温度下的拉伸曲线

高温短时拉伸与室温拉伸相比，有些力学性能指标在高温下会呈现出与室温不相同的规律。在高温拉伸条件下塑性变形出现较早，碳钢的屈服点变得不明显，屈服强度难以测定。

通过拉伸试验可以直接获得两个原始的力学试验参量：试样所受的拉伸载荷 P 及试样伸长量 Δl。根据试样尺寸参数可以把它们转换成两个最基本的力学试验参量，即条件应力 σ 及条件应变 ε。由于试验时试样计算长度的不断增加及试样在变形过程中的不均匀性，上述两个参量 σ 及 ε 并不能反映出真正的应力－应变规律，因而需作适当的修正，由此建立起一组诱导力学试验参量，即真实应力 S 和真实应变 e 及条件均匀应变 λ，经过修正的缩颈真实应力 S_F 等。根据以上这些参量，可建立四种参量曲线，即载荷-伸长（$P-\Delta l$）曲线、条件应力－应变（$\sigma-\varepsilon$）曲线、真实应力-应变（$S-e$）曲线和缩颈真实应力－应变（S_F-e）曲线。选取相应的曲线确定下述四类力学性能指标：

1）强度指标　比例极限 σ_p，屈服点 σ_s，条件屈服强度 $\sigma_{0.2}$，抗拉强度 σ_b。

2）塑性指标　伸长率 δ，断面收缩率 ψ。

3）刚性指标　弹性模量 E。

4）塑性强化指标　强化系数 D，加工硬化指数 n。

1.1.2 应力、应变与温度的关系

在高温条件下，材料的变形机制增多，塑性变形易于进行，表现为强度降低，形变强化系数及形变强化指数下降，塑性变形 δ 及 ψ 增加。这些变化规律可用应力－温度曲线及应变－温度曲线来描述。

1）应力－温度曲线　抗拉强度 σ_b 与温度之间的关系可用 σ_b-T 曲线表示。对于大多数碳钢、CrMoV 钢及耐热不锈钢，σ_b-T 曲线的变化大致可分为三个阶段：在初始阶段，温度较低，σ_b 随温度升高明显下降；在中间阶段，σ_b 缓慢下降；在温度较高的第三阶段，σ_b 急速下降。碳钢及某些低合金钢如 CrMoV 钢，在中间阶段 σ_b 会出现一个峰值，这是时效硬化的结果。峰值温度与蓝脆温度相当。图 6.7-3 为 20 钢、15CrMoV 钢及 18－8 不锈钢的 σ_b-T 曲线。

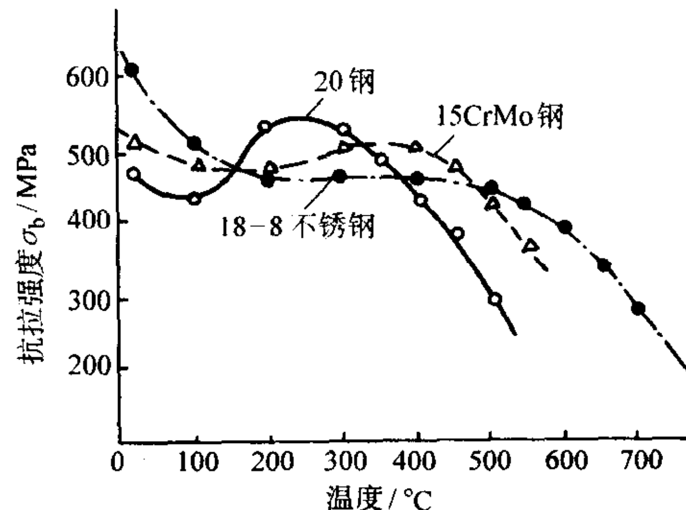


图 6.7-3 20 钢、15CrMo 钢及 18－8 不锈钢的 σ_b-T 曲线

条件屈服强度 $\sigma_{0.2}$ 与温度之间也有类似的关系，但中间阶段的峰值不明显，而在接近 400℃处出现一个小的峰值。图 6.7-4 为上述三种钢的 $\sigma_{0.2}-T$ 曲线。

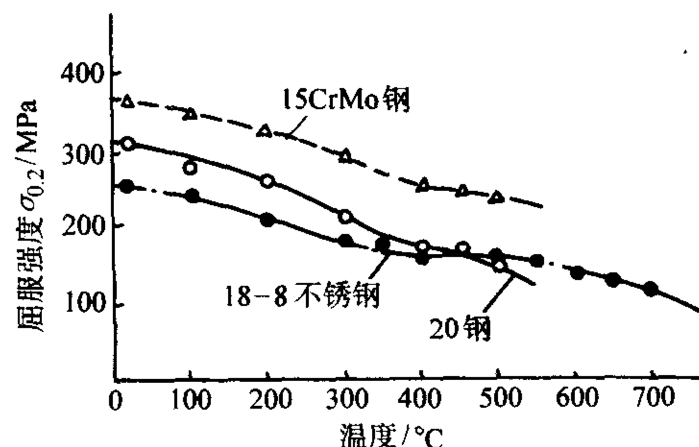


图 6.7-4 20 钢、15CrMo 钢及 18－8 不锈钢的 $\sigma_{0.2}-T$ 曲线

各种金属的屈服强度 σ_s 都随温度的升高而降低，其变化规律如图 6.7-5 所示。

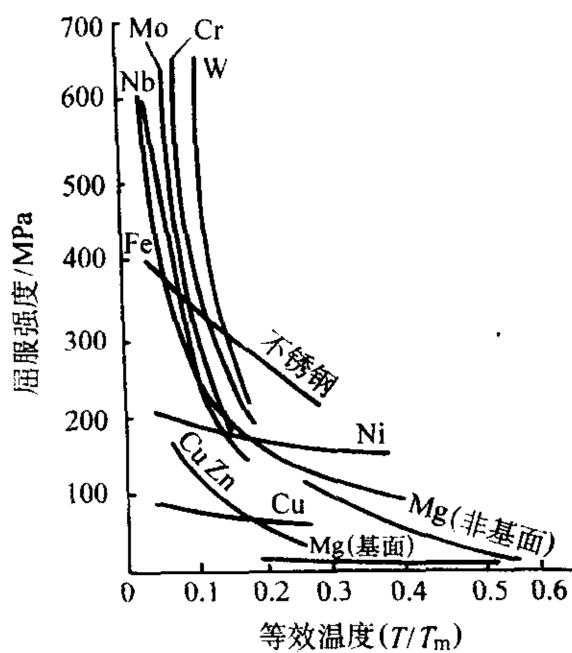


图 6.7-5 不同温度下金属的屈服强度

σ_b 与 $\sigma_{0.2}$ 随温度变化的规律与金属材料的热处理制度有关。图 6.7-6 为不同热处理对抗拉强度随温度下降率的影响。

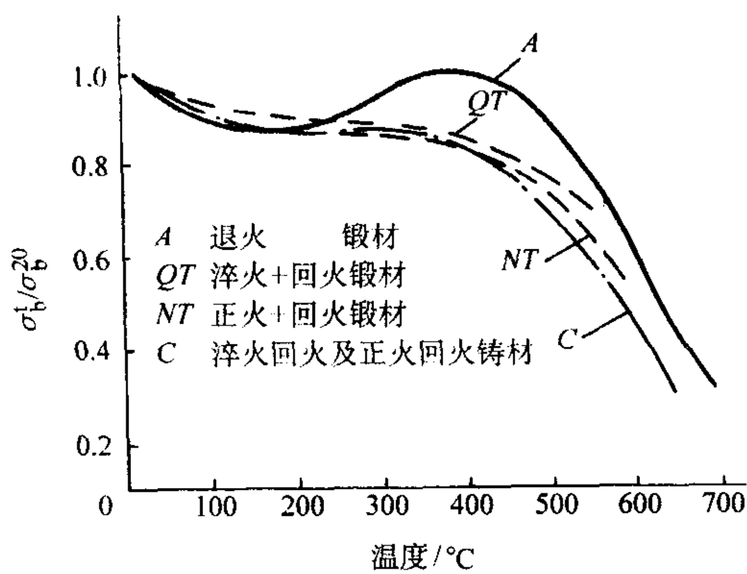


图 6.7-6 不同热处理对 2¼Cr－1Mo 钢抗拉强度随温度下降率的影响

2) 应变－温度曲线　应变随温度变化的规律可用 $\delta-T$ 曲线或 $\psi-T$ 曲线表示。碳钢及低合金钢（如 CrMoV 钢）的伸长率 δ 及断面收缩率 ψ 随温度的变化也可以分成三个阶段：在初始阶段，δ 和 ψ 随温度升高而逐渐下降；在中间阶段，δ 和 ψ 达到一个最低值，然后又开始回升；在第三阶段，随温度升高 δ 和 ψ 值大幅度升高。

图 6.7-7 为 20 钢及 15CrMo 钢的伸长率 δ_5 及 δ_{10} 与温度的关系，而图 6.7-8 为这两种钢的断面收缩率 ψ 与温度的关系。在回升阶段，δ 与 ψ 于接近 400℃处呈现出一个小的波峰，该温度与 $\sigma_{0.2}-T$ 曲线上谷值处的温度相对应。在其他类似的低合金耐热钢中这种现象也普遍存在。

1.2 高温硬度

1.2.1 高温硬度试验

硬度是反应材料抵抗局部塑性能力的力学性能指数，它以其测试简便易行而获得广泛应用，对于高温轴承及其某些工具材料等，材料的高温硬度是重要的力学性能指标之一。

在进行高温硬度试验时，涉及到试样的加热、保湿和防止氧化等一系列问题。目前，在试样温度不太高的情况下，仍用布氏、洛氏和维氏硬度试验方法。在硬度计的工作台上需加装一套试样加热保湿装置，加长压头的压杆，使之伸人加热保湿装置即可。如试验温度较高时，需使用特制的硬度计。在温度不超过 800℃时，硬度计的压头可用金刚石锥体（洛氏和维氏）和硬质合金球（布氏和洛氏）。当试验温度更高时，则应换用人造蓝宝石或刚玉压头。

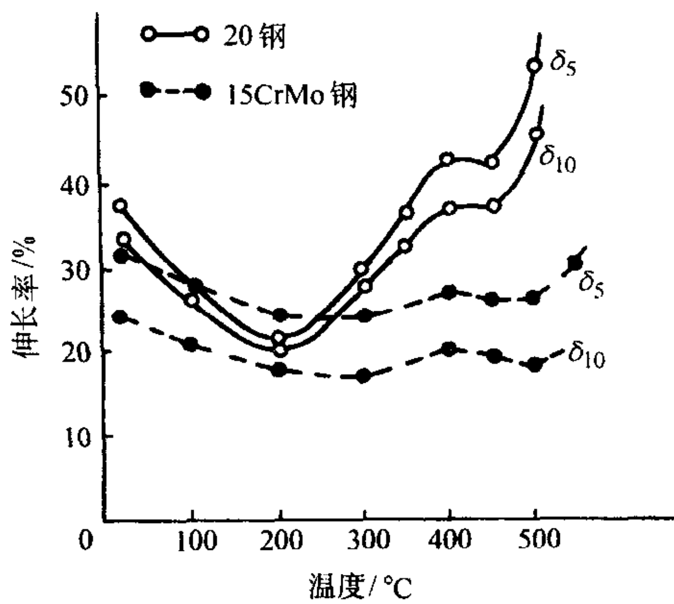


图 6.7-7 20 钢、15CrMo 钢的 δ-T 曲线

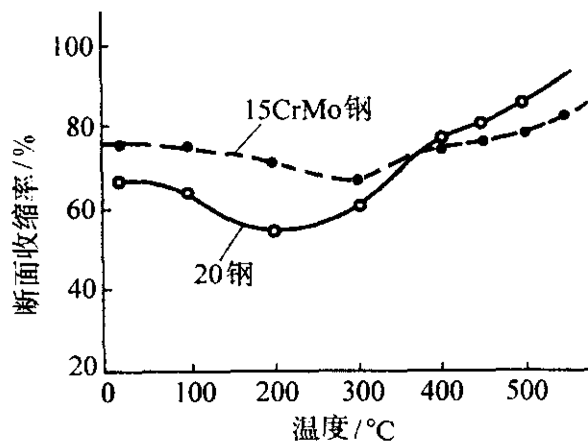


图 6.7-8 20 钢、15CrMo 钢的 $\psi-T$ 曲线

由于试验在较高温度下的硬度较低，所以试验压力不宜过大，并应根据试验温度的高低，改变收验压力的大小，以保证压痕清晰和完整。此外，由于试样在高温下蠕变的影响较大，一般规定加载时间为 30～60 s，但有时为了显示蠕变的影响，将加载时间延长到 1～5 h 时，所得结果叫持久硬度。试样上压痕直径（布氏硬度）或对角线（维氏硬度）的测量，一般是在试样冷却后进行的。

1.2.2 高温硬度与温度的关系

各种纯金属的硬度与温度之间有如下的关系

$$H = Ae^{-BT} \tag{6.7-1}$$

式中，H 为硬度；T 为试验温度；A、B 为常数。

两边取对数

$$\lg H = \lg A - \frac{BT}{2.3} \tag{6.7-2}$$

式（6.7-2）表明，在对数坐标图上，硬度（对数坐标）与温度成直线关系。试验表明，在某个温度处直线有一个转折点，形成两段斜率不同的线段，如图 6.7-9 所示。转折点位置与金属的熔点 T_m 有关，一般发生在（0.4～0.6）T_m 之间。硬度与温度关系直线上的转折点表明了变形机理的转变：由低温的滑移机制转变为高温扩散变形机制（如位错的攀移、多边化、晶粒边界的旋转）。

1.3 蠕变及持久强度

在一定温度下，金属受持续应力的作用而产生缓慢的塑性变形的现象称为金属的蠕变。引起蠕变的这一应力称蠕变应力。在这种持续应力作用下，蠕变变形逐渐增加，最终可

以导致断裂，这种断裂称蠕变断裂。导致断裂的这一初始应力称蠕变断裂应力。在有些情况下（特别是在工程上），把蠕变应力及蠕变断裂应力作为材料在特定条件下的一种强度指标来讨论时，往往又把它们称为蠕变强度及蠕变断裂强度，后者又称为持久强度。

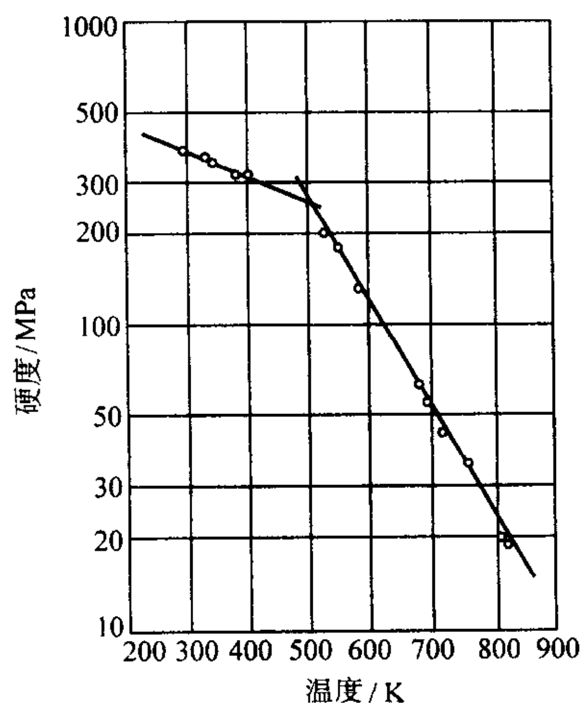


图 6.7-9 典型的纯铝硬度与温度的关系

1.3.1 蠕变曲线

在恒定温度下，一个受单向恒定载荷（拉或压）作用的试样，其变形 ε 与时间 t 的关系可用如图 6.7-10 所示的典型

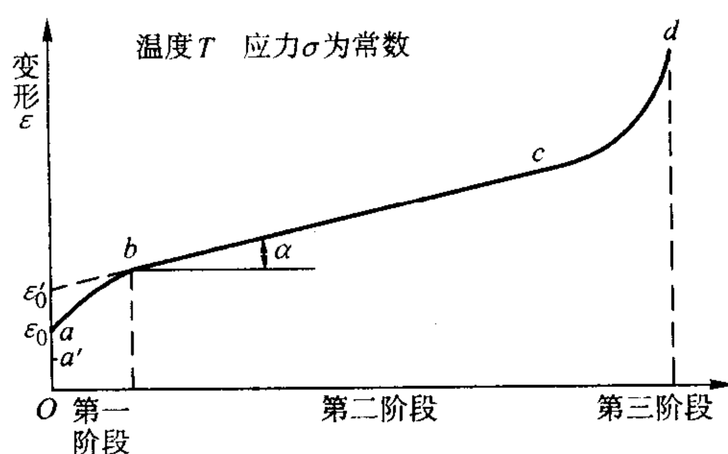


图 6.7-10 典型的蠕变曲线

的蠕变曲线表示。曲线可分下列几个阶段：

Ob 部分包括瞬态变形 *Oa* 和蠕变变形 *ab*，瞬态变形又包括弹性变形 *Oa'* 及瞬态塑性变形 *a'a* 两部分。*ab* 部分为蠕变第一阶段，称蠕变起始阶段。这部分的蠕变速度是逐步减小的，故又称蠕变减速阶段。

bc 部分称蠕变第二阶段，这部分的蠕变变形量与时间成线性关系，即蠕变速度维持恒定，在整个蠕变过程中，这部分的蠕变速度最小，故又称稳态蠕变阶段或最小蠕变速度阶段。

cd 部分称蠕变第三阶段，蠕变变形速度逐渐增加，直至 *d* 点断裂，这部分又称蠕变加速阶段。

尽管材料及其试验温度、应力各不相同，但蠕变曲线这三个阶段的特征却基本相同。温度及应力只影响各个阶段的持续时间及变形大小。温度较低，应力较小时，第二阶段的时间长一些；在高温高应力下，第二阶段较短。

1.3.2 蠕变强度及持久强度

在工程上，需按蠕变强度及持久强度确定许用应力。蠕变强度及持久强度是表示材料抵抗因外力作用导致蠕变变形或蠕变断裂的能力，是材料本身所具有的一种固有性能。

蠕变强度是材料在规定的蠕变条件（在一定的温度下及一定的时间内，达到一定的蠕变变形或蠕变速度）下保持不失效的最大承载应力。在测量中以失效应力表示。通常以试样在恒定温度和恒定拉伸载荷下，在规定时间内伸长（总伸长或残余伸长）率达到某规定值或第二阶段蠕变速度达到某规定值时的蠕变应力表示蠕变强度。根据不同的试验要求，蠕变强度有以下两种表示法。

1）在规定时间内达到规定变形量的蠕变强度，记为 $\sigma_{\delta/\tau}^{T}$，单位为 MPa；其中 T 为温度，℃；δ 为伸长率（总伸长或残余伸长，%）；τ 为持续时间，h。

这种蠕变强度一般用于需要提供总蠕变变形的构件设计。对短时蠕变试验，蠕变速度往往较大，第一阶段的蠕变变形量所占的比例较大，第二阶段的蠕变速度不易确定，所以用总蠕变变形作测量对象比较合适。

2）稳态蠕变速度达到规定值时的蠕变强度，记为 σ_{v}^{T}，单位为 MPa；其中 T 为温度，℃，v 为稳态蠕变速度，%/h。

这种蠕变强度通常用于一般受蠕变变形控制的运行时间较长的构件。因为在这种条件下蠕变速度较小，第一阶段的变形量所占的比例较小，蠕变的第二阶段明显，最小蠕变速度容易测量。

持久强度是材料在规定的蠕变断裂条件（一定的温度和规定的时间）下保持不失效的最大承载应力。通常，以试样在恒定温度和恒定拉伸载荷下到达某规定时间发生断裂时的蠕变断裂应力表示持久强度，记为 σ_{τ}^{T}，单位为 MPa。

1.3.3 高温应力松弛

在一定的温度下，一个受拉或受压的金属零件，若使用过程中总变形保持不变，则应力会自发下降，这种现象称为应力松弛。

一般认为，金属零件中产生应力松弛主要是在总变形不变的条件下，一部分弹性变形转变成塑性变形，致使弹性力减少。弹性变形的减少与塑性变形的增加是同时等量进行的。

在弹性范围内加载时，应力松弛过程可以表示为：

$$\varepsilon_0 = \varepsilon_p + \varepsilon_e = 常数$$
$$\sigma \neq 常数 \tag{6.7-3}$$

式中，ε_0 为总变形；ε_p 为弹性变形；ε_e 为塑性变形；σ 为试样或零件的应力。

根据应力随时间逐步降低的规律，以应力和时间为坐标绘出松弛曲线。典型的松弛曲线如图 6.7-11。图中，一开始瞬时的应力称为初应力 σ_0。以后应力下降逐渐减缓，称为松弛第二阶段。最后，曲线逐渐趋向与时间轴平行，此时的剩余应力称为松弛极根。它表示在一定的初应力和温度下，不再继续发生松弛的剩余应力，如图中 σ_r。

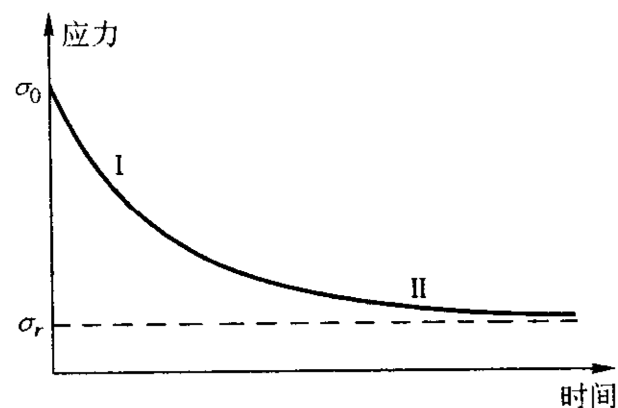


图 6.7-11 典型的松弛曲线

1.4 高温疲劳

广义的高温疲劳是指高于常温的疲劳现象。但由于有些零部件的工作温度虽高于室温，但并不太高。其疲劳设计，只要考虑温度对疲劳极限的影响，仍用室温下的疲劳设计方

法。只有当温度高于 $0.5T_m$（T_m 为以热力学温度表示的熔点），或在再结晶温度以上时，出现了蠕变与机械疲劳复合的疲劳现象，这时才称为高温疲劳。

1.4.1　高温对材料力学性能的影响

高温对材料的抗拉强度 σ_b 和屈服点 σ_s 的影响，见图6.7-12和图6.7-13。高温对材料疲劳极限的影响，见图6.7-14～图6.7-20。

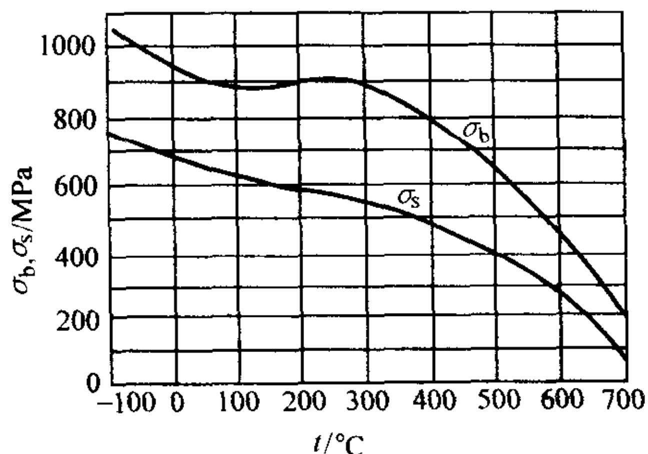


图 6.7-12　温度对钢 σ_b 和 σ_s 影响

（50CrMo钢，600℃回火）

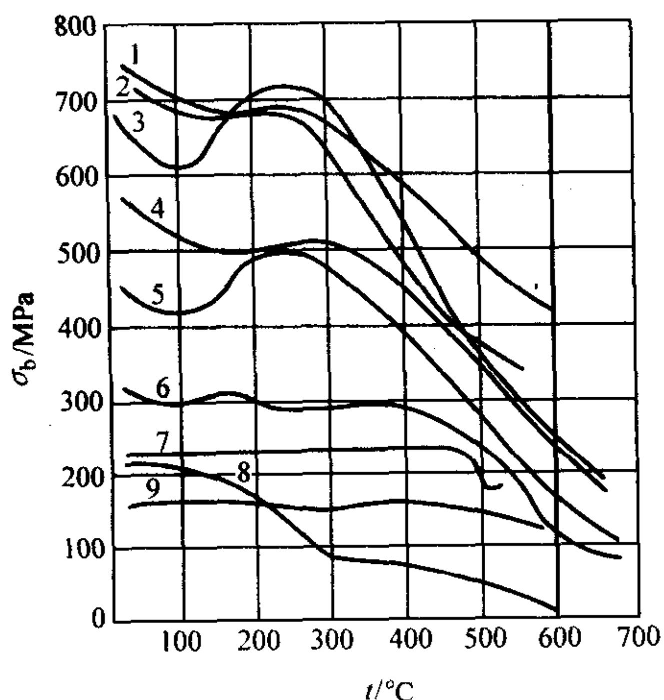


图 6.7-13　高温对材料 σ_b 的影响

1—钢（$w_C=0.4\%$，$w_{Cr}=13\%$）；2—蒙乃尔合金（$w_{Ni}=68\%$，$w_{Cu}=28\%$，$w_{Mn}=1.5\%$，$w_{Fe}=2.5\%$），轧制；3—钢（$w_C=0.4\%$），淬火；4—铸钢（$w_C=0.4\%$）；5—钢（$w_C=0.3\%$）；6—熟铁；7—可锻铸铁；8—黄铜；9—铸铁

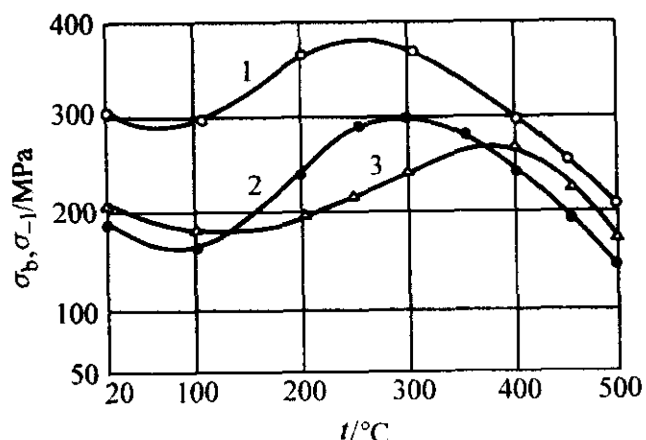


图 6.7-14　温度对 $w_C=0.17\%$ 碳钢疲劳强度的影响

1—抗拉强度 σ_b；2—弯曲疲劳极限，$f=10$ Hz；3—弯曲疲劳极限，$f=2\,000$ Hz

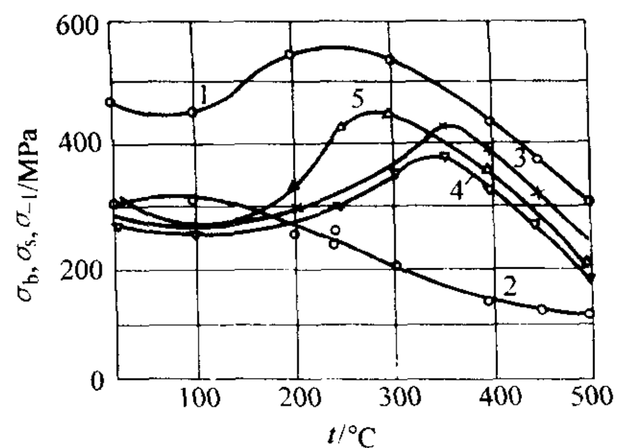


图 6.7-15　温度对 0.17%C 钢的静强度及疲劳强度的影响

1—抗拉强度 σ_b；2—屈服强度（点）σ_s；
3—在 33 Hz 下的旋转弯曲疲劳极限（$N=5\times10^5$）；
4—在 33 Hz 下的旋转弯曲疲劳极限（$N=10^8$）；
5—在 0.17 Hz 下的旋转弯曲疲劳极限（$N=5\times10^5$）

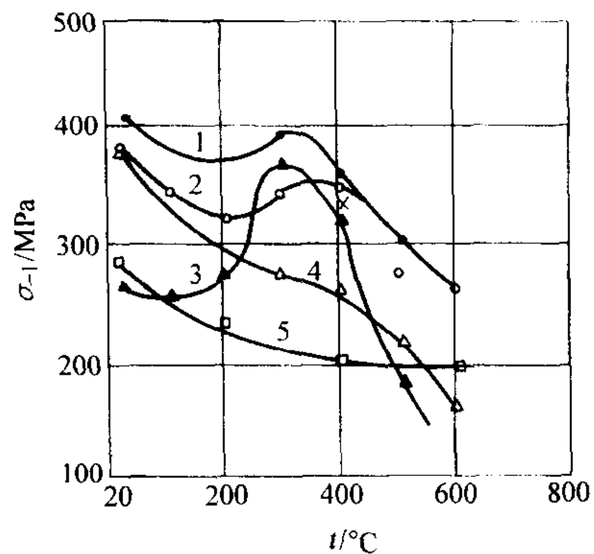


图 6.7-16　温度对材料疲劳极限影响

1—30CrMo钢；2—30CrNiMo钢；3—钢（$w_C=0.17\%$）；4—1Cr13钢；5—1Cr18Ni9Ti钢

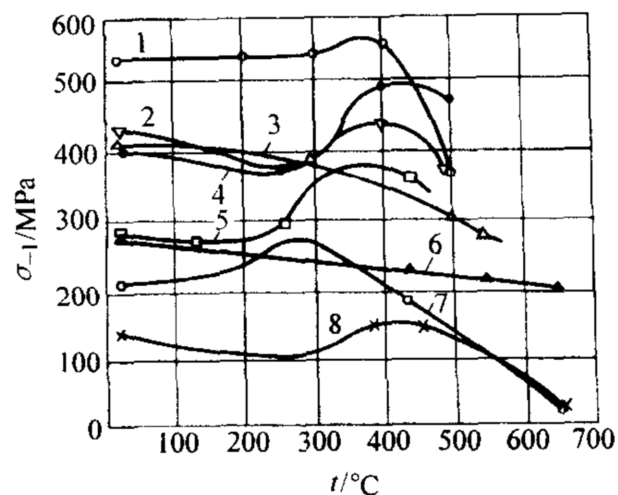


图 6.7-17　温度与旋转弯曲疲劳极限的关系

1—Ni-Cr钢；2—Cr-Mo-V钢；3—钢（$w_C=12\%$）；4—钢（$w_C=0.5\%$）；5—钢（$w_C=0.25\%$）；6—18Cr-8Ni钢；7—钢（$w_C=0.17\%$）；8—铸铁

1.4.2　高温时材料 S－N 曲线

高温时材料的 S－N 曲线包括应力－寿命（$\sigma-N$）曲线和应变－寿命（$\varepsilon-N$）曲线，见图6.7-21～图6.7-32。

1.4.3　影响金属高温疲劳性能的主要因素

（1）材料因素

试验表明，疲劳强度（σ_{-1}）与强度极根（σ_b）之间存在着一定的关系，但是在不同的材料和不同的组织状态下，这种关系可在很宽的范围内变化。材料的疲劳极限与强度极限的比值 σ_{-1}/σ_b，称为疲劳比。对大多数材料，疲劳比随温度的升高而增高。表6.7-1示出了不同材料在不同温度下的疲劳比。由此可见，材料在不同温度下的疲劳极限和强度极限，均需单独试验确定，不宜借助疲劳比相互换算。

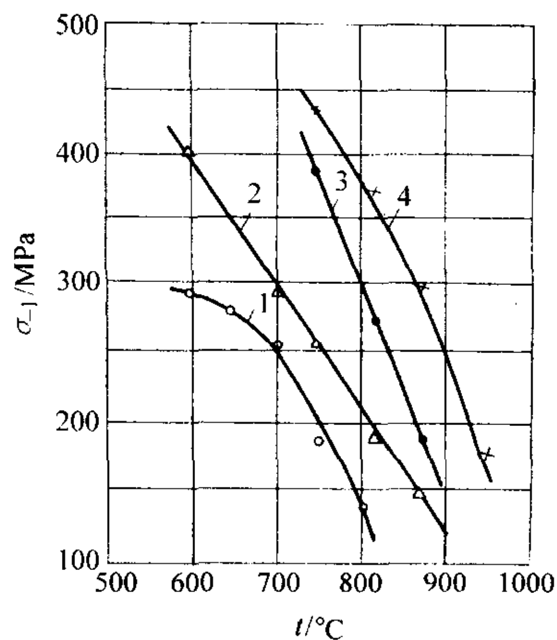


图 6.7-18　温度对尼莫尼克合金疲劳强度的影响

1—尼莫尼克 80，轴向对称循环应力，$N=4\times10^7$；

2—尼莫尼克 90，轴向对称循环应力，$N=3.6\times10^7$；

3—尼莫尼克 90，旋转弯曲应力，$N=3.6\times10^7$；

4—尼莫尼克 100，旋转弯曲应力，$N=4.5\times10^7$

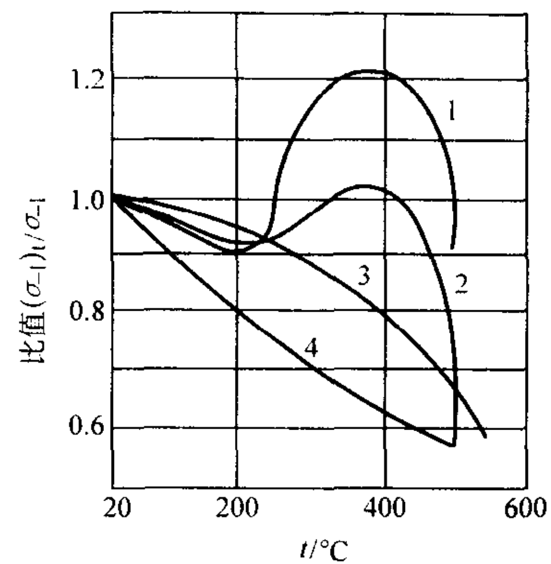


图 6.7-19　温度对材料疲劳极限的影响

1—钢（$w_C=0.48\%$）；2—Cr-Ni-Mo 钢；

3—钢（$w_{Cr}=12\%$）；4—耐热钢；

σ_{-1}—室温下的疲劳极限；

$(\sigma_{-1})_t$—温度 t 时的疲劳极限

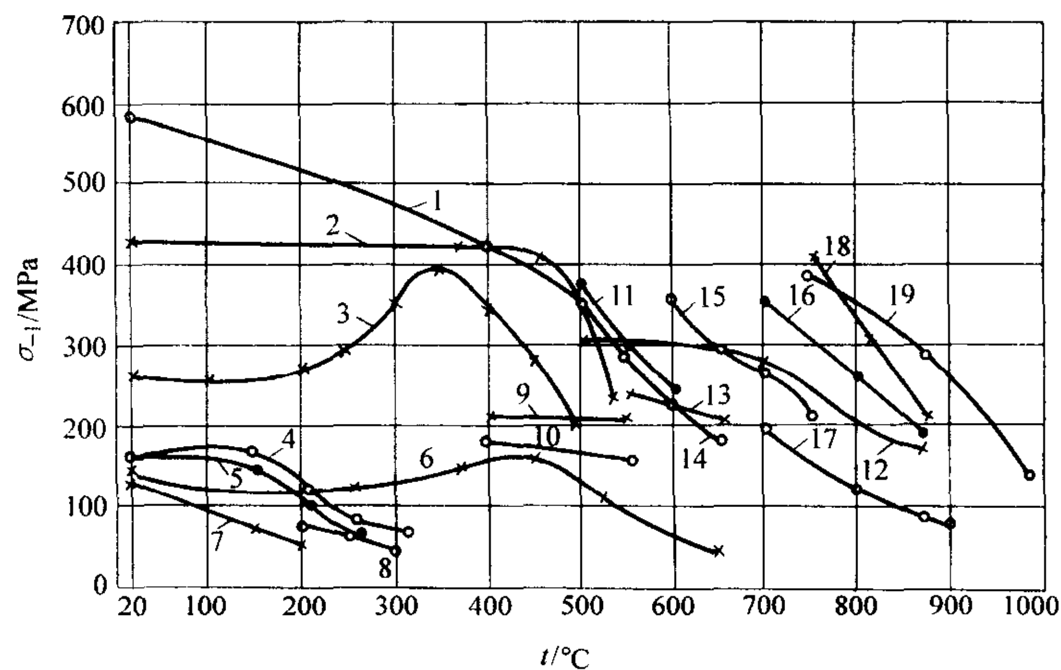


图 6.7-20　温度对金属材料的疲劳强度的影响

各合金及其成分（质量分数,%）如下：

1—钛合金（含铝的钛合金）；2—Ni-Cr-Mo 钢；3—低碳钢（C0.17）；4—铝铜合金；5—铝锌镁合金；

6—高强度铸铁；7—镁铝锌合金；8—镁锌锆钍合金；9—铜镍合金（Ni30，Cr0.5，Al1.5，其余 Cu）；

10—铜镍合金（Ni30，Mn1，Fe1，其余 Cu）；11—合金钢（Cr2.7，Mo0.5，V0.75，W0.5）；

12—奥氏体镍铬钼钢；13—奥氏体钢（Cr18.75，Ni12.0，Nb1.25）；14—合金钢（Cr11.6，Mo0.6，V0.3，Nb0.25）；

15—奥氏体钢（Cr13，Ni13，Co10）；16—钴合金（Cr19，Ni12，Co45）；17—奥氏体钢；

18—镍铬合金（Cr15，Co20，Ti1.2，Al4.5，Mo5，其余 Ni）；19—镍铬合金（20Cr，Co18，Ti2.4，Al1.4，其余 Ni）

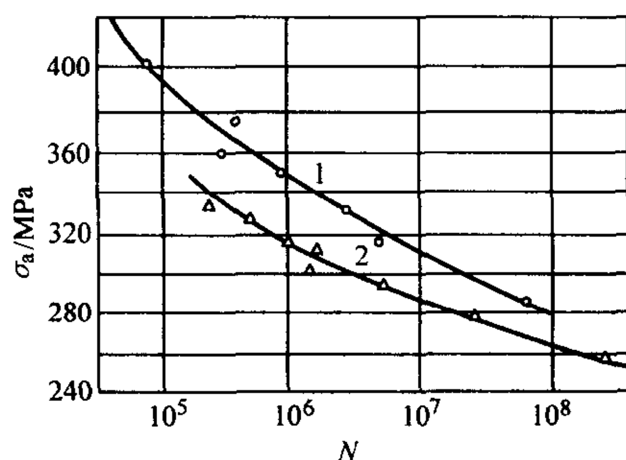


图 6.7-21　低碳钢在 400℃时的 $S-N$ 曲线

1—旋转弯曲疲劳；2—拉压疲劳

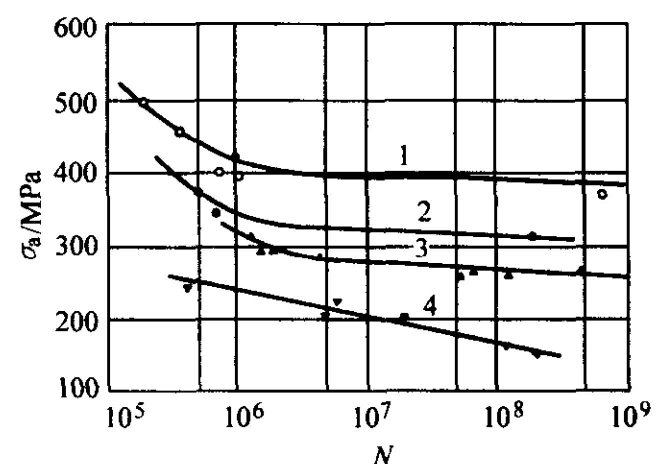


图 6.7-22　铁基合金 N-155 在高温下的旋转弯曲 $S-N$ 曲线

1—温度 $t=20$℃；2—温度 $t=650$℃；

3—温度 $t=730$℃；4—温度 $t=815$℃

N-155 的合金成分（质量分数,%）：C0.08～0.16，

Mn1.0～2.0，Si＜1，Cr20.0～22.5，

Ni19.0～21.0，Co18.5～21.0，Mo2.50～3.50，W2.0～3.0

Nb0.75～1.25，N0.10～0.20

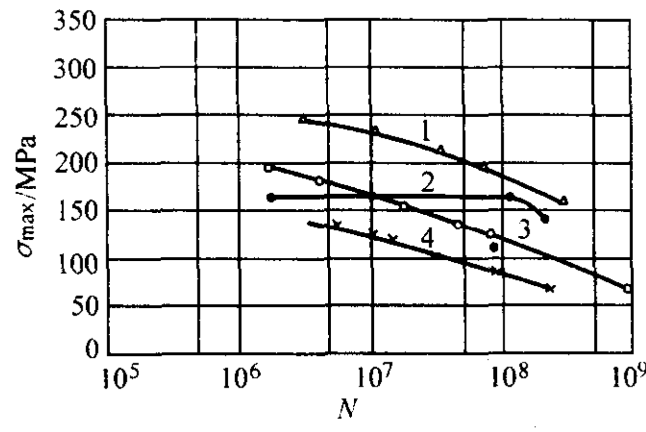


图 6.7-23 铁基合金 N-155 在 815℃时的 $S-N$ 曲线

1—应力比 $r=-0.242$；2—应力比 $r=-1$；

3—应力比 $r=0.6$；4—应力比 $r=1$

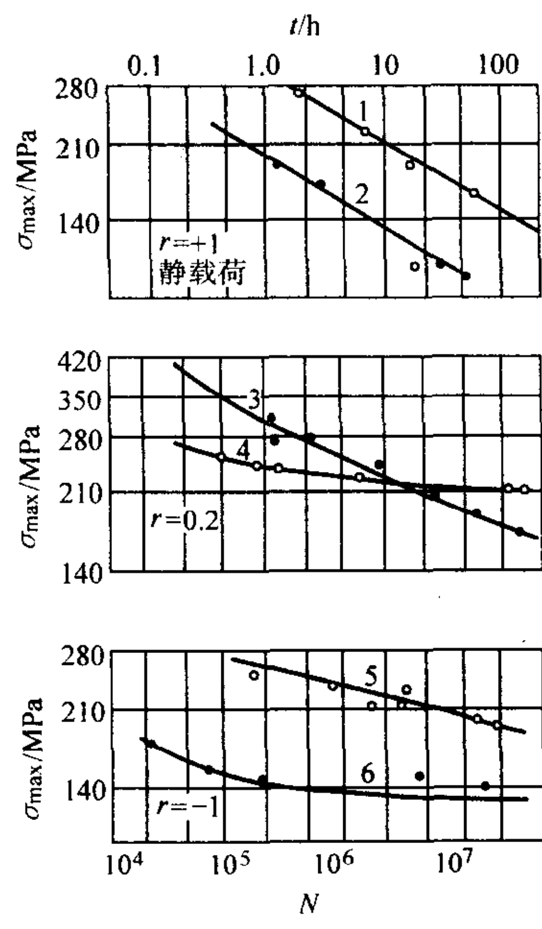


图 6.7-24 缺口对 S-816 合金在 900℃时的 $\sigma-t$ 和 $\sigma-N$ 曲线

1—$r=+1$，$\alpha_\sigma=3.4$；2—$r=+1$，$\alpha_\sigma=1$（光滑试样）；

3—$r=0.2$，$\alpha_\sigma=1$；4—$r=0.2$，$\alpha_\sigma=3.4$；

5—$r=-1$，$\alpha_\sigma=1$；6—$r=-1$，$\alpha_\sigma=3.4$

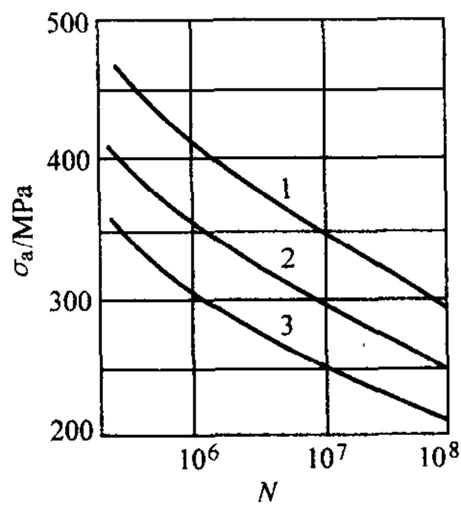


图 6.7-25 GH4037 合金的高温时的 $S-N$ 曲线

1—700℃；2—800℃；3—850℃

(2) 温度因素

随着温度的升高，疲劳强度一般有降低的趋势，越接近熔点，降低趋势越明显。疲劳强度的降低是由于发生了再结晶、扩散和溶解等过程引起的。但也有某些过程能提高疲劳强度，如时效硬化和应变硬化。因此，有些材料在高温时的疲劳强度反而比室温时高，疲劳强度随温度的变化规律比较复杂。表 6.7-2～表 6.7-5 也是温度对疲劳强度的影响数据。

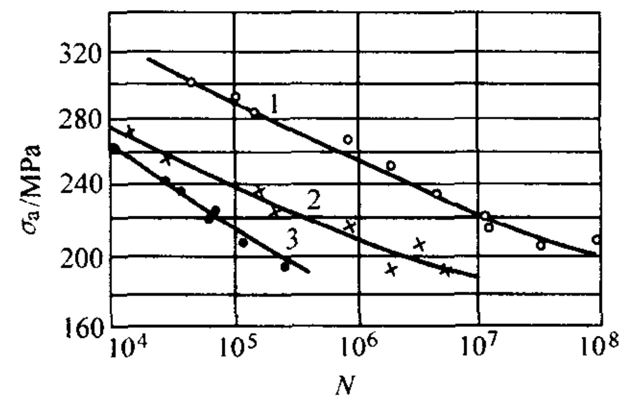


图 6.7-26 碳钢（$w_C=0.17\%$）在 450℃时，频率对拉压疲劳极限的影响

1—试验频率 $f=2\ 000\ \text{min}^{-1}$；

2—试验频率 $f=125\ \text{min}^{-1}$；

3—试验频率 $f=10\ \text{min}^{-1}$

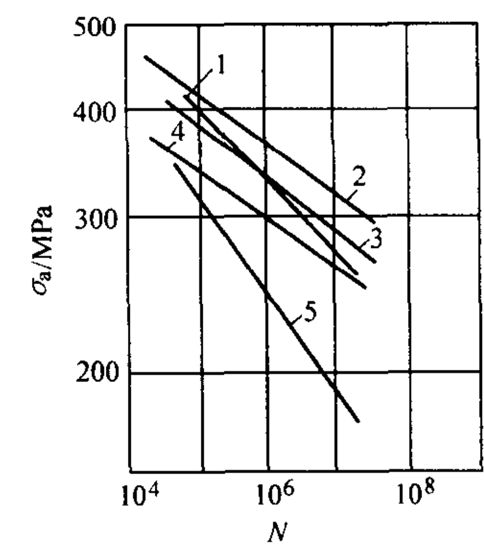


图 6.7-27 镍基高温合金在不同温度下的 $S-N$ 曲线

1—600℃；2—800℃；3—900℃；

4—950℃；5—1 000℃

镍基高温合金化学成分（质量分数，%）：

Cr5，W5，Mo4，Co4.5，Al5.5，

Ti2.8，C0.15，B0.0

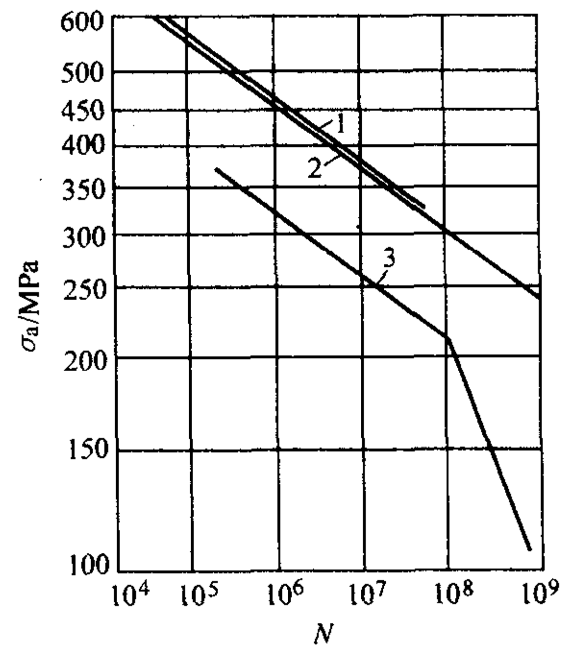


图 6.7-28 GH3032 合金在不同温度下的 $S-N$ 曲线

1—20℃；2—700℃；3—800℃

(3) 频率因素

高温疲劳的频率效应显著，主要是由于存在着蠕变作用的关系。频率低，应力作用的时间长，使蠕变的成分增加，裂纹扩展速度加快。此外，随频率的改变，断裂的特征也不同。频率较高时为穿晶断裂，较低时为沿晶断裂，中间则为混合断裂。图 6.7-33 示出了 A-286 合金的断口形态与频率的关系。图 6.7-34 示出了频率对 U-700 镍基高温合金疲劳寿命的影响。

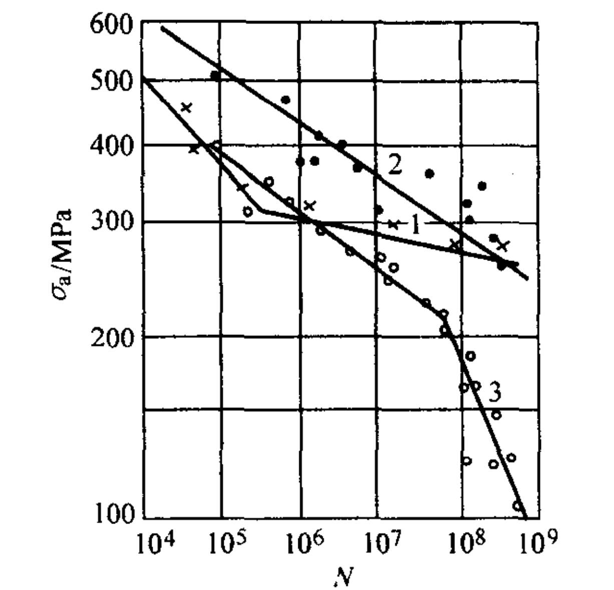


图 6.7-29　材料在高温下的 $S-N$ 曲线

1—钛合金，$t=200$℃；2—镍基合金，$t=700$℃；3—镍基合金，$t=800$℃

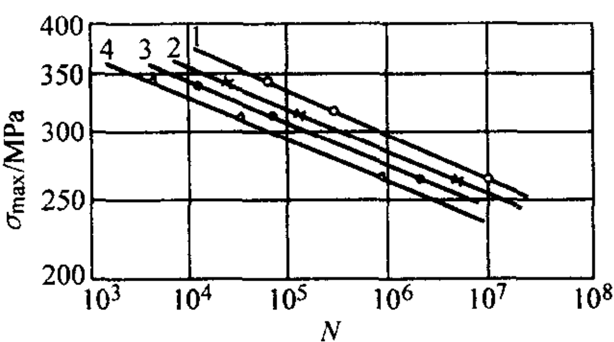


图 6.7-30　Cr2W9V 钢在 800℃时的 $p-S-N$ 曲线

1—存活率 $p=50\%$；2—存活率 $p=68\%$；3—存活率 $p=95.4\%$；4—存活率 $p=99.7\%$

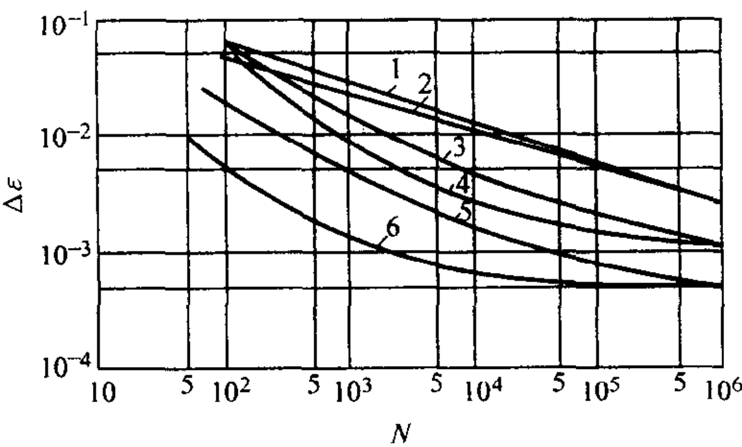


图 6.7-31　温度及频率对 304 奥氏体不锈钢低周疲劳 $\varepsilon-N$ 曲线的影响

1—$f=10\ \mathrm{min}^{-1}$，$t=430$℃；2—$f=10^{-3}\ \mathrm{min}^{-1}$，$t=430$℃；3—$f=10\ \mathrm{min}^{-1}$，$t=650$℃；4—$f=10\ \mathrm{min}^{-1}$，$t=816$℃；5—$f=10^{-3}\ \mathrm{min}^{-1}$，$t=650$℃；6—$f=10^{-3}\ \mathrm{min}^{-1}$，$t=816$℃

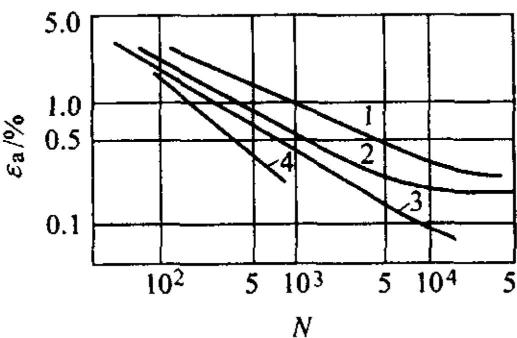


图 6.7-32　2.25Cr-1Mo 钢在高温对称弯曲时保持时间对 $\varepsilon-N$ 曲线的影响

1—室温，保持时间为 0，经过时间 1 min；
2—温度 600℃，保持时间为 0，经过时间 1 min；
3—温度 600℃，保持时间 30 min，经过时间 31 min；
4—温度 600℃，保持时间 300 min，经过时间 301 min

表 6.7-1　不同材料在不同温度下的疲劳比

材料	试验温度/℃	σ_{-1}/MPa	σ_b/MPa	疲劳比
GH3032 型	20	330	1 190	0.28
	600	343	940	0.36
	700	285	770	0.37
	800	235	780	0.30
GH4033 型	20	370	1 020	0.36
	600	360	—	—
	700	390	810	0.48
	800	260	620	0.42
GH4037 型	20	370	1 040	0.36
	700	380	880	0.43
	800	360	750	0.48
	900	280	520	0.54
尼莫尼克 80（Nimonic80）	20	346	820	0.42
	600	299	580	0.52
	650	288	—	—
	700	263	360	0.73
	750	195	—	—
	800	142	200	0.71

表 6.7-2　不同温度下材料的疲劳强度（一）

钢的主要化学成分（质量分数/%）	疲劳极限（$N=10^8$）σ_{-1}/MPa		
	20℃	70℃	100℃
C0.6，Mn0.7	430	370	—
C0.24，Ni3.9，Cr1.0	490	430	—
C0.2，Ni4.7，Cr1.4，Mo0.6	570	—	450

表 6.7-3　不同温度下材料的疲劳强度（二）

材料成分（质量分数）/%	旋转弯曲疲劳极限（$N=10^7$）σ_{-1}/MPa					
	20℃	100℃	200℃	300℃	400℃	500℃
灰铸铁（C3.2，Si1.1）	90	90	90	105	110	95
镍铬钢（Ni4.6，Cr1.6）	535	500	—	485	420	—
钢（C0.35）	298	—	310	330	—	275
钢（C0.60）	370	355	395	505	425	185
低合金钢（C0.14，Mo0.5）	315	—	—	400	370	275

表 6.7-4　不同温度下材料的疲劳强度（三）

铝合金	疲劳极限（1.2×10^8）σ_{-1}/MPa				
	20℃	150℃	200℃	250℃	300℃
DTD683（Zn5.5）	170	115	60	—	—
BSL65（Cu4.5）	130	80	57	39	39
DTD324（Si12）	127	85	60	39	29
DSL64（Cu4.5）	125	90	62	54	39

表 6.7-5　叶片钢的疲劳极限（$N=10^7$）

钢号	热处理	疲劳极限 σ_{-1}/MPa	20℃	200℃	300℃	400℃	500℃	550℃
1Cr13	1 030～1 050℃油淬 680～700℃回火	光滑试样	367	—	271	—	248	191
		缺口试样	183	—	114	—	104	100
2Cr13	1 000～1 020℃油淬 700～720℃回火	光滑试样	362	343	313	304	235	—

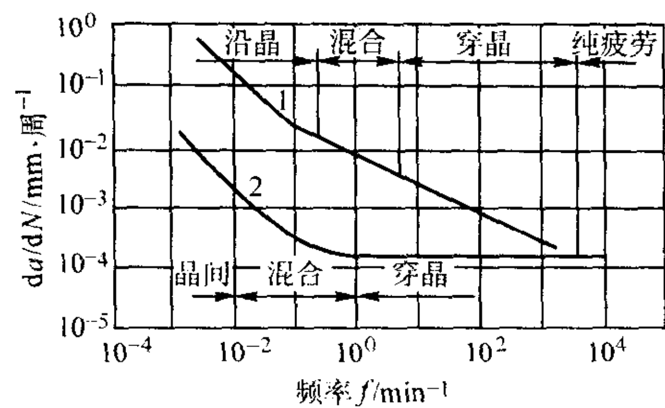


图 6.7-33 A-286 合金在 593℃时断口形貌与频率的关系

1—在空气中；2—在真空中

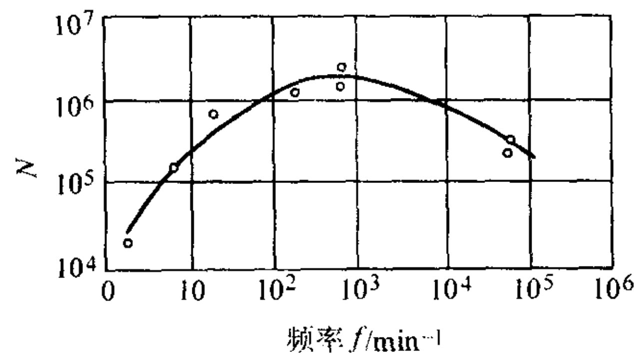


图 6.7-34 频率对 U-700 镍基合金在 760℃时的疲劳寿命的影响

(4) 应力集中因素

在高温下缺口产生的应力集中，大多数情况下会导致疲劳强度降低。缺口越尖锐，应力集中越严重，疲劳强度降低越多。表 6.7-6 为缺口对疲劳强度的影响。

表 6.7-6 缺口对疲劳强度的影响

材料	温度/℃	试验条件	疲劳极限/MPa 光滑试样	疲劳极限/MPa 缺口试样	理论应力集中系数 α_σ	有效应力集中系数 K_σ	敏性系数 q
GH4037	800	纯弯曲 180 kHz 100 h	350	250	2	1.40	0.4
	900		280	190	2	1.48	0.48
GH4033	20	纯弯曲 180 kHz 100 h	370	220	2	1.68	0.68
	600		360	240	2	1.50	0.50
	700		390	230	2	1.70	0.70
	800		260	230	2	1.13	0.13

一般讲，在有缺口时，高温疲劳强度是降低的。但是当应力比 r 不同时，也会出现不同的结果。图 6.7-24 为 S-816 合金在 900℃时的 S-N 曲线。当静载荷时 $r=+1$，$\alpha_\sigma=3.4$ 的缺口试样在同一应力水平下的寿命大于光滑试样的。当 $r=0.2$，即在蠕变和疲劳复合作用的情况下，在低寿命区，缺口试样的疲劳强度低于光滑试样；在高寿命区，缺口试样的疲劳强度高于光滑试样。当 $r=-1$，即在对称应力循环下，缺口试样的疲劳强度低于光滑试样。

图 6.7-35 为在旋转弯曲试验时，钢试样的应力集中敏性系数 q 随温度的变化曲线。图 6.7-36 为高温下碳钢的有效应力集中系数。

(5) 表面状态因素

材料的疲劳强度与表面状态有很大关系。表面粗糙度增加，疲劳强度就降低。各种表面强化工艺，对高温下材料疲劳强度的影响，随着温度的升高而降低。表 6.7-7 为各种加工工艺对镍基合金 GH3032（CrNi77TiAl）试样疲劳寿命的影响。表 6.7-8 为表面喷丸对钴基合金缺口试样疲劳强度的影响，试样为边长 15.2 mm 的方形截面，材料为钴基合金 S-816，进行平面弯曲疲劳试验，缺口为有 60°的 V 形槽，槽深 1.9 mm，槽的根部圆角半径 0.76 mm。将试样先经磨削引入残余拉应力，在经喷丸引入残余正应力。由于槽部磨削引入残余拉应力，使有效应力集中系数 K_σ 在室温下大于 α_σ；喷丸引入残余压应力，使 K_σ 在室温下比 α_σ 值小得多。但随着温度的升高，磨削的有害效应及喷丸的有利效应将逐渐消失。表 6.7-9 为残余压应力对铁基合金疲劳强度的有利影响。

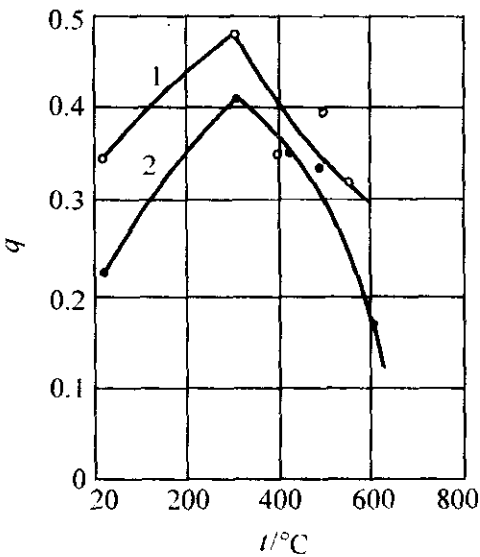


图 6.7-35 钢在高温下的应力集中敏性系数 q

1—1Cr13 钢；2—30CrMo 钢

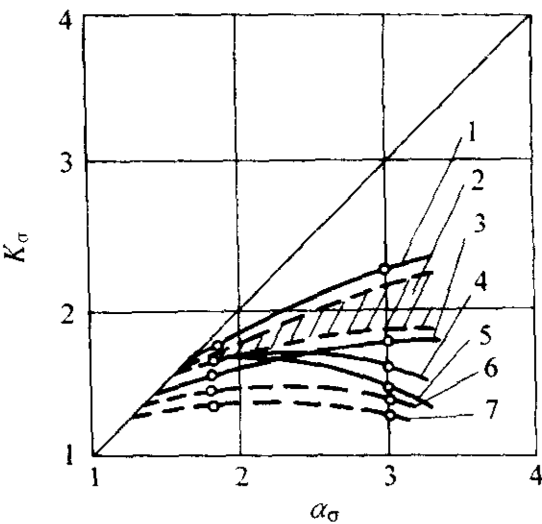


图 6.7-36 高温下碳钢的有效应力集中系数 K_σ

1—$w_C=0.21\%$钢，$f=2\,980\ \text{min}^{-1}$，$t=300℃$；
2—$w_C=0.21\%\sim0.72\%$钢，$f=150\ \text{min}^{-1}$，$t=20℃$；
3—$w_C=0.21\%$钢，$f=2\,980\ \text{min}^{-1}$，$t=500℃$；
4—$w_C=0.72\%$钢，$f=2\,980\ \text{min}^{-1}$，$t=500℃$；
5—$w_C=0.72\%$钢，$f=2\,980\ \text{min}^{-1}$，$t=575℃$；
6—$w_C=0.72\%$钢，$f=150\ \text{min}^{-1}$，$t=575℃$；
7—$w_C=0.21\%$钢，$f=150\ \text{min}^{-1}$，$t=500℃$

表 6.7-7 各种加工工艺对镍基合金 GH3032 试样疲劳寿命的影响

加工工艺	硬层厚度/μm	当 $\sigma_a=412$ MPa 时，到达破坏的循环数 当 20℃时 $N/10^6$	当 20℃时 寿命/%	当 700℃时 $N/10^6$	当 700℃时 寿命/%
电抛光	—	4.85	—	13.4	—
精车	128	2.85	-41	9.01	-34
粗车	185	1.53	-68	5.35	-61
带电车削	91	2.27	-53	7.05	-48
新砂轮磨削	49	3.61	-25	11.7	-13
钝砂轮磨削	37	3.44	-29	10.4	-23
新刀车削后抛光	75	4.28	-11.6	10.0	-26
钝车车削后抛光	139	3.82	-21	8.55	-36
磨削后抛光	37	5.03	+3.7	12.6	-6
辊压	296	7.83	+61	14.3	+6.4
喷丸	189	17.8	+246	15.2	+12.6

注：电抛光试样的寿命设为 100%。

表 6.7-8 喷丸对钴基合金缺口试样疲劳强度的影响

加工工艺	有效应力集中系数 K_σ（$\alpha_\sigma=2.7$，$N=10^8$）		
	室温	482～593℃	649℃
槽部磨削	4.6	2.9	2.4
喷丸	1.3	1.5	1.9

表 6.7-9 表面残余压应力对铁基合金疲劳性能的影响

铁基合金	试样类型	试验温度/℃	残余应力/MPa	σ_{-1}/MPa（$N=10^7$）		σ_{-1}增加率/%
				未喷丸	喷丸	
GH1140	板材 $\alpha_\sigma=1$	550	-1 100	350	460	31
GH2135	缺口 $\alpha_\sigma=2$	450	-950	175	275	57
GH2135	缺口 $\alpha_\sigma=2$	550	-950	240	300	25
GH2036	缺口 $\alpha_\sigma=2$	600	-1 400	≤200	300	≥28
GH2132	缺口 $\alpha_\sigma=2$	650	-1 600	230	255	30

(6) 平均应力因素

平均应力 σ_m 对材料疲劳强度的影响，可用等寿命曲线来表示。在高温疲劳中，随着温度的提高，整个曲线向原点移动，即蠕变强度及疲劳强度都降低。图 6.7-37 为钴基合金 S-816 在室温 24℃及高温下的等寿命曲线，实线为光滑试样，虚线为缺口试样（$\alpha_\sigma=3.4$）。图 6.7-38 为 N-155 合金的等寿命曲线。

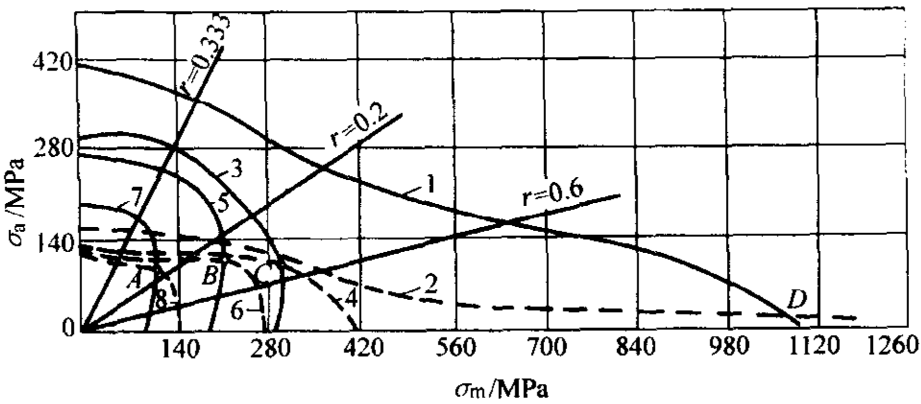


图 6.7-37 钴基合金 S-816，在寿命 100 h 或 2.16×10^7 次循环下，有平均拉应力时的等寿命曲线

1—光滑试样，$t=24$℃；2—缺口试样（$\alpha_\sigma=3.4$），$t=24$℃；3—光滑试样，$t=732$℃；4—缺口试样（$\alpha_\sigma=3.4$），$t=732$℃；5—光滑试样，$t=816$℃；6—缺口试样（$\alpha_\sigma=3.4$），$t=816$℃；7—光滑试样，$t=900$℃；8—缺口试样（$\alpha_\sigma=3.4$），$t=900$℃

A 点—900℃；B 点—816℃；C 点—732℃；D 点—24℃

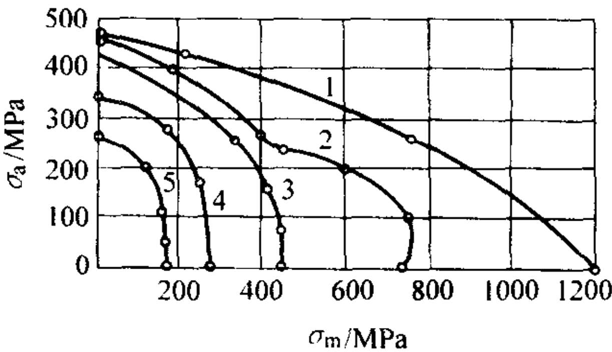


图 6.7-38 N-155 合金光滑试样在 150 h 寿命下有平均应力时的等寿命曲线

1—室温；2—538℃；3—649℃；4—732℃；5—816℃

1.4.4 高温下疲劳强度计算

高温下疲劳强度计算方法有两种：①考虑高温及循环应力（应变）效应的静态计算法；②考虑蠕变与疲劳复合作用的计算法。

(1) 静态计算法

静态计算法是根据长期实践经验，对高温下工作的零部件确定出其许用应力和安全系数。由机械零件的工作温度和受力情况求安全系数有三种情况，即

1) 工作温度下以蠕变极限为基准的安全系数

$$n_{\varepsilon p}^t=\frac{\sigma_\varepsilon^t}{\sigma_p^t} \tag{6.7-4}$$

2) 工作温度下以持久极限为基准的安全系数

$$n_{\tau p}^t=\frac{\sigma_\tau^t}{\sigma_p^t} \tag{6.7-5}$$

3) 工作温度下以屈服强度为基准的安全系数

$$n_{sp}^t=\frac{\sigma_s^t}{\sigma_p^t} \tag{6.7-6}$$

式中，σ_p^t 为在温度 t 时的许用应力；σ_ε^t 为蠕变极限，其中 t 代表温度，ε 为单位小时的变形量；σ_τ^t 为持久极限，其中 t 代表温度，τ 为时间；σ_s^t 为工作温度下的屈服强度。

由于 σ_s^t 的测量比较困难，可以用高温强度极限 σ_b^t 与 σ_s^t 之间的比例关系，近似地算出 σ_s^t 的数值。对于碳钢有

在 300℃时，$\sigma_s^t=0.38\sigma_b^t$；

在 350℃时，$\sigma_s^t=0.36\sigma_b^t$；

在 400℃时，$\sigma_s^t=0.35\sigma_b^t$；

在 450℃时，$\sigma_s^t=0.33\sigma_b^t$；

在 500℃时，$\sigma_s^t=0.25\sigma_b^t$；

对于合金钢，这个比值比碳钢大，在 500℃时，钼钢的比值为 0.34，镍钼钢的比值为 0.42。

涡轮机零部件的安全系数许用值，见表 6.7-10。

表 6.7-10 涡轮机零部件的安全系数许用值

零部件名称	应力状态	材料	安全系数	零部件名称	应力状态	材料	安全系数
涡轮机动叶片	拉、弯、扭疲劳	1Cr13，2Cr13，15Cr11MoV	$n_{sp}^t=1.7\sim2.5$（低温）	涡轮机静叶片	弯、扭疲劳	20，20CrMo，15Cr1Mo1V，12CrMnV，1Cr13	$n_{sp}^t=3$（低温）
			$n_{sp}^t=2$，$n_{\tau p}^t=2$ $n_{\varepsilon p}^t=1.3$（高温）				$n_{sp}^t=3$，$n_{\tau p}^t=2.3$ $n_{\varepsilon p}^t=1.4$（高温）
涡轮机卫带、拉金	拉、弯	—	$n_{sp}^t=1.5\sim2.4$（低温）	涡轮机整体转子	周向应力	34CrMo，34CrNi3Mo，25Cr1Mo1V	$n_{sp}^t=2.2$（低温）
			$n_{sp}^t=2$，$n_{\tau p}^t=2$ $n_{\varepsilon p}^t=1.3$（高温）				$n_{sp}^t=2.2$，$n_{\tau p}^t=1.5$ $n_{\varepsilon p}^t=1.0$（高温）

续表 6.7-10

零部件名称	应力状态	材料	安全系数	零部件名称	应力状态	材料	安全系数
涡轮机焊接转子	周向应力	34CrMo 34CrNi3Mo 25Cr1Mo1V	$n^t_{sp}=2.3$（焊接处） $n^t_{sp}=3$（低温）	涡轮机焊接隔板	弯	ZG20CrMo ZG15Cr1Mo1V	$n^t_{sp}=1.65$，$n^t_{rp}=1.65$ $n^t_{ep}=1.25$（高温）
			$n^t_{sp}=2.2$，$n^t_{rp}=1.65$ $n^t_{ep}=1.25$（高温）	涡轮机机壳（铸件） 涡轮机机壳（锻件）	内压	HT250，Q235A，20，ZG20CrMo，ZG20CrMoV 45，35CrMoA，25Cr2Mo1V	$n^t_{sp}=1.5$（低温）
涡轮机套装叶轮	周向应力	34CrMo 34CrNi3Mo 25Cr1Mo1V	$n^t_{sp}=1.8$（低温）				$n^t_{sp}=2$，$n^t_{rp}=2$ $n^t_{ep}=1.55$（高温）
			$n^t_{sp}=1.8\sim2$，$n^t_{rp}=1.65$ $n^t_{ep}=1.25$（高温）				$n^t_{sp}=1.65$，$n^t_{rp}=1.65$ $n^t_{ep}=1.25$（高温）

注："低温"是指低于蠕变温度；"高温"是指高于蠕变温度。下角 p 为"许用"。

汽轮机用钢的许用应力，见表 6.7-11。

表 6.7-11　汽轮机用钢的许用应力

钢号	在不超过下列温度（℃）的许用应力 σ^t_p/MPa													
	20	200	250	300	350	400	425	450	475	500	525	550	575	600
34CrMo	166	166	166	166	166	156	132	107	83	61				
1Cr13	117	107	103	98	88	83	76	76	67	60	50	37		17
2Cr13	137	127	117	115	112	107	103	95	83	63	44	23		
30Cr2MoV	215	215	215	205	196	181	171	161	127	88	52	30		
33Cr3MoWV	235	235	235	235	225	210	205	181	147	98	63			

动力机械用铸钢的许用应力，见表 6.7-12。

表 6.7-12　动力机械用铸钢的许用应力

钢号	应力种类	在不超过下列温度（℃）的许用应力 σ^t_p/MPa								
		120	200	250	300	350	400	425	450	475
ZG230-450	拉伸	83	83	83	71	62	56	53	43	31
	弯曲	100	100	100	85	75	68	63	52	37
ZG270-500	拉伸	92	92	92	78	68	58	54	44	31
	弯曲	109	109	109	94	82	70	65	53	37

（2）蠕变疲劳复合作用计算法

有些机器的零部件是在高温下受交变载荷的作用，如设备的经常起动和制动，高温压力容器的充气和放气等。对它们进行失效分析时，常常会涉及到蠕变和疲劳交互作用的问题，即同时会产生由于载荷保持不变引起的蠕变损伤和由于交变载荷引起的疲劳损伤。对于这类构件的寿命估算，常常需要考虑零部件所受的载荷、频率和温度等主要因素的影响，工程中应用的有线性累积损伤法、应变幅划分法和频率修正法。

1）线性累积损伤法　该方法是根据 Miner 线性累积损伤理论导出的。假设零部件由于交变载荷引起的疲劳损伤 D_f 和由于载荷保持不变引起的蠕变损伤 D_c 的总和达到临界损伤值 D 时，就会发生失效，即

$$D_f + D_c = D \tag{6.7-7}$$

D_f 和 D_c 可以根据线性累积损伤公式计算，即

$$\left.\begin{aligned} D_f &= \sum_{i=1}^{n}\frac{n_i}{N_i} \\ D_c &= \sum_{j=1}^{m}\frac{t_j}{T_j} \end{aligned}\right\} \tag{6.7-8}$$

式中，n_i 为在应力为 σ_i 时的循环数；N_i 为在应力为 σ_i 时的疲劳失效寿命；t_j 为在应力为 σ_j 时的保持时间，h；T_j 为在应力为 σ_j 时的蠕变失效时间，h。

因此，在设计中应使

$$\sum_{i=1}^{n}\frac{n_i}{N_i} + \sum_{j=1}^{m}\frac{t_j}{T_j} \leqslant D \tag{6.7-9}$$

临界损伤 D 值对于不同的材料和工作温度是不相同的，图 6.7-39 给出了两种不同材料的蠕变-疲劳交互作用累积损伤的实验曲线。在工程中为计算方便，常取 $D=1$，这对于大多数情况是偏于安全的。

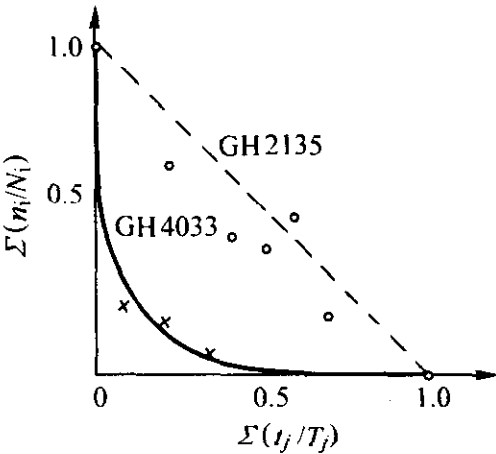


图 6.7-39　蠕变-疲劳复合作用的损伤曲线

2）应变幅划分法　该方法是由曼森（Manson）等人首先提出，认为蠕变-疲劳引起的失效与常温下低周疲劳失效有相类似处，可以用材料的应力-应变循环来估算寿命。并假设在蠕变温度下，零部件的寿命是受应力-应变循环中与时间无关的塑性应变和与时间有关的蠕变应变控制。因此，可对每一个应力-应变滞回环的应变进行划分，分别计算各部分所形成的损伤。

根据高温下可能出现的各种应力-应变循环特性，可以划分成以下四种基本类型。

a）拉伸为塑性应变，压缩为塑性应变，如图 6.7-40a 所示，总应变幅为 $\Delta\varepsilon_{pp}$。

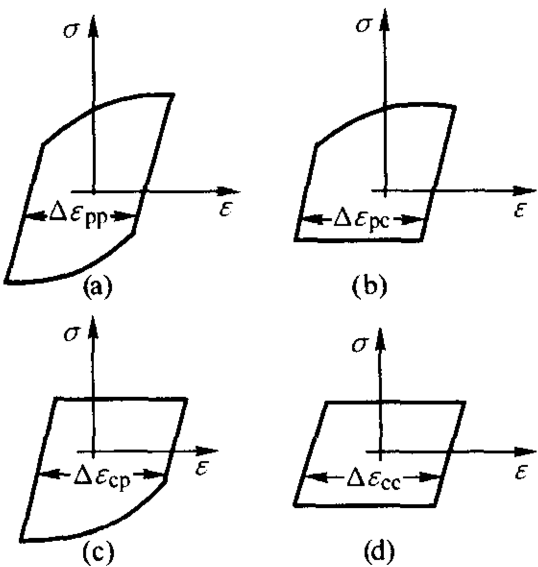


图 6.7-40　四种基本应力-应变循环

b）拉伸为塑性应变，压缩为蠕变应变，如图 6.7-40b 所

示，总应变幅为 $\Delta\varepsilon_{pc}$。

c）拉伸为蠕变应变，压缩为塑性应变，如图 6.7-40c 所示，总应变幅为 $\Delta\varepsilon_{cp}$。

d）拉伸为蠕变应变，压缩为蠕变应变，如图 6.7-40d 所示，总应变幅为 $\Delta\varepsilon_{cc}$。

在每一种基本循环特性下引起的损伤，都符合科芬-曼森（Coffin-Manson）公式，可由试验得到

$$\left.\begin{aligned}\Delta\varepsilon_{pp}N_{pp}^{a_1}&=C_1\\ \Delta\varepsilon_{pc}N_{pc}^{a_2}&=C_2\\ \Delta\varepsilon_{cp}N_{cp}^{a_3}&=C_3\\ \Delta\varepsilon_{cc}N_{cc}^{a_4}&=C_4\end{aligned}\right\} \tag{6.7-10}$$

式中，N_{pp}、N_{pc}、N_{cp}、N_{cc} 分别在应变幅为 $\Delta\varepsilon_{pp}$、$\Delta\varepsilon_{pc}$、$\Delta\varepsilon_{cp}$、$\Delta\varepsilon_{cc}$ 下的失效寿命；a_1、a_2、a_3、a_4 为材料常数；C_1、C_2、C_3、C_4 为材料常数。

对于任意一个应力-应变滞回环，可以按上述四种基本循环进行划分，获得各应变幅分量，如图 6.7-41 所示的一个闭合的应力-应变滞回环，经划分后可得到：

$$\Delta\varepsilon_{pp}=DB,\ \Delta\varepsilon_{cc}=CD,\ \Delta\varepsilon_{pc}=AC-DB$$

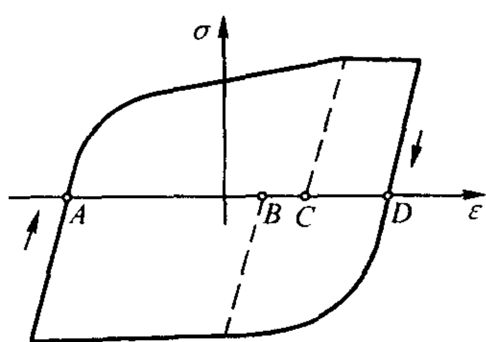


图 6.7-41 应变幅的划分

在一个闭合的应力-应变滞回环中 $\Delta\varepsilon_{pc}$ 和 $\Delta\varepsilon_{cp}$ 不可能同时出现，若拉伸塑性应变大于压缩塑性应变时，仅有 $\Delta\varepsilon_{pc}$ 应变幅分量；反之，仅有 $\Delta\varepsilon_{cp}$ 应变幅分量。

根据应力-应变循环中的各应变幅分量和线性累积损伤假设，现有两种估算失效寿命 N_f 的方法，即

a）方法之一

$$\frac{1}{N_f}=\frac{1}{N_{pp}}+\frac{1}{N_{pc}}+\frac{1}{N_{cp}}+\frac{1}{N_{cc}} \tag{6.7-11}$$

式中的 N_{pp}、N_{pc}、N_{cp} 和 N_{cc} 是分别根据已知的应力幅分量 $\Delta\varepsilon_{pp}$、$\Delta\varepsilon_{pc}$、$\Delta\varepsilon_{cp}$ 和 $\Delta\varepsilon_{cc}$ 由式（6.7-10）计算得到。

b）方法之二

$$\frac{1}{N_f}=\frac{F_{pp}}{N_{pp}}+\frac{F_{pc}}{N_{pc}}+\frac{F_{cp}}{N_{cp}}+\frac{F_{cc}}{N_{cc}} \tag{6.7-12}$$

其中

$$F_{pp}=\frac{\Delta\varepsilon_{pp}}{\Delta\varepsilon},\ F_{pc}=\frac{\Delta\varepsilon_{pc}}{\Delta\varepsilon},$$

$$F_{cp}=\frac{\Delta\varepsilon_{cp}}{\Delta\varepsilon},\ F_{cc}=\frac{\Delta\varepsilon_{cc}}{\Delta\varepsilon}$$

$$\Delta\varepsilon=\Delta\varepsilon_{pp}+\Delta\varepsilon_{pc}+\Delta\varepsilon_{cp}+\Delta\varepsilon_{cc}$$

式中，F_{pp}、F_{pc}、F_{cp}、F_{cc} 为应变幅分数；$\Delta\varepsilon$ 为总应变幅。

式（6.7-12）中的 N_{pp}、N_{pc}、N_{cp} 和 N_{cc} 是以 $\Delta\varepsilon$ 作为应变幅值，代入式（6.7-10）分别计算后得到。

3）频率修正法 频率修正法是对室温下计算疲劳寿命的通用斜率方程用频率项进行修正，以适应高温下具有疲劳和蠕变交互作用的寿命计算。

对于循环中无保持时间的失效寿命 N_f 为

$$\Delta\varepsilon=C_2N_f{}^{-\beta_\nu(1-k)\beta}+\frac{AC_2^n}{E}N_f{}^{-\beta n_\nu k_1+(1-k)\beta n} \tag{6.7-13}$$

式中，$\Delta\varepsilon$ 为总应变幅；E 为弹性模量；ν 为频率；A、C_2、n、β、k、k_1 为常数。

对于循环中具有保持时间情况，需用失效总时间 t_f 表示寿命，若不考虑材料的循环应变硬化，即 $n=1$ 时，t_f 为

$$\Delta\varepsilon=C_{2f}{}^{t-\beta_\tau k\beta}+\frac{AC_2}{E}t_f{}^{-\beta_\tau-(k_1-k\beta)} \tag{6.7-14}$$

$$\tau=\frac{1}{\nu}$$

各符号意义同式（6.7-13）。

2 材料的低温强度

2.1 低温脆断

2.1.1 系列冲击实验与低温脆性

系列冲击实验在材料研究与生产实际中应用较广，因为它比其他实验方法更能灵敏地反映出材料力学性能随内因和外因变化的差异。对某些材料，当冲击实验分别在低温、室温和高温下进行时可以得到一系列冲击吸收功 A_K（或冲击韧度 a_K），将这些冲击值与所对应的实验温度在直角坐标系中标出，然后用光滑曲线将这些实验数据连接起来，可以得到这种材料冲击韧性与温度的关系曲线，即 A_{K-t} 或 a_{K-t}。这种不同温度下的冲击试验称为系列冲击试验。据此可以评定材料的低温脆性。

系列冲击实验证明：体心立方金属及合金或某些密排六方晶体金属及合金，尤其是工程上常用的中、低强度结构钢，当试验温度低于某一温度 t_k 时，材料由韧性状态变为脆性状态，冲击吸收功明显下降，断裂机理由微孔聚集变为穿晶解理，断口特征由纤维状变为结晶状，这就是低温脆性。转变温度 t_k 称为韧脆转变温度或冷脆转变温度。面心立方金属及合金一般没有低温脆性现象，但有实验证明，在 20～42 K 的极低温度下，奥氏体钢及铝合金有冷脆性。高强度钢及超高强度钢在很宽温度范围内冲击吸收功均较低，故韧脆转变不明显。实验中归纳有 3 种不同的冲击吸收功-温度关系曲线。第一类曲线显示材料在很宽的实验温度范围内都是脆性的，如淬火态的高碳马氏体钢。第二类曲线显示具有面心立方结构的金属如铜、铝等材料在很低的温度下仍具有较高的韧性。这类材料的屈服强度对温度和应变速率的变化不敏感。第三类曲线显示材料在一定温度区间产生低温脆性转变，如体心立方金属及其合金、某些密排六方金属及其合金，及许多珠光体-铁素体两相钢。这类材料的屈服强度对温度和应变速率的变化是十分敏感的。

从宏观角度分析，材料低温脆性的产生与其屈服强度 σ_s 和断裂强度 σ_b 随温度的变化有关。断裂强度 σ_b 随温度的变化很小。屈服强度 σ_s 随温度的变化情况与材料的本性有关。具有体心立方或密排六方结构的金属或合金的屈服强度 σ_s 对温度变化十分敏感，温度降低，σ_s 急剧升高，故 σ_s 与 σ_b 随时间变化的两线交于一点，该交点对应的温度即为 t_k。高于 t_k 时，$\sigma_b>\sigma_s$，材料受载后先屈服再断裂，为韧性断裂；低于 t_k 时，外加应力首先达到 σ_b，材料表现为脆性断裂。而面心立方结构材料的 σ_s' 随温度的下降变化不大，故此种材料的脆性断裂现象不明显。

微观上，体心立方金属的低温脆性与位错在晶体中运动的阻力 σ_i 对温度变化非常敏感有关，σ_i 在低温下增加，故该类材料在低温下处于脆性状态。面心立方金属因位错宽度比较大，σ_i 对温度变化不敏感，故一般不显示低温脆性。

体心立方金属的低温脆性还与迟屈服现象有关。即对该

材料施加一大于 σ_s 的高速载荷时材料并不立即产生屈服，而需要经过一段孕育期才开始塑性变形。在孕育期间只产生弹性变形，而没有塑性变形消耗能量，故有利于裂纹的扩展，从而表现为脆性破坏。而具有面心立方结构材料的迟屈服现象不明显，故其低温脆性也不明显。

2.1.2　韧脆转化温度及其评价方法

因为韧性是材料塑性变形和断裂全过程吸收能量的能力，它是强度和塑性的综合表现，因而在特定条件下，能量、强度和塑性都可用来表示韧性。所以，依照试样断裂消耗的功、断裂后塑性变形的大小及断口形貌均可以确定 t_k。目前尚无简单的判据求韧脆转变温度 t_k。通常只是根据能量、塑性变形或断口形貌随温度的变化定义 t_k。为此，需要在不同温度下进行冲击弯曲试验，根据试验结果作出冲击吸收功-温度曲线、试样断裂后塑性变形量和温度的关系曲线、断口形貌中各区所占面积和温度的关系曲线等，根据这些曲线求 t_k。这里只介绍根据能量判据和断口形貌判据定义 t_k 的方法，各种韧脆性转变温度判据见图6.7-42。

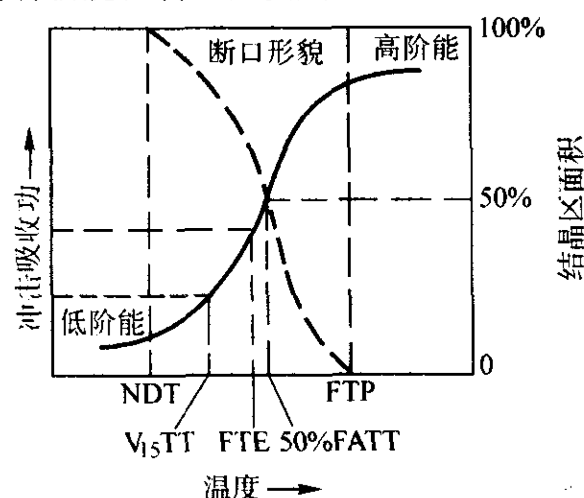


图6.7-42　各种韧脆性转变温度判据

按能量法定义 t_k 的方法有如下几种。

1）当低于某一温度材料吸收的冲击能量基本不随温度而变化，形成一平台，该能量称为“低阶能”。以低阶能开始上升的温度定义 t_k，并记为NDT（nilductility temperature），称为无塑性或零塑性转变温度。这是无预先塑性变形断裂对应的温度，是最易确定 t_k 的判据。在NDT以下，断口由100%结晶区（解理区）组成。

2）高于某一温度材料吸收的能量也基本不变，形成一个上平台，称为“高阶能”。以高阶能对应的温度为 t_k，记为FTP（fracture transition plastic）。高于FTP的断裂，将得到100%的纤维状断口。显然，这是一种最保守定义 t_k 的方法。

3）以低阶能和高阶能平均值对应的温度定义，并记为FTE（fracture transition elastic）。

4）冲击试样冲断后，其断口形貌如图6.7-43所示。如同拉伸试验一样，冲击试样断口也有纤维区、放射区（结晶区）和剪切唇几部分，但在不同试验温度下，3个区之间的相对面积是不同的。温度下降，纤维区面积突然减少，结晶区面积突然增大，材料由韧变脆。通常取结晶区面积占整个断口面积50%时的温度为 t_k，并记为50%FATT（fracture appearance transition temperature）或 $FATT_{50}$、t_{50}。

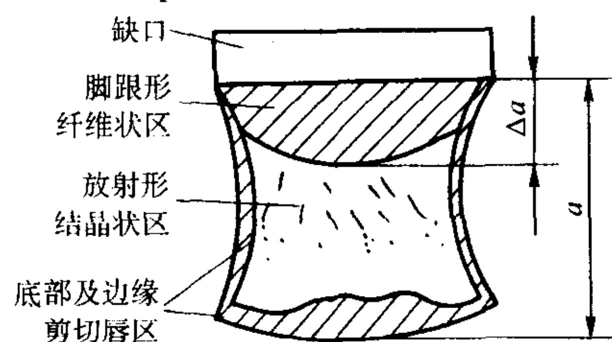


图6.7-43　冲击断口形貌示意图

50%FATT反映了裂纹扩展变化特征，可以定性地评定材料在裂纹扩展过程中吸收能量的能力。实验发现，50%FATT与断裂韧度 K_{IC} 开始急速降低的温度有较好的对应关系，故得到广泛应用。但此方法需要目测评定各区所占面积，受人为影响较大，要求测试人员有较丰富的经验。

韧脆转变温度 t_k 反映了温度对韧脆性的影响，它与 δ、ψ、A_K、NSR一样，也是安全性指标。t_k 是从韧性角度选材的重要依据之一，可用于抗脆断设计，但不能直接用来设计计算机件的承载能力或截面尺寸。对于在低温服役的机件，依据材料的 t_k 值可以直接或间接地估计它们的最低使用温度。显然，机件的最低使用温度必须高于 t_k，两者之差愈大愈安全。为此，选用的材料应该具有一定的韧性温度储备 Δ（$\Delta = t_0 - t_k$，t_0 为材料使用温度），Δ 值常取20～60℃。对于受冲击载荷的重要机件 Δ 值取上限；不受冲击载荷作用的非重要机件 Δ 值取下限。

上述表明，由于定义 t_k 的方法不同，同一材料所得 t_k 亦有差异；同一材料，使用同一定义方法，由于外界因素（如试样尺寸、缺口尖锐度和加载速率等）的改变，t_k 也要变化。所以，在一定条件下用试样测得的 t_k，因为和实际结构工况之间无直接联系，不能说明该材料制成的机件一定在该温度下断裂。

2.1.3　影响材料低温脆性的因素

（1）晶体结构的影响

体心立方金属及其合金存在低温脆性，面心立方金属及其合金一般不存在低温脆性。体心立方金属的低温脆性可能和迟屈服现象有密切关系。所谓迟屈服是指当用高于材料屈服极限的载荷以高加载速度作用于体心立方结构材料时，瞬间并不屈服，需在该应力下保持一定时间后才发生屈服。且温度越低，持续的时间越长，这就为裂纹的发生和传播创造有利条件。中、低强度钢的基体是体心立方结构的铁素体，故都有明显的低温脆性。

（2）化学成分的影响

间隙溶质元素含量增加，高阶能下降，韧脆转变温度提高。这是由于间隙溶质元素溶入基体金属晶格中，通过与位错的交互作用偏聚于位错线附近形成柯氏气团，既增加 σ_i，又使 k_y 增加，致使 σ_s 升高，所以钢的脆性增大。

钢中加入置换型溶质元素（Ni、Mn例外）一般也降低高阶能，提高韧脆转变温度，但这种影响较间隙溶质原子小得多。图6.7-44为合金元素对钢的韧脆转变温度的影响。置换型溶质元素对韧脆转变温度的影响与 σ_i、k_y 及 γ_s 的变化有关。Ni减小低温时的 σ_i 和 k_y，故韧性提高。另外，Ni还增加层错能，促进低温时螺位错交滑移。使裂纹扩展消耗功增加，故韧性增加。若置换型溶质元素降低层错能，促进位错扩展或形成孪晶，使螺位错交滑移困难，则钢的韧性下降。

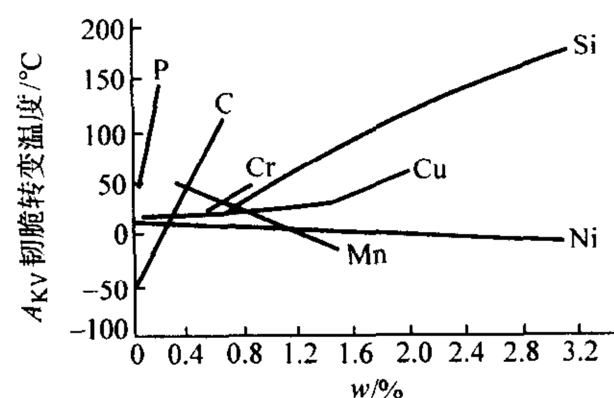


图6.7-44　合金元素对钢的韧性转变温度的影响

杂质元素S、P、Pb、Sn、As等使钢的韧性下降。这是由于它们偏聚于晶界，降低晶界表面能，产生沿晶脆性断

裂，同时降低脆断应力所致。

(3) 显微组织的影响

1) 晶粒大小　细化晶粒可使材料韧性增加。图 6.7-45 为铁素体晶粒直径与韧脆转变温度的关系，这表明韧脆转变温度与 $d^{1/2}$ 成线性关系。研究表明，不仅铁素体晶粒大小与韧脆转变温度呈线性关系，而且马氏体板条束宽度、上贝氏体铁素体条束宽度、原始奥氏体晶粒尺寸和韧脆转变温度之间也呈线性关系。还有人提出，减小亚晶粒和胞状结构尺寸也能提高材料的韧性。

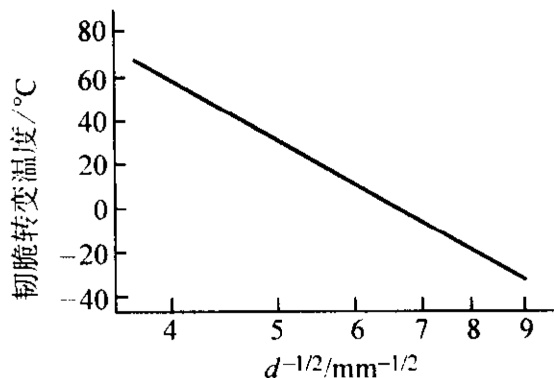


图 6.7-45　韧脆转变温度和铁素体晶粒直径的关系

细化晶粒提高韧性的原因有：晶界是裂纹扩展的阻力；晶界前塞积的位错数减少，有利于降低应力集中；晶界总面积增加，使晶界上杂质浓度减少，避免产生沿晶脆性断裂。

2) 金相组织　在较低强度水平，强度相同而组织不同的钢，其冲击吸收功和韧脆转变温度以回火索氏体最佳，贝氏体回火组织次之，片状珠光体组织最差。此外，球化处理能改善钢的韧性。

在较高强度水平时，中、高碳钢经等温淬火获得下贝氏体组织，其冲击吸收功和韧脆转变温度优于同强度的淬火马氏体并回火的组织。

在相同强度水平，典型上贝氏体的韧脆转变温度高于下贝氏体的韧脆转变温度。但低碳钢低温上贝氏体（B_1）的韧性却高于回火马氏体的韧性，这是由于在低温上贝氏体中渗碳体沿奥氏体晶界的析出受到抑制，减少了晶界裂纹所致。

在低碳合金钢中，经不完全等温处理获得贝氏体和马氏体的混合组织，其韧性比单一马氏体或单一贝氏体组织要好。这是因为贝氏体先于马氏体形成，优先将奥氏体晶粒分割成几部分，使随后形成的马氏体限制在较小范围内，从而获得了极为细小的混合组织。裂纹在此种组织内扩展要多次改变方向，消耗能量较大，故钢的韧性较高。关于中碳合金钢马氏体-贝氏体混合组织的韧性，亦视钢在奥氏体化后的冷却过程中贝氏体和马氏体的形成顺序而定，只有贝氏体先于马氏体形成韧性才可以改善。

在马氏体钢中存在稳定残余奥氏体，可以抑制解理断裂，从而显著改善钢的韧性。马氏体钢中的残余奥氏体膜也有类似作用。

钢中碳化物及夹杂物等第二相对钢的脆性的影响程度取决于第二相质点的大小、形状、分布、第二相性质及其与基体的结合力等因素。一般第二相尺寸增加，材料的韧性下降，韧脆转化温度升高。第二相的形状对材料脆性也有影响，球状第二相材料的韧性较好。

(4) 温度的影响

碳钢和某些合金钢在冲击载荷或静载荷作用下，在一定温度范围内出现脆性。因为在该温度范围内加热钢时，表面氧化色为蓝色，故此现象称为蓝脆。但二者的蓝脆温度范围不同。在静拉伸时，蓝脆的温度范围为 230～370℃；在冲击载荷作用下，蓝脆最严重的温度范围为 525～550℃。

蓝脆是形变时效加速进行的结果，当温度升至某一适当温度时，碳、氮原子扩散速率加快，易于在位错附近偏聚形成柯氏气团。显然，若这一过程的形成速率高于塑性变形速率，则在塑性变形过程中产生时效，使材料强度提高，塑性下降；反之，则材料的塑性提高。在冲击载荷作用下，形变速率较高，碳、氮原子需要在较高温度下才能获得足够的扩散激活能，以形成柯氏气团，故蓝脆温度升高。

(5) 加载速率的影响

提高加载速率如同降低温度，使材料脆性增大，韧脆转变温度提高。

加载速率对钢脆性的影响与钢的强度水平有关。一般，中、低强度钢的韧脆转变温度对加载速率比较敏感，而高强度钢和超高强度钢的韧脆转变温度则对加载速率的敏感性较小。

(6) 试样形状和尺寸的影响

缺口曲率半径越小，t_k 越高，因此，V 形缺口试样的 t_k 高于 U 形试样的 t_k。

当不改变缺口尺寸而只增加试样宽度（或厚度）时，t_k 升高。若试样各部分尺寸按比例增加时，t_k 也升高。这是由于试样尺寸增加时应力状态变硬，且缺陷概率增大，故脆性增大。

2.2　低温疲劳

2.2.1　低温下材料的疲劳极限

表 6.7-13 和表 6.7-14 为低温下材料疲劳极限数据。表 6.7-15 是将各种材料在低温下的疲劳极限处理后得到的平均

表 6.7-13　温度对钢静强度和疲劳极限的影响　MPa

钢种	材料情况	试样	+20℃			-75℃			-183℃		
			σ_b	σ_s	σ_{-1}	σ_b	σ_s	σ_{-1}	σ_b	σ_s	σ_{-1}
w_C = 0.15% 碳钢	正火	光试样	430	315	221	543	437	—	778	718	495
	正火	缺口试样	589	374	166	698	542	210	749	749	294
	粗晶粒	光试样	357	155	166	435	277	—	666	647	—
	粗晶粒	缺口试样	469	221	140	506	357	191	605	605	240
Cr4Ni 钢	商品	光试样	761	585	388	888	680	416	1 106	944	549
	商品	缺口试样	1 022	773	241	1 161	1 011	248	1 106	1 106	274
GCr15 钢	淬火回火	光试样	—	—	828	—	—	818	—	—	—

表 6.7-14　材料的低温疲劳极限

材料	试验循环数 N	疲劳极限 σ_{-1}/MPa					
		20℃	-40℃	-78℃	-188℃	-253℃	-269℃
铜	10^6	98	—	—	142	235	255
黄铜	5×10^7	171	181	—	—	—	—
铸铁	5×10^7	58	73	—	—	—	—
软钢	10^7	181	—	250	559	—	—
碳钢	10^7	225	—	284	612	—	—
镍铬钢	10^7	529	—	568	750	—	—
硬铝	5×10^7	112	142	—	—	—	—
铝合金 2A14	10^7	98	—	—	166	304	—
铝合金 2A11	10^7	122	—	—	152	274	—
铝合金 7A09	10^7	83	—	—	137	235	—

表 6.7-15 低温下金属的疲劳极限比值

材料	低温下的疲劳极限/室温下的疲劳极限（平均值）			缺口试样低温下的疲劳极限/缺口试样室温下的疲劳极限（平均值）		光试样的疲劳极限/光试样的强度极限（平均值）			
	-40℃	-78℃	-186～-196℃	-78℃	-186～-196℃	室温	-40℃	-78℃	-186～-196℃
碳钢	1.20	1.30	2.57	1.10	1.47	0.43	0.47	0.45	0.67
合金钢	1.06	1.13	1.61	1.06	1.23	0.48	0.51	0.48	0.58
合金铸铁	—	1.22	—	1.05	—	0.27	—	0.27	—
不锈钢	1.15	1.21	1.54	—	—	0.52	0.50	0.57	0.59
铝合金	1.14	1.16	1.69	—	1.35	0.42	—	0.46	0.59
钛合金	—	1.11	1.40	1.22	1.41	0.70	—	0.63	0.54

值，表中大多数的数据是在循环数 $N=10^6$ 次循环下试验得到。

2.2.2 低温下材料的 $S-N$ 曲线

图 6.7-46 为温度对铝合金及钢的疲劳极限的影响曲线。图 6.7-47 为 300 K（用符号○表示）以及在 78 K（液态氮，用符号●表示）和 4 K（液态氦，用符号△表示）低温下 5 种材料测得的 $S-N$ 曲线。

2.2.3 低温对应力集中的影响

表 6.7-16 为材料在低温下有效应力集中系数。图 6.7-48 及图 6.7-49 为金属在低温下有效应力集中系数。图 6.7-50 为钢的光滑试样与缺口试样在低温和室温下疲劳极限均值的比值。

2.2.4 低温疲劳强度计算

当温度低于有转折点的温度时，解理断裂会导致疲劳裂纹扩展速度急剧加快。当断裂韧性在低温下大大降低时，裂纹形成寿命可能占有几乎整个低温疲劳寿命。在低温疲劳强度的设计中，一般用室温下疲劳强度设计方法和数据。但必须注意，设计在低温下工作的零件，应避免有尖锐的缺口、裂纹和表面划痕等缺陷，对于短寿命零件，更应避免。

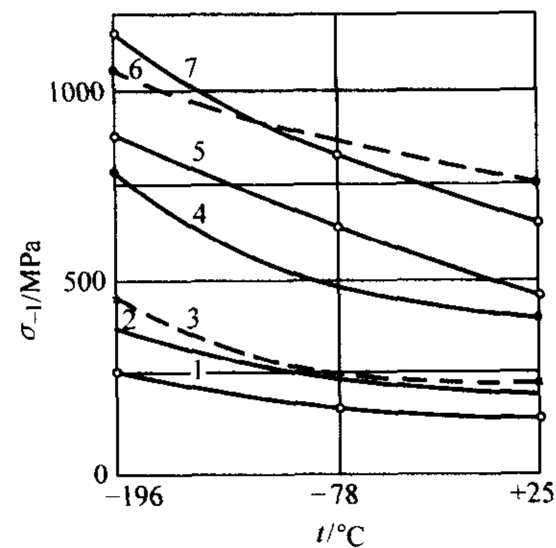


图 6.7-46 温度对铝合金及钢的疲劳极限的影响（$N=10^6$）

各合金及化学成分（质量分数,%）如下：1—铝合金（Mg1.0，Cu0.25，Si0.6，Cr0.25）；2—铝合金（Mn0.6，Mg1.5，Cu4.5）；3—铝合金（Mg2.5，Cu1.6，Cr0.3，Zn5.6）；4—合金钢（C0.3，Mn0.7，Ni3.5）；5—合金钢（C0.3，Mn0.8，Si0.3，Ni0.6，Cr0.53，Mo0.18）；6—合金钢（C0.07，Cr17，Ni6.5，Ti0.37，Al0.12）；7—18/8 奥氏体钢（Cr18，Ni8）

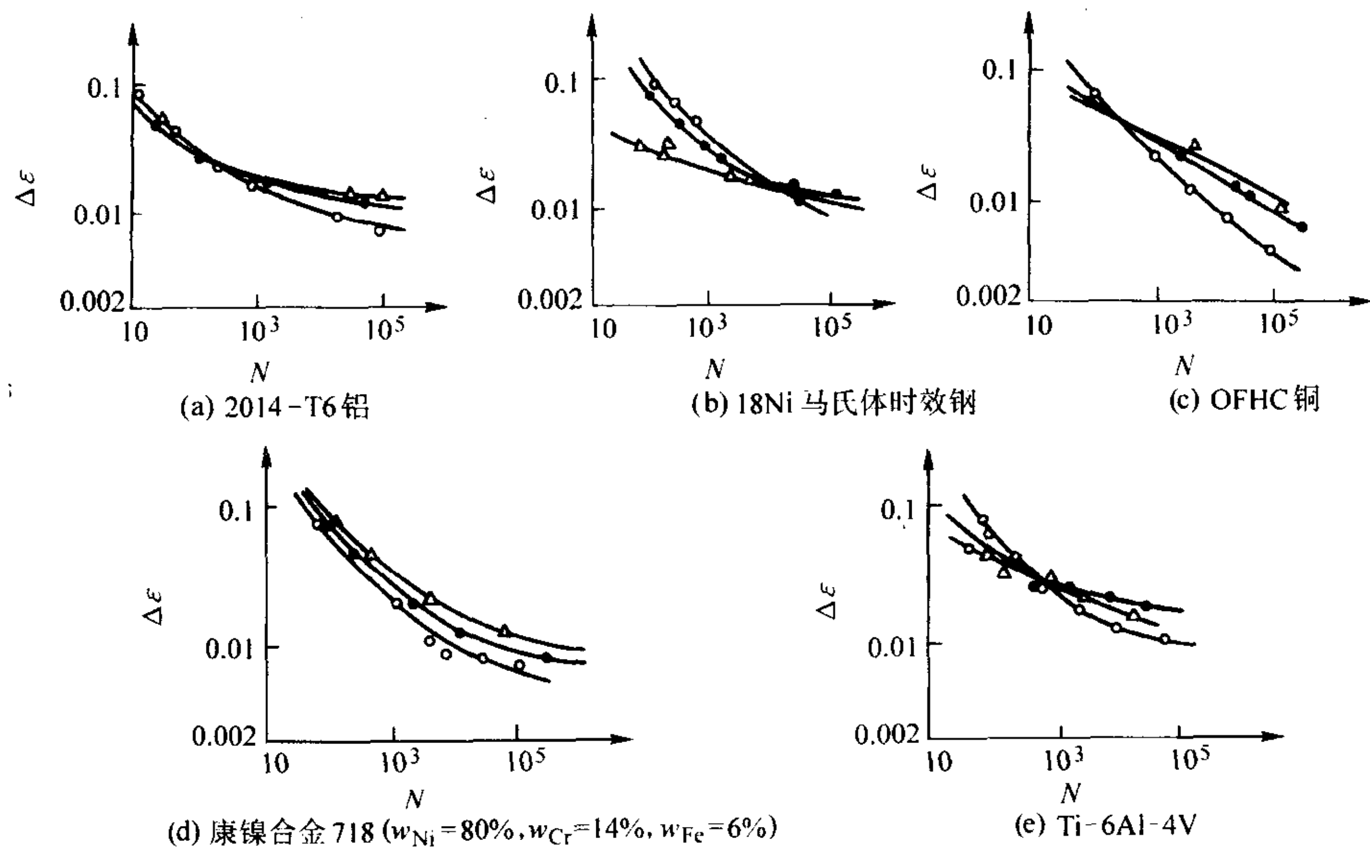


图 6.7-47 低温对低周疲劳的影响

试验温度：○ 300 K（室温）；● 78 K（液氮）；△ 4 K（液氦）

表 6.7-16　材料在低温下的有效应力集中系数

材料	有效应力集中系数 K_σ					
	试验循环数 $N=10^4$		试验循环数 $N=10^5$		试验循环数 $N=10^7$	
	20℃	-196℃	20℃	-196℃	20℃	-196℃
镍钢（500℃回火）	1.16	2.04	1.59	3.42	4.26	3.12
低合金钢	1.09	2.27	1.36	2.46	2.33	3.58
18/8 不锈钢	1.64	2.31	2.61	3.62	4.77	3.86
镍铬钢（650℃回火）	1.09	1.93	1.55	3.0	3.68	5.76
镍铬钢（440℃回火）	1.63	3.4	2.44	3.7	1.82	3.35
钛合金	1.51	1.73	1.55	1.7	2.68	2.5
铝合金 2A12	1.32	1.74	1.42	1.9	2.28	2.24
铝合金 7A09	1.55	2.0	1.51	2.17	2.0	2.78
镁合金	1.31	1.75	1.7	1.95	2.41	2.5

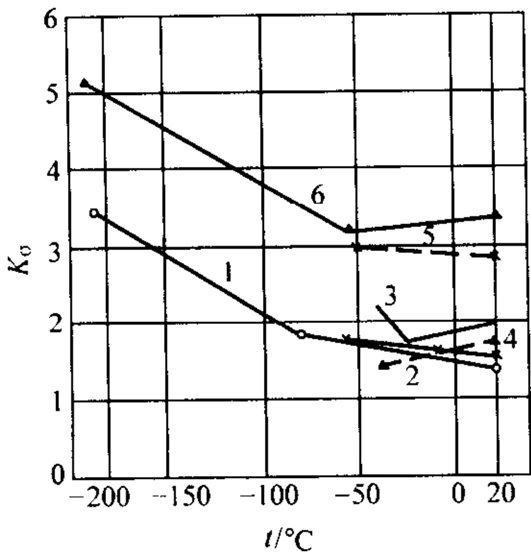


图 6.7-48　碳钢在低温下的有效应力集中系数

1—低碳钢（$w_C=0.08\%$）的拉压疲劳；
2—低碳钢（$w_C=0.08\%$）的旋转弯曲疲劳；
3—中碳钢（$w_C=0.6\%$）的旋转弯曲疲劳；
4—焊接结构轧材，$\sigma_b=402$ MPa，$\alpha_\sigma=2$，钢的拉压疲劳；
5—焊接结构轧材，$\alpha_\sigma=4$ 加拉压疲劳；
6—焊接结构轧材，$\alpha_\sigma=5.6$ 加拉压疲劳

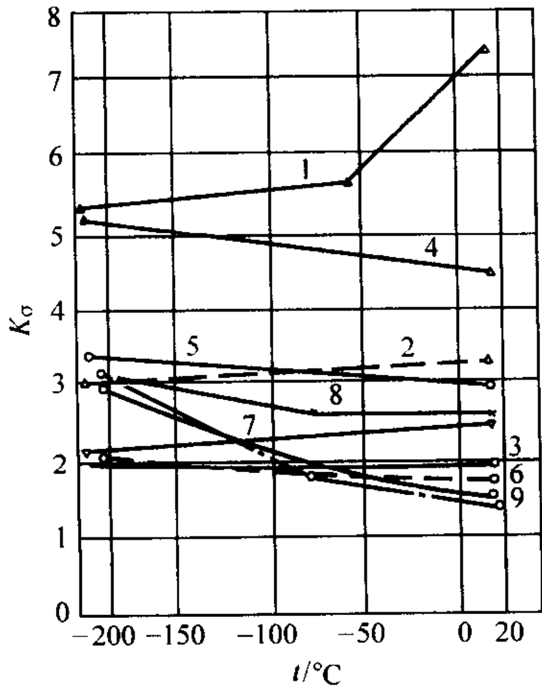


图 6.7-49　材料在低温下的有效应力集中系数

1—耐腐蚀铝合金，$\alpha_\sigma=6$，拉压疲劳；2—耐腐蚀铝合金，$\alpha_\sigma=4$；3—耐腐蚀铝合金，$\alpha_\sigma=2$；4—镍钢（$w_{Ni}=9\%$），$\alpha_\sigma=6$，拉压疲劳；5—镍钢（$w_{Ni}=9\%$），$\alpha_\sigma=4$；6—镍钢（$w_{Ni}=9\%$），$\alpha_\sigma=2$；7—不锈钢酸钢，拉压疲劳；8—铬钼钢（$w_{Cr}=0.83\%$，$w_{Mo}=0.22\%$），拉压疲劳；9—60 钢，拉压疲劳

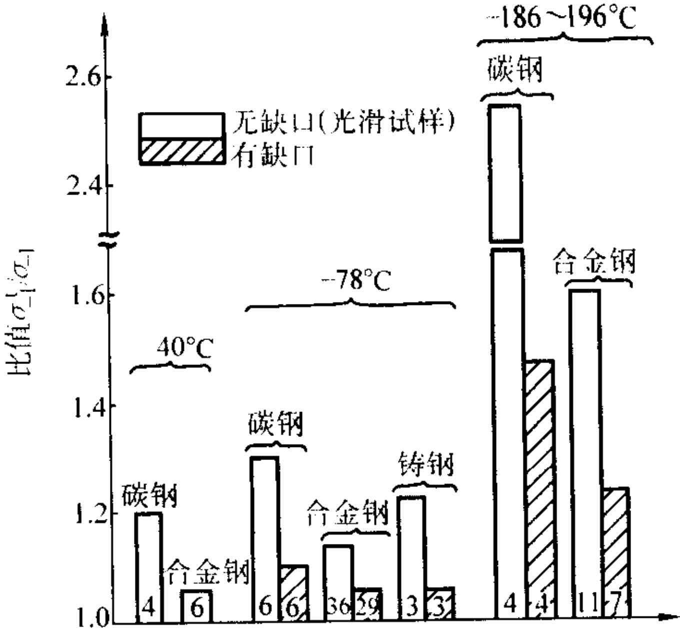


图 6.7-50　无缺口钢（光滑试样）和有缺口钢（缺口试样）在低温下的疲劳极限与在室温下的疲劳极限的均值的比较
（各纵行底部示出所用材料种类及数目）

3　材料热疲劳强度

3.1　热应力与热疲劳

产生热应力情况主要有两种：①零件的热胀冷缩受到固持零件的外加约束而产生热应力；②虽然没有外加约束，但零件各部位温度不一致，存在着温度梯度，导致各部位热胀冷缩不一致而产生热应力。

当热应变越过弹性极限时，热应力与热应变就不成线性关系，此时求解热应力就要按弹塑性关系处理。

由于温度循环变化产生循环热应力所导致的材料或零件的疲劳称为热疲劳。例如热作模具，热轧机的轧辊，热交换管子和锅炉管子等都能产生热疲劳裂纹。

影响热应力大小的因素如下。

1）热应力的大小与线胀系数成正比，线胀系数越大，热应力越大。所以在选材时要考虑线胀系数，在进行机械零部件的配合和焊接时应考虑材料的匹配，即不同材料线胀系数的差别不能太大。例如，铁素体钢和奥氏体钢焊接在一起构成的管道容易破裂，就是因为两者的线胀系数值相差较大，因此所产生的热应力也较大，在多次循环作用下就会引起破坏。又如材料基体中若含有线胀系数不同的第二相，则在温度循环作用下，也会产生局部热应力而引起开裂。

2）在相同的热应变条件下，材料的弹性模量越大，热应力就越大。

3）温度循环变化越大，即上下限温差越大，则热应力就越大。

4）材料的导热系数越低，则快速加热或冷却过程中，温度梯度越陡，热应力也越大。

3.2　热疲劳强度计算

3.2.1　最大温度 - 寿命曲线

对于一般选材及在提高热疲劳强度的材料工艺研究中，可采用 $T_{max}-N$ 曲线，这里 T_{max} 为一个循环的最大温度。图 6.7-51 为某些耐热材料的 $T_{max}-N$ 曲线，在 N 的对数坐标中为直线，可写出下面公式，即

$$T_{max}=A-n\lg N \qquad (6.7\text{-}15)$$

式中，N 为到达破坏的循环数；n 和 A 为材料常数。常数 n 和 A 的取值见表 6.7-17。

3.2.2　应变幅度 - 寿命曲线

对于零件寿命计算需定量给出应变幅度（或应力幅度）与寿命的关系，通常是 $\Delta\epsilon-N$ 曲线。在变温下获得的 $\Delta\epsilon-N$ 曲线称为热疲劳曲线，在双对数坐标中为直线形式。科

芬-曼森提出的热疲劳公式为

$$\Delta\varepsilon_p N^K = C \qquad (6.7\text{-}16)$$

式中，$\Delta\varepsilon_p$ 为塑性应变幅度；K、C 为材料常数。

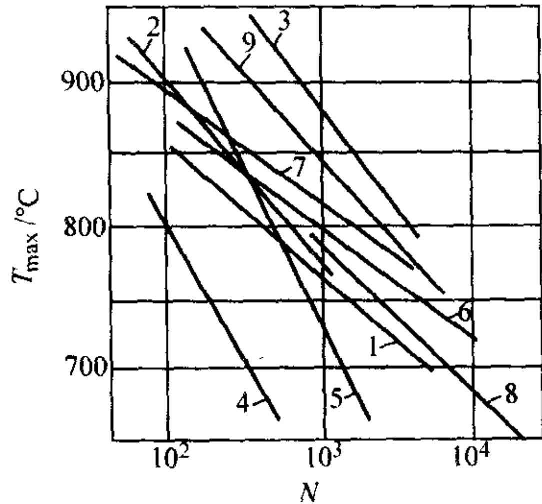


图 6.7-51 几种变形合金的热疲劳曲线

（图中 1～9 编号，对应于表 6.7-17 中材料的序号）

表 6.7-17 式（6.7-15）中常数 *A* 和 *n* 值

序号	材 料	*A*	*n*
1	Nimonic75（镍铬钛耐热合金）	1 400	235
2	Nimonic90（镍铬钛耐热合金）	1 350	150
3	12Cr18Ni9Ti	1 320	250
4	CrNi56WMoCoAl	1 300	150
5	CrNi62WMoCoAl	1 130	100
6	S-816	1 110	109
7	CrNi77TiAlB	1 010	89
8	Inconel550（镍铬铁耐热合金）	1 000	73
9	CrNi70	935	71

对于变温情况，应变的弹性部分必须考虑：

$$\Delta\varepsilon N^K = C \qquad (6.7\text{-}17)$$

式中，$\Delta\varepsilon$ 为总应变幅度，$\Delta\varepsilon = \Delta\varepsilon_e + \Delta\varepsilon_p$；$K$、$C$ 见表 6.7-18 和表 6.7-19。

表 6.7-18 式（6.7-17）中的常数 *K* 和 *C* 值

序号	材料	T_{max}/℃	*K*	*C*
1	CrNi77TiAlB	750	0.825	1 072
		800	0.918	1 096
		850	0.526	48
2	CrNi70WMoTiAl	800	0.875	1 175
		850	0.936	807
		900	1.68	8 260
3	CrNi60WTi	800	0.554	46
		900	0.874	129
4	CrNi62NbMoCoTiAl	800	0.468	56
5	37Cr12Ni8Mn8MoVNb	700	0.414	25.1
6	12Cr18Ni9Ti	700	0.56	57.5
		750	0.60	56.3
		800	0.82	162

表 6.7-19 式（6.7-17）中的常数 *K* 和 *C* 值

序号	材料	T_{max}/℃	*K*	*C*
1	ЖC6y	850	0.247	11.6
		950	0.215	6.2
		950	0.148	4.5
		1 050	0.635	58
2	ЖC6y	1 050	0.688	82
3	ЖC6ф	1 050	0.314	10.7
4	ВЖЛ12y	1 050	0.400	32.1

续表 6.7-19

序号	材料	T_{max}/℃	*K*	*C*
5	ВЖЛ12y	950	0.338	19.5
		1 050	0.125	2.9
		950	0.745	686
			0.625	123
			0.388	20.6
	ВЖЛ12y	950	0.723	170
			0.400	13.7
			0.756	525
6	CrNi62WMoCoAl	850	0.69	189
		900	0.878	475

对某些变形合金和钢进行了热疲劳试验，结果如图 6.7-52 所示。

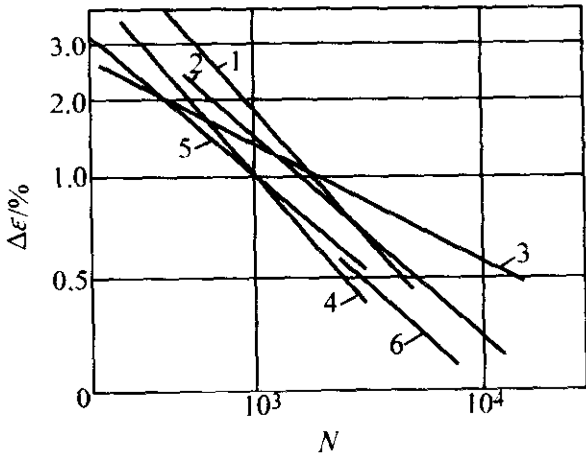


图 6.7-52 几种变形合金的热疲劳曲线

（图中 1～6 编号对应于表 6.7-18 中材料排列序号）

几种铸造合金的热疲劳试验结果，如图 6.7-53 所示。

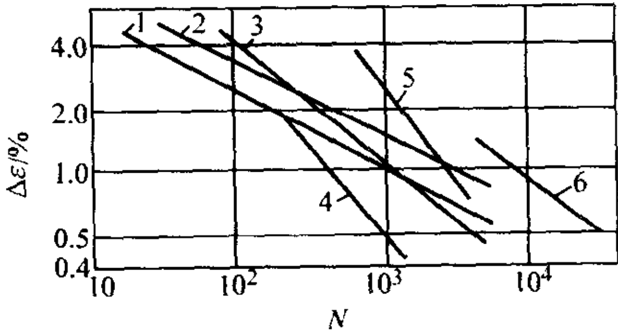


图 6.7-53 几种铸造合金的热疲劳曲线

（图中 1～6 编号对应于表 6.7-19 中材料排列序号）

上面的试验数据，都是在 T_{max} 无保持时间下进行的，而在实际的机器运转中，要有不同的运转的时间（保持时间），特别是民用机器启动一次运转的时间较长，所以要研究带有保持时间情况下的寿命。由于在高温下保持时间内，材料产生蠕变应变（ε_c）的累积损伤，可用下式表示，即

$$\Delta\varepsilon_{e+p+c} N^{K_1} = C_1 \qquad (6.7\text{-}18)$$

一种变形合金 CrNi77TiAlB 的试验结果如图 6.7-54 所示。式（6.7-18）中的 K_1 和 C_1 常数列于表 6.7-20 中。

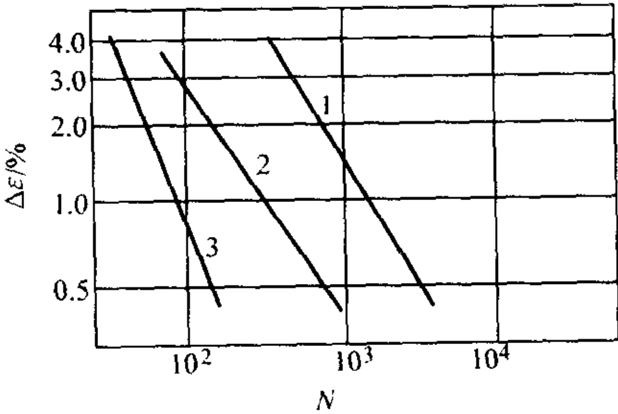


图 6.7-54 CrNi77TiAlB 合金的热疲劳曲线

1—保持时间 = 0；2—保持时间 = 1.5 s；3—保持时间 = 10.7 s

表 6.7-20 式 (6.7-18) 中的 K_1 和 C_1 值

材料	T_{max}/℃	保持时间/s	K_1	C_1
CrNi77TiAlB	800	0	0.918	1 096
		1.5	0.725	76
		10.7	1.66	1 807

图 6.7-55 为三种合金钢的 $\Delta\varepsilon_p-N$ 曲线。

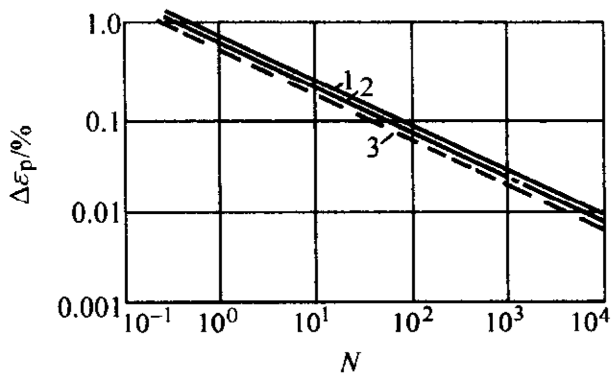


图 6.7-55 三种合金钢的塑性应变幅度与寿命的关系
1—18/8Cr - Ni 钢；2—w_{Cr} = 13%钢；3—Cr - Mo 钢

3.2.3 热疲劳强度设计要考虑的主要问题

热疲劳的寿命估算，可用应变 - 寿命曲线直接得出。但由于影响热疲劳强度的因素很多，而 $\Delta\varepsilon-N$ 曲线可提供给设计应用很少，所以，在热疲劳强度设计中，首先要考虑下列问题。

1）合理选用材料，线胀系数小，导热系数大的材料，对于降低热应力是有效的。而高温持久极限高、韧性好的材料，能提高热疲劳强度。

2）注意结构设计中的问题

① 结构要设计得富有伸缩性。

② 尽量避免有应力集中的结构。

③ 设计焊接结构时，要特别注意坡口形状和背面錾平方法。

④ 焊接不同金属时，采用线胀系数差别小的金属。

4 材料在腐蚀介质中的强度

4.1 应力腐蚀开裂

在腐蚀和静拉应力联合作用下出现的脆性断裂，称为应力腐蚀开裂。应力腐蚀开裂常用英文的三个字头“SCC”表示。不论是韧性材料还是脆性材料，都可能发生应力腐蚀断裂。

4.1.1 应力腐蚀的特征

应力腐蚀断裂一般都是在特定的条件下发生的，这些条件可以归纳如下。

1）只有在拉应力作用下才会发生应力腐蚀断裂。这种拉应力可以是外加载荷引起的，也可以是残余应力，如焊接残余应力、热处理残余应力和装配应力等。一般情况下，发生应力腐蚀时的拉应力都很低，如果没有腐蚀介质的联合作用，零件可以在该应力水平下长期工作而不发生断裂。

2）发生应力腐蚀的环境总是存在腐蚀介质。这种腐蚀介质一般都很弱，如果没有拉应力同时作用，材料在这种介质下腐蚀的速度很慢。产生应力腐蚀的介质一般都是特定的，也就是说，每种材料只对某些介质敏感，而这些介质对其他材料可能没有明显作用。如黄铜在氨气中以及不锈钢在具有氯离子的腐蚀介质中容易发生应力腐蚀，但反过来不锈钢对氨气、黄铜对具有氯离子的介质就不敏感。

3）一般只有合金才产生应力腐蚀，纯金属不会发生这种现象。合金也只在拉应力与特定的腐蚀介质联合作用下才会发生应力腐蚀。

应力腐蚀也是经裂纹形成和扩展这些过程进行的。一般认为，裂纹形成约占全部寿命的 90%左右，而裂纹扩展仅占 10%左右。

应力腐蚀断裂可以是沿晶断裂，也可以是穿晶断裂，这取决于合金成分和腐蚀介质。在一般情况下，低碳钢、普通低合金钢、铝合金和 α 黄铜都是沿晶断裂，其裂纹大致沿垂直于拉应力轴的晶界向材料深处扩展。β 黄铜和暴露在氯化物中的奥氏体不锈钢，在大多数情况下是穿晶断裂。

应力腐蚀的断口的宏观形貌属于脆性断裂，有时带有少量塑性撕裂的痕迹。裂纹源可能有几个，但常常是在垂直于主应力的面上的那个裂纹源才引起断裂。由于介质的腐蚀作用，裂纹源及扩展区常呈黑色或灰黑色，脆断区的断口常有放射花样或人字纹。典型的应力腐蚀断裂断口的微观形态一般为沿晶断裂形态，晶面上有撕裂脊。

4.1.2 断裂力学在应力腐蚀中的应用

常规的应力腐蚀研究通常采用光滑试样，根据在应力和腐蚀介质同时作用下发生的延滞断裂的时间，来判断材料抗应力腐蚀性能的好坏。这虽然也能得到材料的临界应力，供工程设计应用，但裂纹形成和裂纹扩展两部分时间却无法分开。实际零件不可避免地存在着裂纹，因此，常规的研究方法不能反映带裂纹的金属材料抗应力腐蚀的性能。

断裂力学的观点认为，裂纹尖端的应力场也可以用应力强度因子来描述。受应力腐蚀作用的材料也存在着一个临界应力强度因子。由于腐蚀介质的作用，其值很小。又由于应力腐蚀断裂是一种与时间有关的延滞性断裂，所以可以用裂纹扩展速度 da/dt 来描述应力腐蚀裂纹的扩展。

实践证明，在拉应力和腐蚀介质同时作用下，材料发生延滞断裂的时间 t 与应力强度因子 K_I 之间有如图 6.7-56 所示的关系。当裂纹尖端的应力强度因子 $K_I=K_{IC}$时，立刻发生断裂，此时的 $t=0$；当 $K_I<K_{IC}$，但 $K_I=K_{I1}$时，必须经过时间 t_1 后，使裂纹尖端的 K_I 随着裂纹的扩展达到 K_{IC}，这时才发生断裂；当 $K_I=K_{I2}$时，须经过时间 t_2 后才发生断裂，当 K_I 降低到某一定值后，材料就不会由于应力腐蚀而发生断裂，即材料有无限寿命。此时的 K_I 就称为应力腐蚀临界应力强度因子，或称为应力腐蚀门坎值，并以 K_{ISCC}表示。对于一定的材料在一定的介质下，K_{ISCC}为一常数。

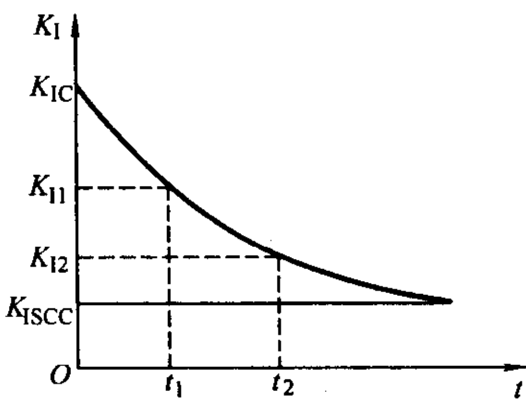


图 6.7-56 应力强度因子 K_I 与延滞断裂时间 t 的关系曲线

K_{ISCC}是材料性能的一个指标，可以用它来建立材料发生应力腐蚀断裂的判据。当裂纹尖端的应力强度因子 K_I 大于材料的 K_{ISCC}时，材料就可能发生应力腐蚀而导致破坏，其可能开裂的判据为

$$K_I \geqslant K_{ISCC} \tag{6.7-19}$$

或

$$\sigma \geqslant \frac{K_{ISCC}}{\alpha\sqrt{\pi a}}$$

式中，K_I 为裂纹尖端的应力强度因子；K_{ISCC}为应力腐蚀临界应力强度因子；a 为裂纹长度之半；α 为应力强度因子系数，由零件的形状和尺寸、裂纹的形状和尺寸、裂纹所在位

置和载荷形式等决定。

当裂纹尖端的 $K_I > K_{ISCC}$ 时，裂纹就会随时间而扩展。单位时间内裂纹的扩展量称为应力腐蚀裂纹扩展速度，用 da/dt 表示。实验证明，da/dt 为裂纹尖端的应力强度因子的函数，即

$$da/dt = f(K_I) \quad (6.7\text{-}20)$$

在 da/dt 与 K_I 的坐标平面上，两者的关系曲线如图6.7-57所示。曲线一般由三段组成：第一阶段——当 K_I 超过 K_{ISCC} 时，裂纹经过一段孕育期后，突然加速扩展；第二阶段——曲线出现水平段，da/dt 与 K_I 几乎无关，因为这一阶段裂纹尖端变钝，裂纹扩展主要受电化学过程控制；第三阶段——裂纹长度已接近临界裂纹尺寸，da/dt 又明显地依赖于 K_I，da/dt 随 K_I 的增大而加快，这是裂纹走向快速扩展的过渡区，当 K_I 达到 K_{IC} 时，裂纹便发生失稳扩展，材料断裂。

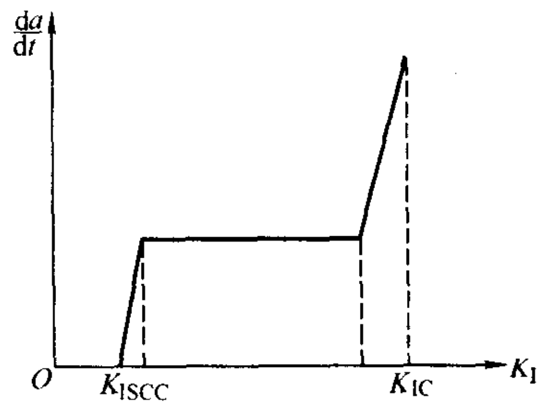


图6.7-57 da/dt 与 K_I 的关系曲线

4.1.3 应力腐蚀的预防措施

应力腐蚀断裂是一种低应力脆性断裂，其危害性很大，应尽量避免或减轻。预防措施如下。

1）在设计中考虑降低零件的工作应力，如增大零件的横截面面积，改进结构，降低应力集中部位的峰值拉应力。采用喷丸或表面热处理方法，使零件表层产生残余压应力，以提高应力腐蚀强度。考虑到工程上发生的应力腐蚀断裂，大部分是由制造和装配过程中产生的残余拉应力引起的，因此，采用退火方法消除残余应力是行之有效的。

2）改变介质条件可以减小或消除材料的应力腐蚀断裂敏感性。主要手段是减小或消除助长应力腐蚀断裂的有害化学离子。例如，通过水的净化处理，降低冷却水与蒸汽中的氯离子含量，对预防奥氏体不锈钢的应力腐蚀断裂是有效的；在一些环境中添加防腐剂也是一种有效的方法。

3）由于一定的合金只在相应的介质中才显示应力腐蚀断裂敏感性，所以可选取合适的合金来避免应力腐蚀，即应尽量选用在某种介质中有较高 K_{ISCC} 的合金，以提高零件抗应力腐蚀的性能。

4）采用外加电位的方法，使金属在介质中的电位远离应力腐蚀断裂的敏感电位区域，从而完成电化学保护。一般的阳极保护或阴极保护就是应用这种方法的实例。但应注意，对于高强度钢或其他对氢脆敏感的材料，不能采用阴极保护法。

4.2 腐蚀疲劳

腐蚀介质与循环应力交互作用，能大大降低材料和零构件的疲劳强度。腐蚀介质和循环应力（应变）的复合作用所导致的疲劳称腐蚀疲劳。

应力腐蚀和腐蚀疲劳的区别在于，应力腐蚀只有在特定的腐蚀环境中才发生，而腐蚀疲劳在任何腐蚀环境及循环应力复合作用下，都会发生腐蚀疲劳断裂。应力腐蚀开裂，有一个临界应力强度因子 K_{ISCC}，当应力强度因子 $K_I \leq K_{ISCC}$，就不发生应力腐蚀开裂。但腐蚀疲劳不存在临界应力强度因子，只要在腐蚀环境中有循环应力继续作用，断裂总是会发生的。

腐蚀疲劳与空气中疲劳的区别，在腐蚀疲劳过程中，除不锈钢和渗氮钢以外，机械零部件表面均变色。腐蚀疲劳形成的裂纹数目较多，即呈多裂纹。腐蚀疲劳的 $S-N$ 曲线没有水平部分，因此，对于腐蚀疲劳极限，一定要指出是某一寿命（即达到破坏的循环数）下的值，即只存在条件腐蚀疲劳极限。影响腐蚀疲劳强度的因素要比空气中的疲劳多而且复杂，如在空气中，疲劳试验频率小于1 000 Hz时，频率基本上对疲劳极限没有影响，但腐蚀疲劳在频率的整个范围内都有影响。

4.2.1 腐蚀疲劳的 $S-N$ 曲线

图6.7-58～图6.7-78是腐蚀疲劳的 $S-N$ 曲线和频率影响、预腐蚀时间、温度的影响曲线。

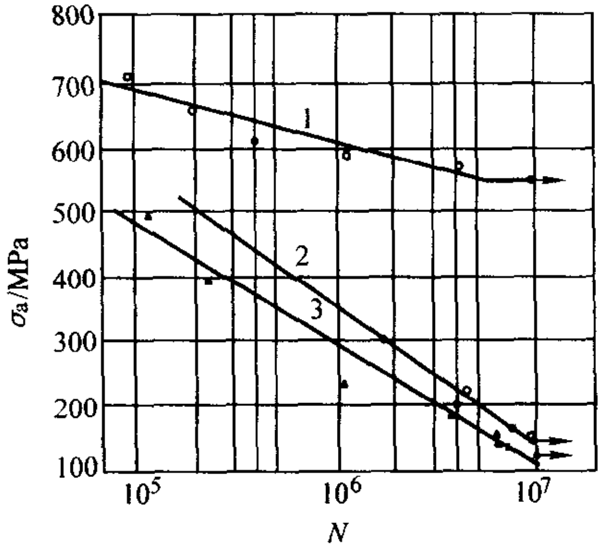


图6.7-58 40Cr钢的腐蚀疲劳 $S-N$ 曲线

1—在室温大气中；2—流动自来水，17℃；3—$w_{NaCl}=3\%$ 水溶液（17℃）中

40Cr钢热处理：840℃油淬，500℃保温，油冷，

力学性能：$\sigma_b = 1\ 147$ MPa 光滑试样（$\alpha_\sigma = 1$），

旋转弯曲试验（$r=-1$），应力频率 $f = 3\ 000\ \text{min}^{-1}$

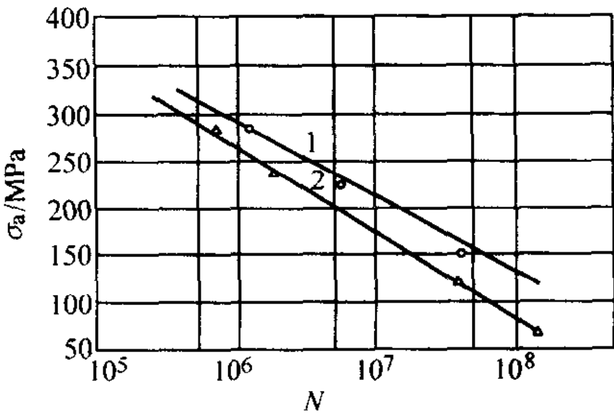


图6.7-59 40Cr钢在天然海水中的 $S-N$ 曲线

1—应力频率，$f = 3\ 000\ \text{min}^{-1}$；2—应力频率，$f = 1\ 000\ \text{min}^{-1}$

力学性能：$\sigma_b = 1\ 147$ MPa 光滑试样，旋转弯曲疲劳试验，室温，试样直径 ϕ7.0 mm

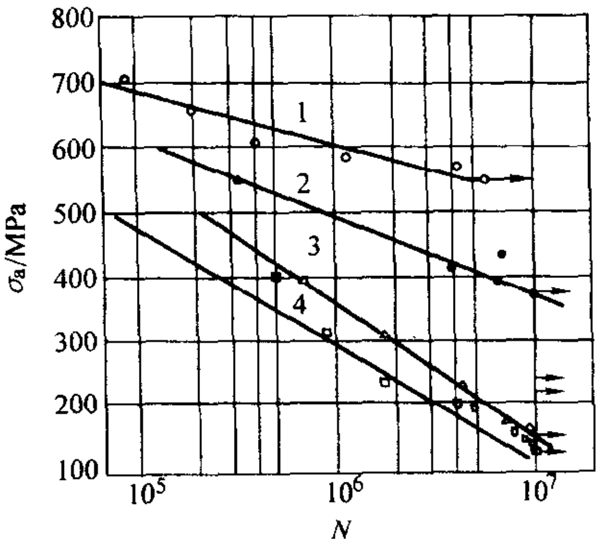


图6.7-60 40Cr钢在不同温度淡水下的 $S-N$ 曲线

1—在室温大气中；2—4℃流动自来水中；3—17℃流动自来水中；4—24℃流动自来水中

40Cr钢热处理：840℃油碎，500℃保温，

油冷 力学性能：$\sigma_b = 1\ 147$ MPa，光滑试样（$\alpha_\sigma = 1$），旋转弯曲试验（$r=-1$），应力频率 $f = 3\ 000\ \text{min}^{-1}$

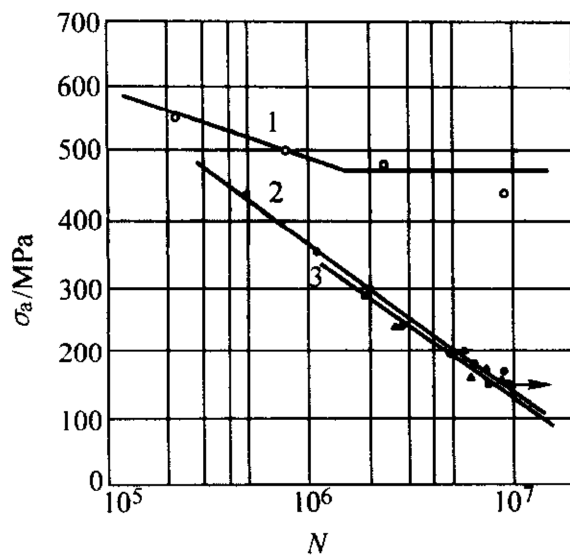


图 6.7-61　20CrMo 钢在淡水中的 $S-N$ 曲线

1—在室温大气中；2—热处理Ⅰ，淡水中；3—热处理Ⅱ，淡水中
20CrMo 钢：热处理Ⅰ（500℃回火）$\sigma_b = 986$ MPa，热处理Ⅱ（580℃回火）$\sigma_b = 934$ MPa，光滑试样（$\alpha_\sigma = 1$），旋转弯曲试验（$r = -1$），应力频率 $f = 3\,000$ min^{-1}
腐蚀介质：流动自来水，17℃

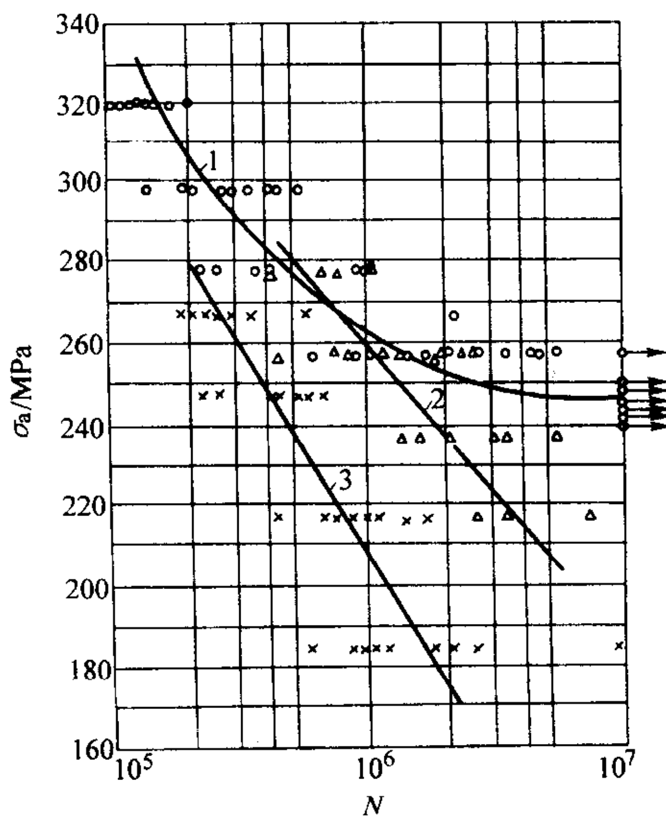


图 6.7-62　12CrNiMo 钢腐蚀疲劳的 $S-N$ 曲线

1—空气中；2—自来水中；3—人造海水中
材料规格：$\delta = 25$ mm 力学性能：$\sigma_b = 725.2$ MPa　缺口试样（$\alpha_\sigma = 2.05$），旋转弯曲试验（$r = -1$），应力频率 $f = 3\,000$ min^{-1}

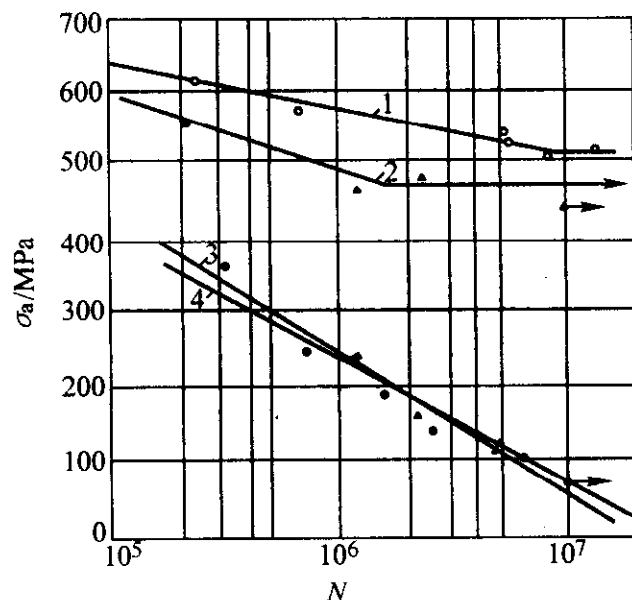


图 6.7-63　船用钢在海水中的 $S-N$ 曲线

1—402 船用钢在室温大气下；2—20CrMo 钢在室温大气下；3—402 船用钢在 24℃天然海水中；4—20CrMo 钢在 24℃天然海水中
402 船用钢 $\sigma_b = 936$ MPa　20CrMo 钢 $\sigma_b = 986$ MPa　热处理：880℃水淬，500℃水冷光滑试样，旋转弯曲试验（$r = -1$），应力频率 $f = 3\,000$ min^{-1}

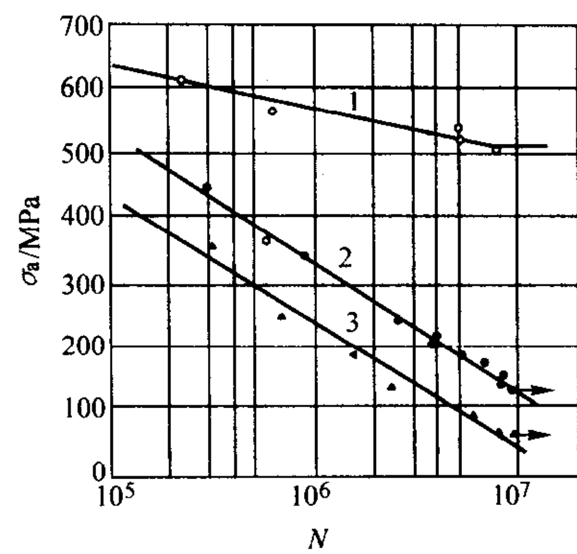


图 6.7-64　402 船用钢在室温大气中，24℃海水中和自来水中的 $S-N$ 曲线

1—室温大气下；2—流动自来水中；3—天然海水（葫芦岛）中
402 船用钢 $\sigma_b = 936$ MPa　热处理：860℃油淬，600℃油冷，光滑试样（$\alpha_\sigma = 1$），旋转弯曲试验（$r = -1$），应力频率 $f = 3\,000$ min^{-1}

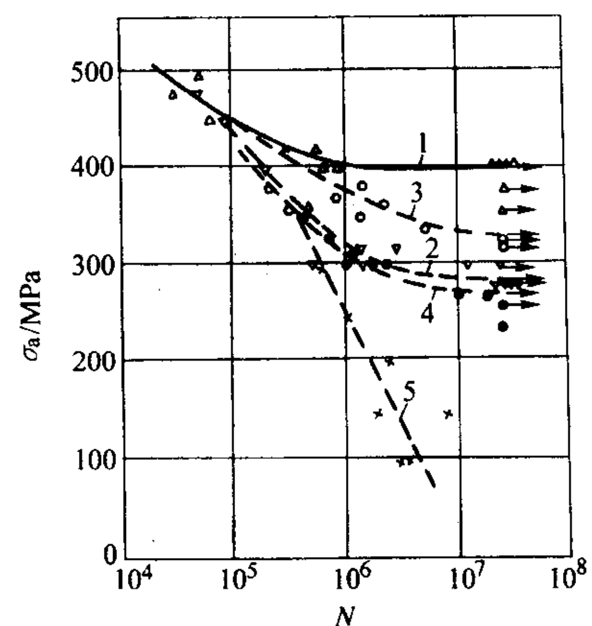


图 6.7-65　21/7 铬镍不锈钢的 $S-N$ 曲线

1—空气中，2～5 都在 $w_{NaCl} = 3\%$ 水溶液中；
2—电路切断；3—相对于溶液加上电位 0 mV；
4—加上 50 mV；5—加上 200 mV

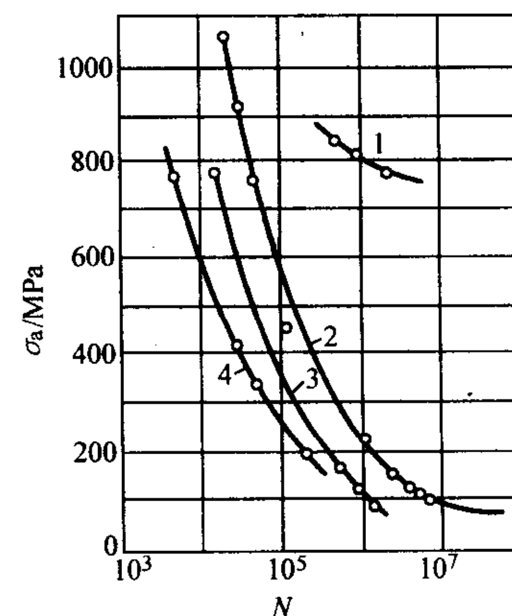


图 6.7-66　镍硅钢的腐蚀疲劳的 $S-N$ 曲线

1—在空气中；2—淡水中，$f = 1\,450$ min^{-1}；
3—淡水中，$f = 50$ min^{-1}；4—淡水中，$f = 5 \sim 8$ min^{-1}

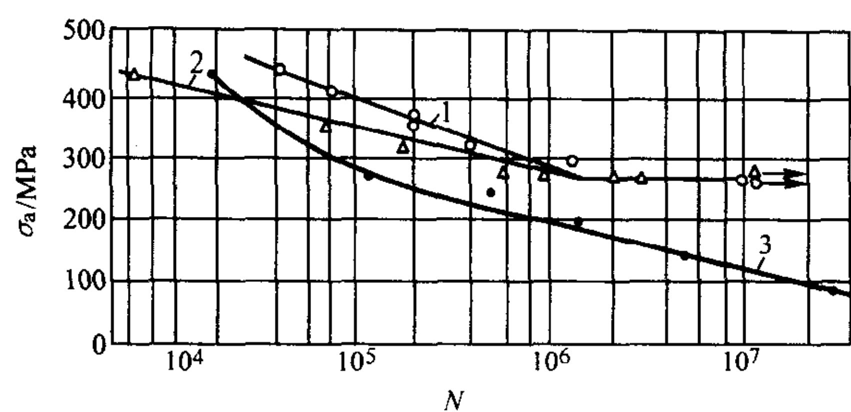


图 6.7-67 13Cr 不锈钢腐蚀疲劳的 $S-N$ 曲线
1—在空气中；2—在蒸汽中；
3—在蒸汽加 $w_{NaCl}=3\%$ 的腐蚀环境

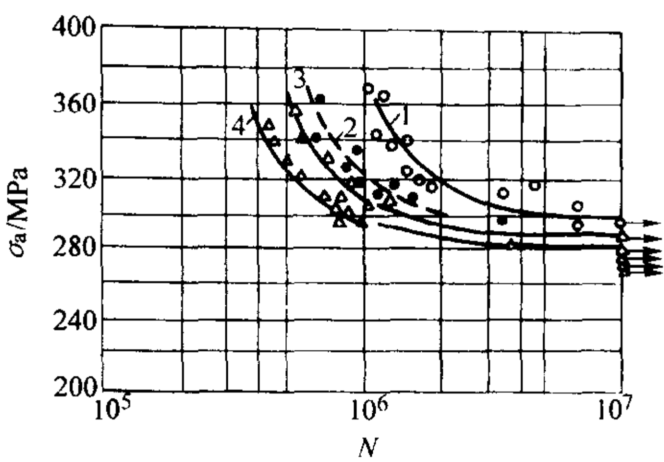


图 6.7-70 空气、氧和氩对低碳钢高周疲劳的 $S-N$ 曲线的影响
1—干氩；2—湿氩；3—湿氧；4—干氧，空气；
温度 25℃时有疲劳极限 $\sigma_{-1}=283$ MPa

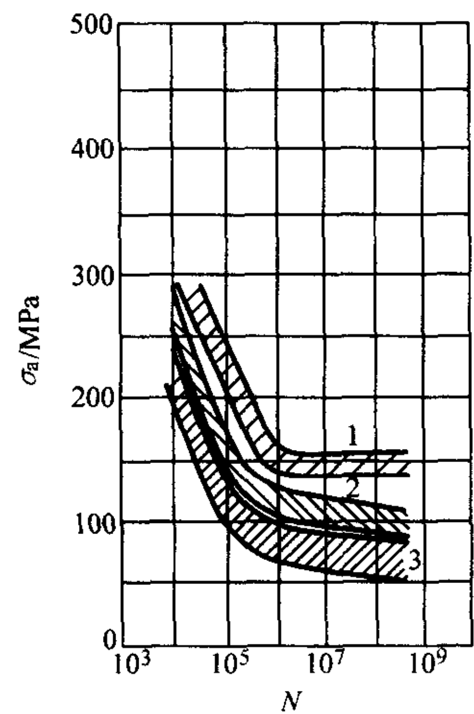


图 6.7-68 13Cr 不锈钢缺口试样的 $S-N$ 曲线
1—空气中；2—蒸馏水中；3—$w_{NaCl}=1\%$水溶液中
13Cr 钢力学性能：$\sigma_b=760\sim830$ MPa，$\sigma_s=610\sim650$ MPa，
旋转弯曲试验（$r=-1$），$f=50$ Hz，温度 23℃

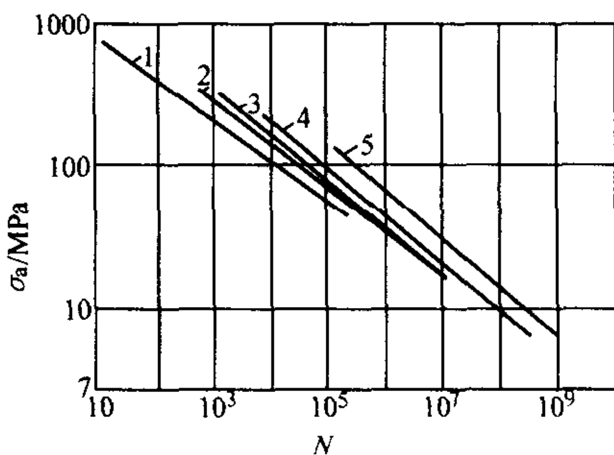


图 6.7-71 碳钢及低合金钢在淡水流中试验频率对疲劳强度的影响
1—频率 $f=0.083\ \text{min}^{-1}$；2—频率 $f=10\ \text{min}^{-1}$；3—频率 $f=50\ \text{min}^{-1}$；4—频率 $f=500\ \text{min}^{-1}$；5—频率 $f=10\,000\ \text{min}^{-1}$

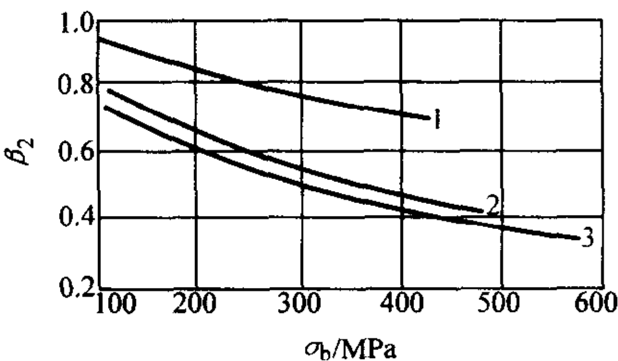


图 6.7-72 预腐蚀对铝合金疲劳极限的影响
1—10 天；2—50 天；3—100 天（天数—试验前将试样浸于淡水中的天数）
试验循环次数 10^7，旋转弯曲试验

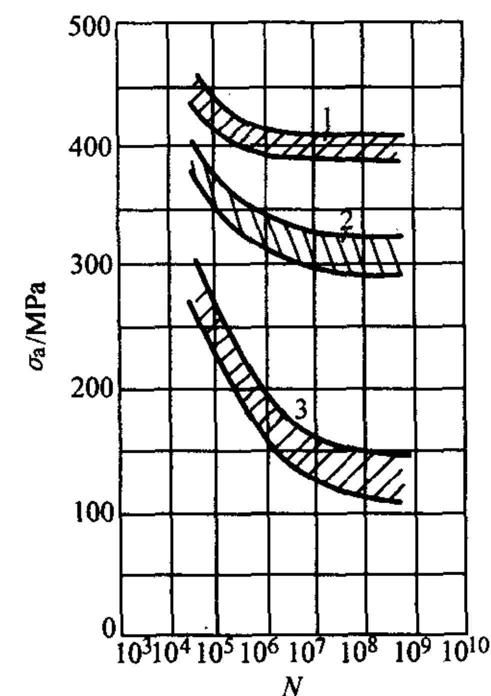


图 6.7-69 13Cr 不锈钢光滑试样的 $S-N$ 曲线
1—空气中；2—蒸馏水中；3—$w_{NaCl}=1\%$水溶液中
13Cr 钢化学成分：$w_{Cr}=13\%$，$w_C=0.20\%$
$\sigma_b=760\sim830$ MPa，$\sigma_s=610\sim650$ MPa，
旋转弯曲试验（$r=-1$），$f=50$ Hz，温度 23℃

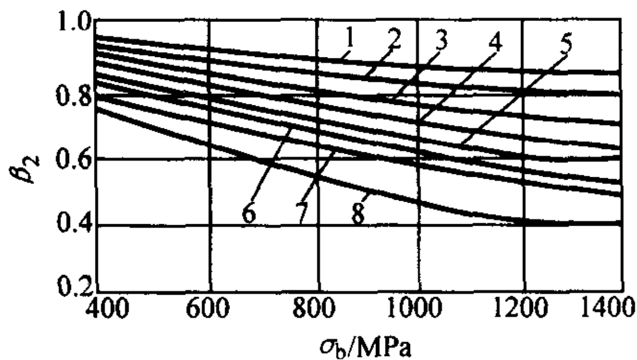


图 6.7-73 预腐蚀对钢试样疲劳极限的影响
1—1 天；2—2 天；3—4 天；4—7 天；5—10 天；
6—25 天；7—50 天；8—200 天
（天数—试验前将试样浸于淡水中的天数）
试验循环次数 10^7，旋转弯曲试验

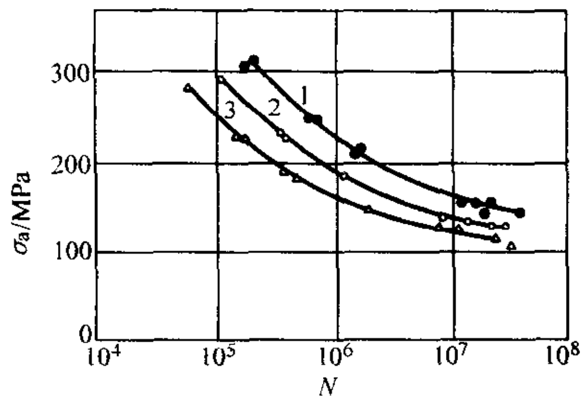


图 6.7-74　18Cr11Ni2Mo 不锈钢在 $w_{H_2SO_4}=5\%$ 水溶液中的 S—N 曲线

1—25℃；2—50℃；3—75℃

力学性能：$\sigma_s=255$ MPa

悬臂弯曲试验（$r=-1$），应力频率 $f=1\ 700\ min^{-1}$

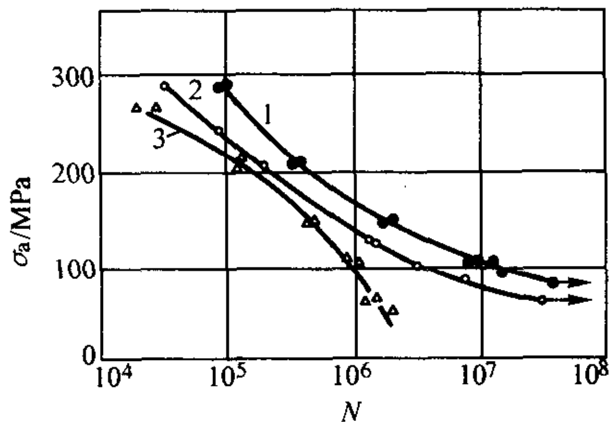


图 6.7-75　19Cr12Ni 不锈钢在 $w_{H_2SO_4}=5\%$ 水溶液中的 S—N 曲线

1—25℃；2—50℃；3—75℃

力学性能：$\sigma_s=220.7$ MPa

悬臂弯曲试验（$r=-1$），应力频率 $f=1700\ min^{-1}$

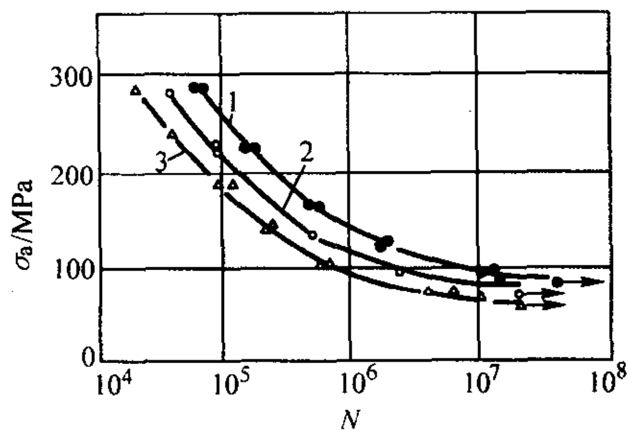


图 6.7-76　19Cr12Ni 不锈钢在 $w_{H_2SO_4}=5\%$ 水溶液中的 S—N 曲线

1—25℃；2—50℃；3—75℃

力学性能：$\sigma_s=216$ MPa

悬臂弯曲试验（$r=-1$），应力频率 $f=1\ 700\ min^{-1}$

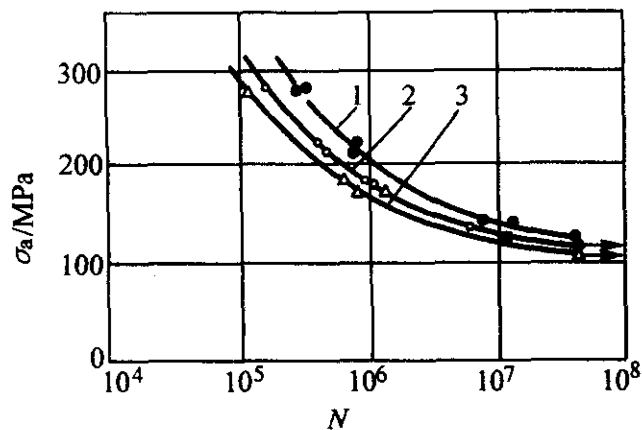


图 6.7-77　18Cr13Ni2Mo 不锈钢在 $w_{H_2SO_4}=5\%$ 水溶液中的 S - N 曲线

1—25℃；2—50℃；3—75℃

力学性能：$\sigma_s=225.6$ MPa

悬臂弯曲试验（$r=-1$），应力频率 $f=1\ 700\ min^{-1}$

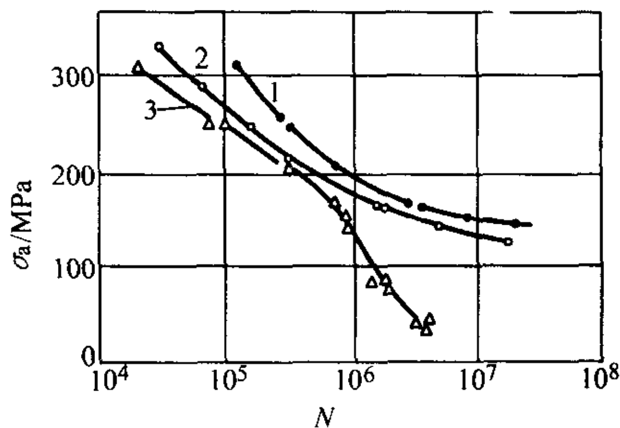


图 6.7-78　18Cr11Ni1Mo 不锈钢在 $w_{H_2SO_4}=5\%$ 水溶液中的 S—N 曲线

1—25℃；2—50℃；3—75℃

力学性能：$\sigma_s=240.3$ MPa

悬臂弯曲试验（$r=-1$），应力频率 $f=1\ 700\ min^{-1}$

4.2.2　腐蚀疲劳极限

表 6.7-21 ~ 表 6.7-24 给出在腐蚀条件下的疲劳极限数据。

表 6.7-21　某些国产钢种的腐蚀疲劳极限 $(\sigma_{-1})_{cf}$

钢种	抗拉强度 σ_b/MPa	试验频率 f/min^{-1}	腐蚀环境	试验循环次数 N	腐蚀疲劳极限 $(\sigma_{-1})_{cf}$/MPa
40Cr	1 170	3 000	$w_{NaCl}=3\%$ 水溶液	10^7	130
			自来水	10^7	155
20CrMo	954	3 000	海水	10^7	110
			自来水	10^7	150
ZG20-SiMn	510	5 000	自来水滴水	10^7	175
			浸水	10^7	178
ZG0Cr13-Ni4Mo	784	5 000	自来水滴水	10^7	200
			浸水	10^7	218

表 6.7-22　蒸汽对钢试样腐蚀疲劳的影响

（试验循环次数 $N=5\times10^7$；旋转弯曲；应力频率 $f=2\ 200\ min^{-1}$）

材料	σ_b /MPa	疲劳极限/MPa				
		在空气中 σ_{-1}	在空气中喷蒸汽 $(\sigma_{-1})_{cf}$	已知温度及蒸汽压力		
				100℃ 0 MPa	149℃ 0.41 MPa	371℃ 1.51 MPa
$w_{Ni}=3.5\%$钢	725	316	161	—	246	239
$w_{Ni}=3.5\%$钢	814	401	161	402	369	362
$w_{Ni}=3.5\%$钢，镀铬	—	—	285	—	315	—
$w_{Cr}=12.5\%$不锈钢	696	416	223	369	377	369
氮化钢：$w_C=0.36\%$，$w_{Cr}=1.5\%$，$w_{Al}=1.2\%$	853	510	—	—	439	345
氮化钢：$w_C=0.36\%$，$w_{Cr}=1.5\%$，$w_{Al}=1.2\%$ 经氮化	—	625	500	—	478	402

表 6.7-23　钢的腐蚀疲劳试验数据

材　　料	热处理	σ_b /MPa	试验方式	应力频率 f/min^{-1}	腐蚀环境	试验循环次数	σ_{-1}③ /MPa	$(\sigma_{-1})_{cf}$④ /MPa	$\beta_2=\frac{(\sigma_{-1})_{cf}}{\sigma_{-1}}$
软钢 18/8Cr-Ni-W 钢	正火 退火		旋转 弯曲		淡水 滴注	10^8	268 277	32 175	0.12 0.63
$w_C=0.21\%$碳钢	退火	500	旋转 弯曲 拉伸	1 300 1 500	海水	10^8	225 142	30 39	0.13 0.27
$w_{Cr}=0.12.5\%$铬钢 18/8 不锈钢 $w_{Cr}=18.5\%$铬钢 $w_C=0.48\%$碳钢，镀镉	退火	1 020 1 320 790 1 040	拉伸	360	淡水	2.5×10^7	257 194 246 203	126 83 194 52	0.49 0.43 0.79 0.26
$w_C=0.35\%$碳钢		610	旋转 弯曲	1 750	盐水① 盐水②	10^7	285	173 74	0.61 0.26
$w_C=0.50\%$碳钢		660	旋转 弯曲	1 750	盐水① 盐水②	10^7	222	140 77	0.63 0.35
$w_C=0.50\%$碳钢	调质	910	旋转 弯曲	1 750	盐水① 盐水②	10^7	424	178 97	0.42 0.23
合金钢（$w_{Cr}=0.8\%\sim1.1\%$，$w_{Mo}=0.15\%\sim0.25\%$）	调质	900	旋转 弯曲	1 750	盐水① 盐水②	10^7	493	189 99	0.38 0.20
合金钢（$w_C=0.55\%\sim0.65\%$，$w_{Si}=1.8\%\sim2.2\%$）	正火	1 010	旋转 弯曲	1 750	盐水① 盐水②	10^7	507	175 104	0.35 0.20
$w_{Cr}=5\%$铬钢	调质	910	旋转 弯曲	1 750	盐水① 盐水②	10^7	520	371 109	0.71 0.21
熟铁		330	旋转 弯曲	1 750	盐水① 盐水②	10^7	215	137 115	0.64 0.54

① $w_{NaCl}=6.8\%$盐水溶液，试样整体浸入。
② $w_{NaCl}=6.8\%$盐水与饱和 H_2S，试样整体浸入。
③ 在空气中的疲劳极限。
④ 在腐蚀环境中的疲劳极限。

表 6.7-24　有色金属的腐蚀疲劳试验数据

材　　料	热处理	σ_b /MPa	试验方式	应力频率 f/min^{-1}	腐蚀环境	试验循环次数	σ_{-1}③ /MPa	$(\sigma_{-1})_{cf}$④ /MPa	$\beta_2=\frac{(\sigma_{-1})_{cf}}{\sigma_{-1}}$
铝	退火	75	旋转弯曲	1 450	淡水① 含盐量为海水的 1/3 的河水②	2×10^7	37	— 15②	— 0.41
	半硬化	96					44	— 22②	— 0.50
	硬化	124					64	37① 30②	0.58 0.47
硬铝（铝铜镁合金）	退火	206					107	52① 45②	0.49 0.42
	已热处理	427					110	62① 52②	0.56 0.47
电解铜，热轧	退火	193					62	— 64②	— 1.03
电解铜，冷轧	回火	289					104	107① 107②	1.03 1.03
铜镍合金（$w_{Cu}=78\%$，$w_{Ni}=21\%$），冷轧	退火	289					110	117① 117②	1.06 1.06
	回火	379					160	147① 160②	0.92 1.00
铜镍合金（$w_{Cu}=48\%$，$w_{Ni}=48\%$），冷轧		475					234	179① 202②	0.76 0.86
铜镍合金（$w_{Ni}=67\%$，$w_{Cu}=30\%$），冷轧	退火	503					222	165① 179②	0.74 0.81
	回火	779					325	190① 215②	0.58 0.66
镍，冷轧	退火	475					209	154① 142②	0.74 0.68
	回火	806					319	184① 165②	0.58 0.52
铜锌合金（$w_{Cu}=62\%$，$w_{Zn}=37\%$），冷拔	退火	324					137	— 117②	— 0.85
	回火	517					147	110① 110②	0.75 0.75

续表 6.7-24

材料	热处理	σ_b /MPa	试验方式	应力频率 f/min^{-1}	腐蚀环境	试验循环次数	σ_{-1}③ /MPa	$(\sigma_{-1})_{cf}$④ /MPa	$\beta_2=\frac{(\sigma_{-1})_{cf}}{\sigma_{-1}}$
硬铝（$w_{Mg}=2.5\%$）	轧制	386	旋转弯曲 轴向加载	1 450	w_{NaCl} = 3%盐雾	5×10^7	126 110	46 35	0.37 0.32
		227	旋转弯曲 轴向加载			10^7	89 75	13 13	0.15 0.17
碲铅（$w_{Te}=0.05\%$，$w_{Cu}=0.06\%$）			旋转弯曲		$w_{H_2SO_4}$ = 38%滴流	4×10^7	3.7	2.4	0.65
锑铅（$w_{Sb}=1\%$）							5.2	4.5	0.87
蓄电池铅							12	11	0.93
AZG 镁铝锌					自来水	2×10^7	70	34	0.49
AM537 镁铝锰							70	44	0.63
AZ855 镁铝锌⑦							131	48	0.37
AM503 镁铝锰⑧					3%盐水		49	17	0.35
AZM 镁铝锰⑨							136	11	0.08
磷青铜（$w_{Sn}=4.2\%$）	轧和拉拔，正火	379			3%盐水	5×10^7	137	163	1.19
铝青铜（$w_{Al}=9.8\%$，$w_{Zn}=1.4\%$）	挤压和拉拔	489					200	135	0.68
耐蚀高强度铜合金⑤	挤压和拉锻	572					227	246	1.08
$w_{Al}=9.2\%$，$w_{Ni}=5.0\%$，$w_{Fe}=5.4\%$		710					310	201	0.65
铝青铜（$w_{Al}=9.3\%$）	淬火 淬火，热处理	203 448					157 136	120 107	0.76 0.79
铍青铜（$w_{Be}=2.2\%$）	溶液处理 热理处	441 1 117					246 274	187 219	0.76 0.80
铝-锌-镁合金⑥ DTD683（7075）	溶液处理 热处理 时效	255 427 379			w_{NaCl} = 3%盐溶液，液体薄膜	10^7	124 172 151	62 69 69	0.50 0.40 0.46
纯铝					$w_{H_2SO_4}$ = 38%硫酸滴流	4×10^7	2.6	—	—

① 淡水。
② 含盐量为海水的 1/3 的河水。
③ 在空气中的疲劳极限。
④ 在腐蚀环境中的疲劳极限。
⑤ 耐蚀高强度铜合金的化学成分：$w_{Al}=8.5\%\sim10.5\%$，$w_{Fe}=4\%\sim6\%$，$w_{Ni}=4\%\sim6\%$，其余铜。
⑥ DTD683（7075），相当于中国的铝合金号 7A09。
⑦ AZ855 镁合金的化学成分：$w_{Al}=8.0\%$，$w_{Zn}=0.4\%$，$w_{Mn}=0.3\%$。
⑧ AM503 镁合金的化学成分：$w_{Mn}=1.5\%$。
⑨ AZM 镁合金的化学成分：$w_{Al}=6.0\%$，$w_{Zn}=1.0\%$，$w_{Mn}=0.3\%$。

4.2.3 影响腐蚀疲劳的因素

影响腐蚀疲劳的因素有：加载频率和应力波形、加载方式、平均应力、应力集中、试样尺寸、表面状态及处理等。有些影响因素在上述的 $S-N$ 曲线中已有阐述。影响结果见表 6.7-25 ~ 表 6.7-33。

表 6.7-25　低碳钢（$w_C=20\%$）试样的旋转弯曲的腐蚀疲劳极限

试样直径/mm	在空气中的疲劳极限 σ_{-1}/MPa	浸在盐水中的腐蚀疲劳极限（$N=6\times10^7$）$(\sigma_{-1})_{cf}$/MPa
10	205	49
130	191	112

表 6.7-26　弯曲及拉压的疲劳极限

材　料	σ_b /MPa	空气中的疲劳极限 σ_{-1}/MPa		$(\sigma_{-1})_{cf}$/MPa w_{NaCl} = 3%盐溶液喷雾中（$N=5\times10^7$，f = 2 200 min^{-1}）	
		弯曲	拉压	弯曲	拉压
碳钢（$w_C=0.48\%$）	975	386	237	43	37
不锈钢（$w_C=0.12\%$，$w_{Cr}=14.7\%$）	619	380	339	139	169
奥氏体不锈钢（$w_C=0.11\%$，$w_{Cr}=18.3\%$，$w_{Ni}=8.2\%$）	1 023	366	370	244	228
不锈钢（$w_C=0.25\%$，$w_{Cr}=17\%$，$w_{Ni}=1.16\%$）	843	505	439	190	240
硬铝	435	139	123	53	40

表 6.7-27 20Cr 钢的尺寸对腐蚀疲劳极限的影响

环境	材料性能	试样直径		
		$d=16$ mm	$d=32$ mm	$d=40$ mm
空气 ($N=5\times10^6$)	σ_b/MPa β_2 ε	264 1.0 1.0	248 1.0 0.937	240 1.0 0.907
机油 ($N=10^7$)	$(\sigma_{-1})_{cf}$/MPa β_2 ε	243 0.92 1.0	235 0.95 0.964	230 0.96 0.945
淡水 ($N=2\times10^7$)	$(\sigma_{-1})_{cf}$/MPa β_2 ε	122 0.462 1.0	140 0.565 1.14	154 0.64 1.26

注：悬臂式旋转弯曲试验，频率 $f=2\,000\ \text{min}^{-1}$。

4.2.4 腐蚀疲劳的寿命估算

腐蚀疲劳的 $S-N$ 曲线没有水平部分，所以腐蚀疲劳只有有限寿命设计。腐蚀疲劳的寿命估算方法有二，即用 $S-N$ 曲线的常规疲劳设计方法和用断裂力学的裂纹扩展理论估算寿命。由于在腐蚀环境和循环载荷复合作用下，无裂纹寿命很短，因此，腐蚀疲劳的寿命，主要是裂纹的扩展寿命。

腐蚀疲劳的 $S-N$ 曲线，其影响因素比空气中的多而且复杂。在空气中影响材料 $S-N$ 曲线的主要因素，有应力集中系数、尺寸系数和表面系数三种。在腐蚀疲劳中，应力集中和表面粗糙度的影响要比空气中的严重，而尺寸越小，腐蚀疲劳强度降低越多，这些都是在指定寿命（循环数）的基础上讲的，可以参阅上面给出的腐蚀疲劳极限数据和 $S-N$ 曲线。所以，提高零件腐蚀疲劳强度的措施，较有限的是进行表面强化工艺和镀或涂保护层。

表 6.7-28 腐蚀环境及应力集中同时作用的疲劳极限

材料及试验方式	试样 d/mm	疲劳极限/MPa		有效应力集中系数		腐蚀系数 β_2
		空气中 σ_{-1} 或 τ_{-1}	腐蚀环境中 $(\sigma_{-1})_{cf}$ 或 $(\tau_{-1})_{cf}$	空气中 K_σ 或 K_τ	腐蚀环境中 $K_{\sigma f}$ 或 $K_{\tau f}$	
20Cr 弯曲	光滑试样，$d=8$ 缺口试样，$d=14$	318 151	210 151	2.11	2.11	0.66
20Cr 弯曲	光滑试样，$d=20$ 缺口试样，$d=20$	285 133	166 122	2.07	2.25	0.61
40Cr（正火）弯曲	光滑试样，$d=8$ 缺口试样，$d=8$	426 266	364 248	1.6	1.72	0.85
铸铁 弯曲	光滑试样，$d=20$ 缺口试样，$d=20$	117 105	107 89	1.11	1.32	0.92
镍铬钢 扭转 $\sigma_b=784.6$ MPa	光滑试样 有肩试样 有肩试样 有孔试样	302 196 192 151	223 188 205 93	— 1.54 1.57 2.00	— 1.60 1.47 3.25	0.74 — — —
镍铬钢[①] 扭转 $\sigma_b=1\,108$ MPa	光滑试样 有肩试样 有孔试样	384 254 205	223 137 137	— 1.51 1.87	— 2.8 2.8	0.58 — —
镍铬钢 弯曲 $\sigma_b=872.8$ MPa	光滑试样 有肩试样 有孔试样	439 247 212	233 130 109	— 1.78 2.07	— 3.37 4.0	0.53 — —
镍铬钢 弯曲 $\sigma_b=1\,079$ MPa	光滑试样 有肩试样 有孔试样	617 247 212	89 75 61	— 2.5 2.9	— 8.18 10.0	0.145 — —
灰铸铁 $\sigma_b=274.6$ MPa	光滑试样 缺口试样	120 103	97 89	— 1.17	— 1.35	0.8 —
钢 弯曲 $\sigma_b=539.4$ MPa	光滑试样 有肩试样 有孔试样	370 168 171	199 89 123	— 2.2 2.16	— 4.15 3.0	0.54 — —
钢 弯曲 $\sigma_b=485.4$ MPa	光滑试样 有肩试样 有孔试样	343 164 162	164 96 116	— 2.08 2.11	— 3.57 2.94	0.48 — —
钢 弯曲 $\sigma_b=627.6$ MPa	光滑试样 有肩试样 有孔试样	374 178 182	130 103 109	— 2.1 2.05	— 3.64 3.41	0.35 — —
钢 弯曲 $\sigma_b=858.1$ MPa	光滑试样 有肩试样 有孔试样	436 205 171	96 75 89	— 2.12 2.54	— 5.77 4.88	0.22 — —

① 镍铬钢成分：$w_C=0.4\%$，$w_{Mn}=0.75\%$，$w_{Ni}=1.0\%\sim1.5\%$，$w_{Cr}=0.45\%\sim0.75\%$。

表 6.7-29 拉压脉动循环的疲劳极限

试样和试验条件	脉动循环疲劳极限/MPa	
	脉动拉伸	脉动压缩
经过磨削，在大气中	1 177	1 618
有缺口（深 0.03 mm），在大气中	931	1 500
经过磨削，在淡水中	147	1 540

表 6.7-30 表面高频淬火对 45Cr 钢疲劳极限的影响

试样处理方法	疲劳极限（$N=10^7$）			
	在大气中		在 $w_{NaCl}=3\%$ 的溶液中	
	MPa	%	MPa	%
正火（原始状态）	252	100	98	100
电解镀铬	199	79	85	87
同上，预先经过高频淬火	339	134	294	300

此外，试验频率和腐蚀环境的温度，都对零件疲劳强度的影响也很显著。目前进行腐蚀疲劳试验的频率，一般为 1 500～3 000 min^{-1}，太低的频率很少用，主要考虑用低频率进行试验劳动量大，所以，遇到有些文献中没有给定试验频率数据时，可假设在上述的频率范围内。

假使腐蚀疲劳的 $S-N$ 曲线，是模拟零件实际使用条件进行试验得出的，那么用这个 $S-N$ 曲线可以直接估算而得到该零件的寿命。在一般情况下，$S-N$ 曲线是用试样在同样腐蚀环境下得到的，则在零件的寿命估算中，需要考虑到由试样到零件，存在着应力集中、尺寸等的影响差异，需要进行修正。

如只有在空气中的 $S-N$ 曲线而需要进行腐蚀疲劳寿命估算，或指定寿命下校核安全系数。这时，可用本章第 4 节高周疲劳的方法，但此时应考虑腐蚀系数 β_2。考虑的方法是，没有进行强化工艺时，不论表面粗糙度如何，都用 $\beta=\beta_2$；有强化时，用强化或镀后试样在腐蚀环境中的腐蚀系数为表面系数 β，不要将 β_1、β_2、β_3 相乘作为 β。

表 6.7-31 镀层对试样的腐蚀疲劳极限的影响

材料	腐蚀环境，试验循环数，试样，应力频率 f/min^{-1}，d/mm	镀层金属	镀层厚度/mm	腐蚀系数 β_2
钢：$w_C=0.36\%$，$w_{Si}=0.28\%$，$w_{Mn}=0.73\%$；在 840～860℃下正火	淡火，光滑试样，$N=10^7$，$f=1\ 450$，$d=10$	Zn	0.030	0.94
钢：$w_C=0.37\%$，$w_{Mn}=0.74\%$，$w_{Cr}=0.61\%$，$w_{Si}=0.21\%$，$w_{Ni}=1.4\%$，淬火回火（$\sigma_b=853.2$ MPa）	淡水，光滑试样，$N=10^8$，$f=1\ 450$，$d=9$	Zn	0.004 0	0.41
		Cd	0.002 5	0.25
		Cd	0.012 5	0.45
		Pb	0.012 5	0.33
50 钢，冷拔，$\sigma_b=980.7$ MPa	$w_{NaCl}=3\%$ 溶液，光滑试样，$N=2\times10^7$，$f=2\ 200$，$d=7$	Zn	0.014	0.87
		Cd	0.013	0.77
50 钢，正火，$\sigma_b=637.5$ MPa		Zn	0.014	0.90
		Cd	0.013	0.84
硬铝：$w_{Cu}=4\%～4.5\%$，$w_{Mn}=0.64\%$，$w_{Mg}=0.63\%$，$w_{Fe}=0.84\%$，$w_{Si}=0.22\%$，$\sigma_b=382.5$ MPa	$w_{NaCl}=3\%$ 溶液，光滑试样，$N=5\times10^7$，$f=2\ 000$，$d=8$	Zn	—	0.71
		Zn + 合成橡胶清漆	—	0.65
		Cd	—	<0.5

表 6.7-32 45 钢经表面强化后在 $w_{NaCl}=3\%$ 溶液中的腐蚀疲劳极限

试样处理方式	疲劳极限（$N=10^7$）			
	MPa		%	
	在大气中	在 $w_{NaCl}=3\%$ 的溶液中	在大气中	在 $w_{NaCl}=3\%$ 的溶液中
磨削	250	98	100	100
喷丸	291	198	116	202
辊压	276	247	111	252
高频淬火	191	351	187	358

表 6.7-33 表面处理对腐蚀疲劳极限的影响（旋转弯曲试验）

材料		σ_b/MPa	表面处理	保护层厚度/mm	应力频率/min^{-1}	腐蚀环境	试验循环次数 N	疲劳极限 σ_{-1}/MPa 未处理	疲劳极限 σ_{-1}/MPa 处理	腐蚀疲劳极限 $(\sigma_{-1})_{cf}$/MPa 未处理	腐蚀疲劳极限 $(\sigma_{-1})_{cf}$/MPa 处理
钢：$w_C=0.5\%$	冷拉 正火	1 992 713	涂瓷漆		2 200	$w_{NaCl}=1\%$盐雾	2×10^7	337 227	317 234	48 55	144 151
	冷拉 正火		电镀锌	0.048 3					344 206		317 227
	冷拉 正火		表面锌化	0.127					310 200		337 206
	冷拉 正火		电解镀锌	0.014 2					337 220		289 206
	冷拉 正火		电解镀镉	0.013 2					317 206		234 186
	冷拉 正火		电解镀镉涂瓷漆	0.012 7					317 220		241 186
	冷拉 正火		电解镀锌	0.012 7					289 213		206 179

续表 6.7-33

材料		σ_b/MPa	表面处理	保护层厚度/mm	应力频率/min^{-1}	腐蚀环境	试验循环次数 N	疲劳极限 σ_{-1}/MPa		腐蚀疲劳极限 $(\sigma_{-1})_{cf}$/MPa	
								未处理	处理	未处理	处理
钢： $w_C=0.5\%$	冷拉 正火	磷酸盐水 处 理 涂瓷漆							310 248		144 179
	冷拉		铝雾	0.050 8					351		275
	冷拉		铝雾涂 瓷 漆	0.050 8		淡水			351		331
中碳钢		772	热 浸 低焊料 热 浸 敷镉层 电镀镍 电镀铬	0.010 2 0.020 3 0.203 0.203		滴流	10^8	193	227 200 137 200	96	813 151 137 200
		818	表面辊压			淡水		227	255	89	131
			表面辊压	0.508		淡水	2×10^8	255	317	<137	262
氮化钢：$w_{Cr}=1.6\%$，$w_{Al}=0.9\%$，$w_{Mo}=0.3\%$			氮化			河水滴流	10^8	455	510	<69	344
铬钒钢：$w_C=0.2\%$，$w_{Cr}=0.9\%$，$w_V=0.1\%$		1 698	氮化		1 450	自来水喷射	10^8		648		524
钢：$w_C=0.47\%$		1 451	电镀锌 表面锌化 镀 锌 镀 镉			淡水	2×10^7	372		124	268 268 303 282
钢：$w_C=0.38\%$			抛光镀锌 韧性镀锌	0.012 7 0.0254 0.012 7 0.025 4		浸入油池中 盐水 液态碳化 物浸湿	10^7	344		74	124 137 117 124

编写：王雷（东北大学）
李佳（东北大学）

第 8 章　特殊材料的强度问题

1　复合材料

复合材料的强度计算参见本大典复合材料卷第 14 篇复合材料的力学问题与设计。

2　聚合物

聚合物具有很大的相对分子质量，高分子化合物的分子运动远比低分子复杂，具有明显的松弛特征，对温度的依赖性很大，聚合物的力学性能不仅取决于分子结构，在更大程度上还取决于这些分子排列、堆砌的聚集态结构。聚合物材料的原料丰富，合成方便，易于加工。近年来按人们希望的性质来设计聚合物也取得突破性进展，使之成为结构材料。由于这些原因，聚合物材料发展迅速，在整个使用材料中，聚合物所占的比例不断增加。

2.1　聚合物的结构特点与力学状态

高分子材料可以是天然的，如纤维素、蛋白质、天然橡胶和天然树脂等；也可以是人工合成的。合成的种类更多，成为人们使用的主要材料之一。聚合物按结构性能和用途，可概略地分为四类，即热塑性塑料、纤维、橡胶和热固性塑料，其结构形状如图 6.8-1 所示。a、b 为热塑性塑料和纤维；c 为橡胶；d 为热固性塑料。实际上，塑料、纤维和橡胶之间并无严格的界限，有的聚合物既可制成纤维，也可用作塑料等。

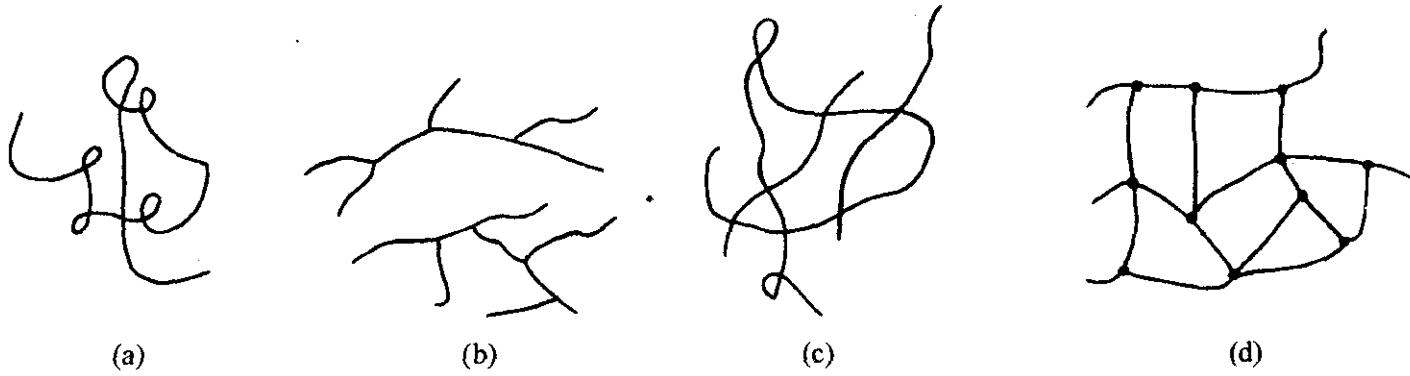


图 6.8-1　不同结构形状的聚合物分子示意图

2.1.1　结构单元

结构单元是组成聚合物的链节。根据高分子主链的结构，聚合物可分为三类。

1）碳链高分子　主链由碳原子单种元素组成，如聚乙烯、聚丙烯等，主要的碳链聚合物见表 6.8-1，大多数碳链高分子的可塑性好，易于加工成形，但耐热性差，易老化，易燃烧。

表 6.8-1　碳链聚合物

聚合物（缩写符号）	结构式	聚合物（缩写符号）	结构式
聚乙烯（乙塑）（PE）	$-CH_2-CH_2-$	聚甲基丙烯酸甲酯（PMMA）	$-CH_2-C(CH_3)(COOCH_3)-$
聚丙烯（丙纶）（PP）	$-CH_2-CH(CH_3)-$	聚丙烯腈（腈纶）（PAN）	$-CH_2-CH(CN)-$
聚异丁烯（PIB）	$-CH_2-C(CH_3)(CH_3)-$	聚醋酸乙烯酯（白胶水）（PVAC）	$-CH_2-CH(COOCH_3)-$
聚苯乙烯（苯塑）（PS）	$-CH_2-CH(C_6H_5)-$	聚乙烯醇（PVA）	$-CH_2-CH(OH)-$
聚氯乙烯（PVC）	$-CH_2-CH(Cl)-$	聚丁二烯（PB）	$-CH_2-CH=CH-CH_2-$
聚偏氯乙烯（PVDC）	$-CH_2-C(Cl)(Cl)-$	聚四氟乙烯（PTFE）	$-CF_2-CF_2-$
聚氟乙烯（PVF）	$-CH_2-CH(F)-$	聚三氟氯乙烯（PCTEF）	$-CF_2-CF(Cl)-$
聚丙烯酸（PAA）	$-CH_2-CH(COOH)-$	聚异戊二烯（PIP）	$-CH_2-C(CH_3)=CH-CH_2-$
聚丙烯酸甲酯（PMA）	$CH_2-CH(COOCH_3)$	聚氯丁二烯（PCB）	$-CH_3-C(Cl)=CH-CH_2-$

2）杂链高分子 主链中除碳原子外，还有氧、氮、硫等其他元素，如聚酰胺、聚酯等。杂链高分子具有较高的耐热性和机械强度，但因分子链带有极性基团，故易于水解。

3）元素有机高分子 主链中没有碳原子，而由硅、氧、氮、铝、硼、钛等元素组成，侧基则为有机基团，如甲基、乙基、乙烯基和芳基等，因此元素高分子兼有无机高分子和有机高分子的特性，即一方面有高的耐热耐寒性，另一方面又有较高的弹性和可塑性。如主链中无碳原子，侧基也不含有机基团，完全由其他元素组成，便是无机高分子，这类元素的成链能力较弱，故聚合物分子量不高，且易水解。

主要的杂链及元素有机聚合物见表6.8-2。

表6.8-2 杂链和元素有机聚合物

聚合物	结构式
聚甲醛	$—O—CH_2—$
聚环氧乙烷	$—OCH_2—CH_2—$
聚对苯二甲酸乙二酯（涤纶）	$—O—CH_2CH_2O—C(=O)—C_6H_4—C(=O)—$
环氧树脂	$—O—C_6H_4—C(CH_3)_2—C_6H_4—O—CH_2CH(OH)CH_2—$
聚碳酸酯	$—O—C_6H_4—C(CH_3)_2—C_6H_4—O—C(=O)—$
双酚A聚砜（聚砜）	$—O—C_6H_4—C(CH_3)_2—C_6H_4—O—C_6H_4—S(=O)_2—C_6H_4—$
聚己内酰胺（尼龙6）	$—NH(CH_2)_5CO—$
聚己二酰己二胺（尼龙66）	$—NH(CH_2)_6—CO(CH_2)_cO—$
聚酰亚胺	$—N[(C=O)_2C_6H_2(C=O)_2]N—C_6H_4—O—C_6H_4—$
聚氨酯（泡沫塑料）	$—O(CH_2)_2OC(=O)NH(CH_2)_6NHC(=O)—$
脲醛树脂（电玉粉）	$—NHC(=O)NH—CH_2$
聚合物	结构式
酚醛树脂（电木）	$—C_6H_3(OH)—CH_2—$
聚硫橡胶	$—CH_2CH_2—S(=S)—S—S(=S)—$
硅橡胶	$—O—Si(CH_3)_2—$

2.1.2 聚合度

聚合物的强度、熔点等和相对分子质量的大小有直接关系，因而需要一个参数表示相对分子质量的大小，通常选用聚合度 X，它是聚合物的相对分子质量 M 除以链节（聚合物中重复的结构单元）的相对分子质量 M_0，即分子中所含链节的数目：$X = M/M_0$。聚合过程中，由于所有的链节不可能增长到相同的长度，所以聚合物中相对分子质量的分散性很大，呈现类似图6.8-2的分布。若设 N_i 为相对分子质量的 M_i 的分子占据总分子数的百分数，则按分子数分布的统计平均，或数均相对分子质量 M_c 为

$$M_c = \sum_{i=1}^{\infty} N_i M_i,\ \sum_{i=1}^{\infty} = 1 \tag{6.8-1}$$

因此，平均聚合度为：$\overline{X} = M_c / M_0$。

若令 W_i 是相对分子质量为 M_i 的分子所占质量分数，则质均相对分子质量 M_w 为：

$$M_w = \sum W_i M_i ; \sum W_i = 1 \tag{6.8-2}$$

质均值和数均值是不同的，两者均可由实验测定。

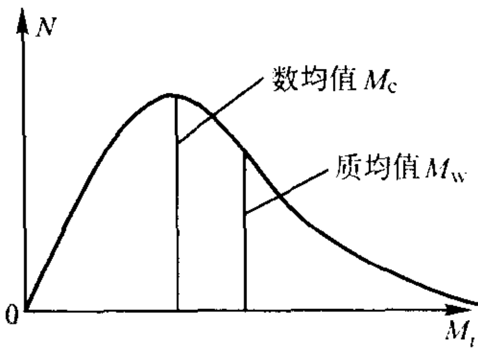


图6.8-2 相对分子质量分布曲线

2.1.3 高分子的聚集态结构

聚合物内大分子间的排列构成聚集态结构。普通物质的聚集态可分为气体、液体和固体，但聚合物内分子很大，因而分子间作用力也大，很容易聚集成固体或高温熔体，进一步提高温度，聚合物便分解，不存在气体。

聚合物分为晶体和非晶体两类。无规聚合物通常难以结晶，在实际的冷却速度下进入非结晶的玻璃态；规整聚合物易于结晶，只能通过速冷，使之成为"冻结"的过冷液体而进入玻璃态。一般情况下，规整聚合物由晶区和非晶区混杂组成。晶区是分子链结构有规则地来回折叠以薄片形式构成的晶体组成，通过高温退火或提高结晶温度都能使晶体尺寸增大。如聚乙烯晶体在100℃时晶体厚度约10 *nm*，而在130℃加热几小时后，其厚度可增加到40 *nm*，要把大分子链全部规则地排列起来是困难的，所以结晶聚合物都是部分结晶，同时存在晶区和非晶区（图6.8-3）。晶区所占质量分数称为结晶度，如涤纶、尼龙等结晶度为10%～40%。具有对称或立构规整链的聚合物（如聚乙烯或等规立构聚丙烯）一般是可结晶的，具有不规整主链或无规排列侧基的聚合物（如丁二烯—苯乙烯或无规聚甲基丙烯酸甲酯）是不可结晶的。结晶聚合物的有序度（结晶度）视其可结晶性和加工历史而定。在非晶态聚合物中，分子链也不是完全无规则的，而是具有不同程度短程有序，即分子链中的一段是排列有序的。

晶区的分子排列较紧密，增强了分子间的作用力，因而使聚合物的强度、硬度、刚度、相对密度、熔点等性能提高，但与链运动相关的性能，如弹性、延伸率等下降。

聚合物晶体中也存在类似于其他晶体中存在的缺陷：点缺陷、位错以及其他缺陷。这类缺陷都与高分子链的性质有关，如点缺陷往往出现在链的端部、短的支链、链折叠处和分子链间缠结处等。

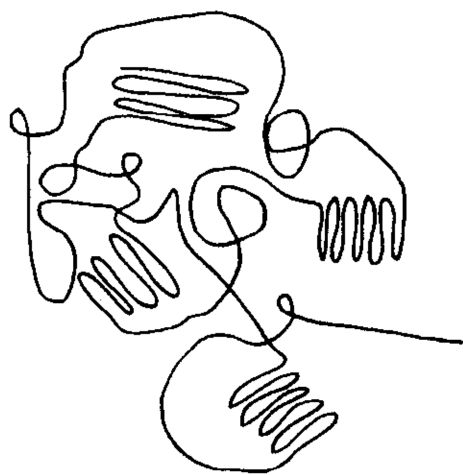


图 6.8-3　聚合物的晶区与非晶区示意图

2.1.4　聚合物的力学状态与转变

聚合物的物理性质依赖于温度和时间。随着温度的变化，在一些特定的温度，某些热力学性质会发生突然的变化，这种现象称为转变，转变时的温度成为转变温度。

在 $t=0$ 时给定阶跃应变 ε_0，然后在 t 秒时测量 $\sigma(t)$ 值，在不同温度 T 下重复上述试验，便得出 t 秒时的松弛模量 $E(t)=\sigma(t)/\varepsilon_0$ 对 T 的曲线。由于聚合物的力学性能是时间相关的，因此 t 取不同值时得到的 $E(t)-T$ 曲线也不同，通常取 $t=10$ s 或 5 s。图 6.8-4 表示非晶态聚合物的 $E(t)-T$ 曲线的一般特征。由图明显可见具有不同性质的五个区域。

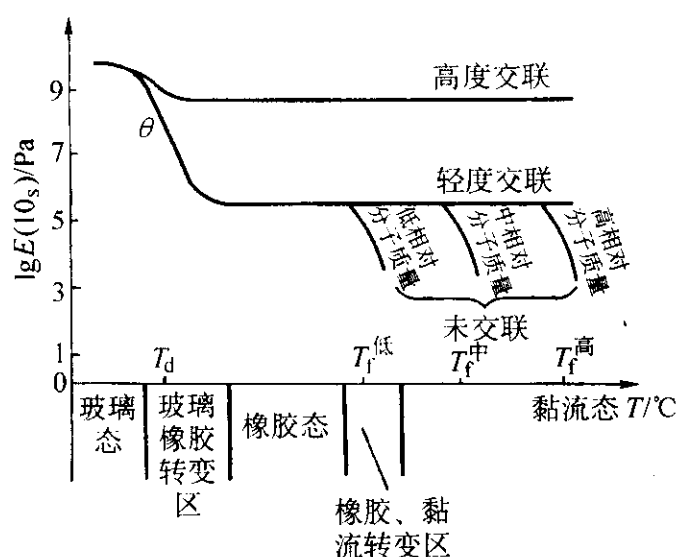


图 6.8-4　聚合物的 10 s 模量曲线

1）玻璃态　此时高分子链的整链和链段都被冻结，不能运动，只能发生高分子链角和键长的变化，在外力作用下的变形是瞬时的，E（10 s）随温度的下降很小，$E=E_g$，其值约为 $10^{9.5}$ N/m² 的量级。普通塑料室温时便处于这种状态，可以保持一定的几何形状，且有一定的承载能力。

2）玻璃-橡胶转变区　这一区域高分子链节的振幅加大，键的内旋转开始，这一转变区的宽度约 5～20℃，松弛模量可变化几个数量级，黏弹性的特征表现特别明显。这一转变的特征温度是玻璃化转变温度 T_g。这一转变区也用曲线的别点温度 $T_i[E(10\text{ s})]$ 约为 10^8 N/m² 处和在该点的负斜率 $-\tan\theta$ 来表征，通常 T_i 和 T_g 仅相差几度。

3）橡胶态（高弹态）　这一区域松弛模量下降到一新的平台，$E=E_r$，其值约为 $10^{5.5}$ N/m² 量级。此时高分子键的内旋转正常进行，链段发生运动，使构像发生变化，短程扩散运动远比观察时间为快，但整个分子链的长程运动还未开始。在此区域内材料具有高弹性，受力后伸长率可达 100%～1000%，卸载后可恢复，正常使用的橡胶便处于这种状态。

4）橡胶-黏流转变区　模量再次下降，约为 $10^{4.5}$ N/m² 量级，分子链段的运动规模加大，致使整个分子链开始运动。这一转变的特征温度是黏流温度 T_f。随着分子量的增大，T_f 值升高；分子量分散性增大时，转变区亦增宽。

5）黏流态　此时整个分子链产生运动，分子链间的阻力再也不能阻止流动。黏流态适合于聚合物的加工成形。

有些非晶态聚合物的分解温度低于 T_f，如纤维和聚丙烯腈等，所以没有黏流态。而有些热固性塑料，如酚醛树脂，由于交联程度很高，只存在玻璃态。

2.2　聚合物的时间效应和时－温等效原理

聚合物的力学性能随时间变化的现象称为松弛，具有黏弹性的特征。转变和松弛是相互联系的现象，都与聚合物内部的分子运动状态相关，是其宏观表现形式，是在宏观结构不变的情况下，分子运动从一种模式转变到另一种模式的表现。聚合物的转变受加热、冷却和应变速率的影响而变动，与热力学的热平衡相变不同，是一种非平衡热力学状态的转变。

聚合物的力学性质是温度和时间的函数，在一维松弛试验中有

$$\sigma(T,t)=E(T,t)\varepsilon_0 \tag{6.8-3}$$

式中，ε_0 为 $t=0$ 时施加的阶跃应变；$E(T,t)$ 为松弛模量。

2.2.1　时间响应

$t=0$ 时聚合物受到阶跃应变 ε_0 作用，高分子键内的键角和键长立即发生变化，引起瞬时应力，这时 $\lg E-\lg t$ 图上的曲线和 $\lg E-T$ 曲线上的玻璃态对应。经过一段时间后，卷曲的高分子链通过链段运动逐步舒张，高弹变形逐步增加，应力不断下降，这和玻璃态到橡胶态的转变区对应。时间进一步增长后，高分子链间发生相互滑移，整链发生运动，产生黏性流动，时间非常长以后，由于黏性流动，应力可降到零，这对应于黏流态。如果聚合物存在交联，应力可以不松弛到零，这对应于不发生黏流态的情形。

在动态应力下，高应变率对应于玻璃态，低应变率对应于橡胶态，中应变率对应转变区，对无交联的聚合物，应变率无限小时对应于黏流态。中应变率时粘弹性效应最为明显。

2.2.2　时间－温度等效原理

比较聚合物的模量对时间的响应和对温度的响应，发现两者存在对应关系，因为两者都和高分子链的运动形式密切相关。缩短实验时间或提高频率对模量的影响相当于降低温度的影响。实验发现，对非晶态聚合物，在各种温度下所得到的模量松弛数据，都可通过时间标度的适当移动而叠加在一起，这种普遍的现象称为时温等效原理。这一理论认为在温度 T_1 和 T_2 测得的松弛模量可按下式转换

$$\frac{E(T_1,t)}{\rho_1 T_1}=\frac{E(T_2,t/a_t)}{\rho_2 T_2} \tag{6.8-4}$$

式中，T_1 和 T_2 是热力学温度；ρ_1 和 ρ_2 分别为 T_1 和 T_2 时的密度，当 T_1 和 T_2 相差不太大时，$\rho_1 T_1$ 和 $\rho_2 T_2$ 的差别是不大的；a_t 为时间移动因子，只是温度的函数，它代表在 $\lg t$ 坐标轴上的位移：令 $t_2=t_1/a_t$，则

$$\lg(t_2)-\lg(t_1)=-\lg a_t \tag{6.8-5}$$

因此，如要绘制某一温度下的长时期的模量－时间曲线时，可选用该温度为参照温度 T_0，然后在许多温度 T 下作短期试验，测得相应 $E(T,t)$，便可采用式（6.8-5）换算

$$E(T_0,t)=(\rho_0 T_0/\rho T)E(T,t/a_t) \tag{6.8-6}$$

Williams、Landel 和 Ferry 发现，上式的 a_t 可表成下列 WLF 方程

$$\lg a_t=\frac{-C_1(T-T_0)}{C_2+(T-T_0)} \tag{6.8-7}$$

实用上常取 T_g 或室温为 T_0。取 T_g 时某些聚合物的 C_1 和 C_2 值示于表 6.8-3。选其他温度为 T_0 时，C_1 和 C_2 的取值，

可由实验确定。

表 6.8-3 WLF方程中的参数（T_g 为参照温度）

聚合物	$\frac{T_g}{K}$	C_1	C_2	备 注
聚苯乙烯	373	13.7	50.0	
聚异丁烯	205	16.6	104.4	
天然橡胶	200	16.8	53.6	
聚1，4-丁二烯	161	11.3	60	顺/反/乙烯基=
丁苯橡胶	210	20.3	25.6	96.5/1.9/1.6
聚甲基丙烯酸甲酯（无规立构）	381	34.0	80	苯乙烯/丁二烯=
聚乙基丙烯酸甲酸	335	17.6	65.5	23.5/76.5（重）
"普适"常数		17.44	51.6	

2.3 聚合物的力学性能

1960年以前，聚合物的屈服现象未引起人们的重视，把屈服看成是由于材料局部变形引起温升而产生的软化现象；20世纪60年代以来，人们认识到屈服是聚合物的一种力学行为，可应用现有经典的塑性理论来处理；同时，观察到聚合物的"滑移带"和"缠结带"，以及和金属不相同的屈服现象。聚合物的应力-应变曲线依赖于时间和温度，还依赖于其他因素，由于试验条件的不同，可以表现出不同的力学性能。

2.3.1 拉伸应力-应变曲线

图6.8-5a表示两种非晶态聚合物聚氯乙烯和聚苯乙烯的拉伸应力-应变曲线，图7.8-50b表示三种晶态聚合物聚四氟乙烯、聚乙烯和聚三氟氯乙烯的拉伸应力-应变曲线。从这两张图可以见到几种变形方式。

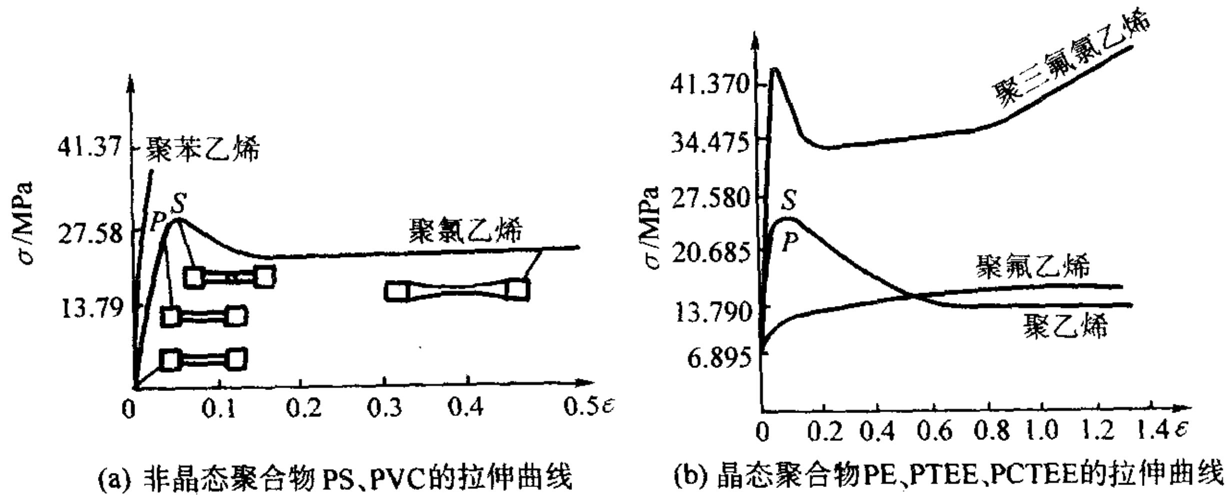


图 6.8-5 非晶态聚合物PS、PVC的拉伸曲线和晶态聚合物PE、PTFE、PCTEE的拉伸曲线

1）聚苯乙烯是一种硬而脆的材料，具有高的模量和强度，断裂伸长率很低，没有明显的屈服。

2）聚氯乙烯、聚四氟乙烯和聚三氟氯乙烯，属于强而韧的材料，这是处于玻璃态的塑料在某一段温度和速度范围内典型的弹塑性变形应力应变曲线。图6.8-5中的 P 点称为弹性极限，S 点为屈服强度，极限强度图中未画。在 P 之前，应力和应变是线性关系，变形是由于分子链中的键长和键角的变化引起的可恢复的弹性变形。从 P 点到 S 点的变形属强迫高弹变形，因此时外力足够大，能克服分子链段运动的壁垒，发生链段的强迫运动；除去外力后，链段不能运动，因而高弹形变被固定下来，成为永久变形。聚合物在 T_g 温度以下，随着温度的降低，链段间的相互作用力增强，因而产生强迫高弹性变形的应力也要增加。但是当温度降到脆化温度 T_β 值以下，强迫高弹形变所需的应力便超过了聚合物的断裂强度，此时外力使聚合物断裂，而不能产生高弹形变。所以实际上强迫高弹形变只能在 T_g 和 T_β 之间产生，T_β 是塑料可使用的最低温度。在 S 点之后，链段沿外力方向运动的同时，发生分子链间的滑动，此时外力几乎不增加而应变增加很快，材料处于塑性区。材料应力在到达屈服应力 σ_s 之前，变形是均匀的，达到 σ_s 后，开始出现不均匀形变，沿试样的某些特殊点上开始颈缩。随后，颈缩区的局部变形增加，非晶态区的链构像发生变化和明显的再取向，颈缩区的强度和刚度增加，如取向强化而增加的屈服应力超过因面积减小而增加的应力，则颈缩区的截面积停止缩小，且以比初始直径稍小的直径稳定下来，以不变的名义应力继续变形过程，这一名义应力称为屈服下限应力或冷拉应力，颈缩便沿试样的长度方向传播，这一细颈伸展的过程称为冷拉。但如取向强化导致的屈服应力升高不足以抵消颈缩区因面积减小而增加的应力，那末颈缩区的截面积将进一步减小，最终形成局部断裂。

虽然非晶态和晶态聚合物的拉伸曲线外形相似，但存在重要差别。首先非晶态聚合物拉伸只发生分子链的取向变化，而晶态聚合物却包含结晶的破坏、取向和再结晶等过程，其次，非晶态的冷拉区在 $T_g \sim T_\beta$ 之间，而晶态的冷拉区却在 $T_g \sim T_m$ 之间。

3）聚四氟乙烯的拉伸曲线，没有局部颈缩，变形过程中横截面均匀地减小。

上面的三点简单地讨论了聚合物的四种变形方式，其中的屈服点很难给以确切定义。如拉伸曲线上应力不出现极大值，则定义2%应变处的应力为屈服应力。聚合物断裂时的伸长率可以从某种聚苯乙烯的1.5%变化到某种聚乙烯的500%或更大。

随着试验条件的改变，聚合物试样可表现出不同的变形方式。图6.8-6表示试验温度对聚甲基丙烯酸甲酯（PMMA）应力-应变曲线的影响。低温时硬而脆，40℃出现颈缩局部断裂，60℃出现冷拉现象，等等。图6.8-7表示变形速度对聚氯乙烯（PVC）应力-应变曲线的影响，变形速度愈快，断裂应变愈小，愈呈现脆性性质。

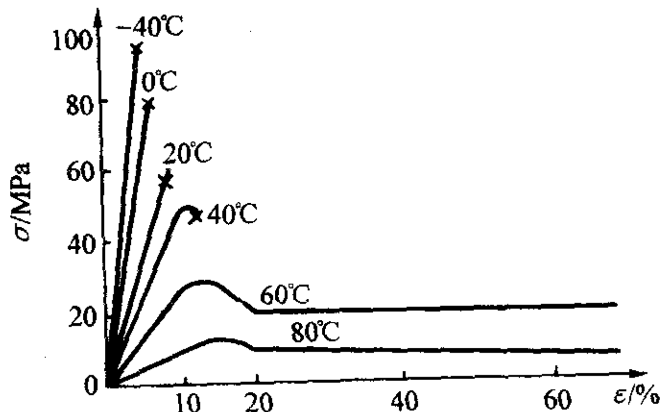


图 6.8-6 试验温度对聚甲基丙烯酸甲酯应力-应变曲线的影响

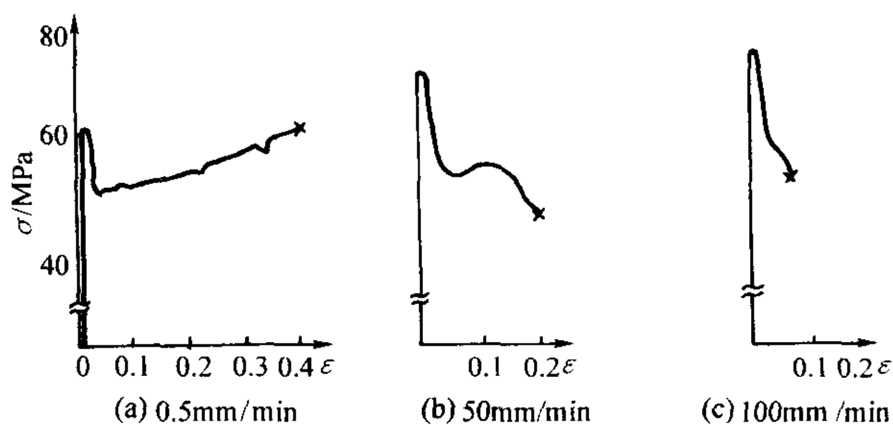


图 6.8-7　拉伸速度对聚氯乙烯应力-应变曲线的影响
（试验温度 20℃）

2.3.2　压缩应力 – 应变曲线

图 6.8-8 表示两种典型的非晶态聚合物 PVC 和乙酸纤维素和两种典型的结晶聚合物聚四氟乙烯（PTFE）和聚三氟氯乙烯的压缩应力-应变曲线。两种非晶态聚合物应力有极大值，此即屈服应力，且显示了真正的软化；反之，两种晶态聚合物应力并无明显的极大值，因而以 2% 应变处的应力为屈服应力。对同一聚合物，比较一下拉伸和压缩屈服应力可知，一般讲，压缩屈服应力比拉伸屈服应力要高 20% 左右，即拉压屈服强度是不同的，这就表明流体静应力对屈服强度的影响。

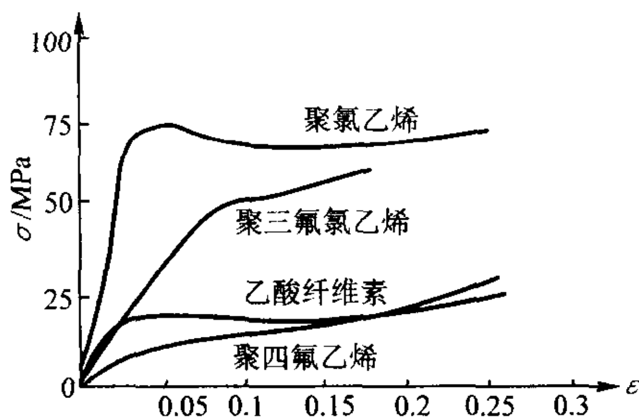


图 6.8-8　几种聚合物的压缩应力 – 应变曲线

在拉伸时表现为脆性的一些热固性塑料，在压缩时却有相当大的塑性流动。图 6.8-9 表示酚醛树脂铸件压缩时的应力 – 应变曲线，图中有一屈服点，断裂应变可达 40% 左右；而在拉伸情况下，应变约为 2% 时便断裂，且无明显的屈服点。

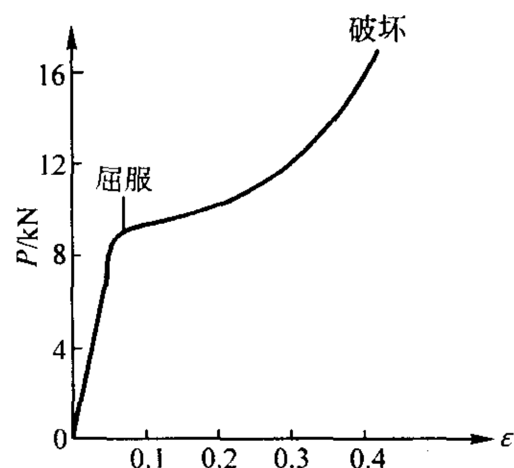


图 6.8-9　酚醛树脂的压缩应力 – 应变曲线

2.3.3　拉伸和压缩时的力学性能数据

目前有众多的均聚物、带添加剂的均聚物、共聚物、嵌段共聚物、增强的聚合物等，很难用一张简单的表格来评价这些材料。然而，了解室温下按测试标准测得的聚合物的力学性能范围，对了解哪些结构特征能增加强度、刚度和韧性是有益的。

表 6.8-4 给出某些塑料的拉、压力学性能数据，由于压缩试验难以做到断裂，故取用屈服应变作指标，而拉伸则用断裂应变。从表中大体可以看出如下几种。

1）对非极性聚合物，如聚乙烯和聚四氟乙烯的弹性模量值较低，但韧性较好。聚丙烯中由于结构单元中甲基的存在增加了位阻，因而强度和刚度要高些。

2）像在聚氯乙烯中用氯取代氢一样，聚三氟氯乙烯中用氯取代氟，增强了极性力，因而强度和模量都有较大的提高。而像尼龙、聚甲醛和聚碳酸酯等强极性聚合物，则具有良好的刚度、强度和韧性。

3）像聚苯乙烯和聚甲基丙烯酸甲酯等非晶态聚合物，庞大的极性的侧基使分子的刚性较大，因而强度和刚度较高，但韧性较差。

4）在热固性塑料中增加合适的填料或加入玻璃纤维等组成复合材料，可得高强度和高模量的材料。

表 6.8-4　某些塑料的拉伸和压缩力学性能

聚合物	弹性模量 /MPa	拉伸屈服强度 /MPa	抗拉强度 /MPa	拉伸断裂伸长率 /%	压缩屈服应力 /MPa	压缩屈服应变 /%
低密度聚乙烯	138 ~ 276	6.8 ~ 13.6	10.2 ~ 17.2	400 ~ 700		
高密度聚乙烯	414 ~ 1 035	17 ~ 34	17.2 ~ 37.4	100 ~ 600	20.4 ~ 34	
聚四氟乙烯	414	10.2 ~ 13.6	13.6 ~ 27.2	100 ~ 350	10.2 ~ 13.6	
聚三氟氯乙烯	1 035 ~ 2 070	27.2 ~ 34	30.6 ~ 40.8	80 ~ 250	34 ~ 54.4	8
聚丙烯	1 035 ~ 1 552	20.4 ~ 27.2	23.8 ~ 37.4	200 ~ 600	34 ~ 54.4	
尼龙 – 66	1 242 ~ 2 760	57.8 ~ 78.2	61.2 ~ 81.6	60 ~ 300	54.4 ~ 88.4	4 ~ 20
聚碳酸酯	2 415	54.4 ~ 68	54.4 ~ 68	60 ~ 120	68 ~ 81.6	
聚甲基丙烯酸甲酯	2 415 ~ 3 450	47.6 ~ 61.2	47.6 ~ 68	2 ~ 10	74.8 ~ 95.2	5 ~ 10
聚甲醛	2 760	47.6 ~ 54.4	61.2 ~ 68	20 ~ 80	68 ~ 108.8	
聚苯乙烯	2 760 ~ 3 450		37.4 ~ 54.4	1 ~ 2.5	74.8 ~ 108.8	4 ~ 6
硬聚氯乙烯	2 070 ~ 4 140	54.4 ~ 68	40.8 ~ 74.8	5 ~ 60	68 ~ 74.8	5 ~ 6
酚醛树脂铸件	2 760 ~ 3 450		40.8 ~ 61.2	1.5 ~ 2	95.2 ~ 122.4	4 ~ 6
矿物填充的酚醛塑料	6 900 ~ 13 800		27.2 ~ 47.6	0.2 ~ 0.5	102 ~ 204	4 ~ 6
纤维填充的三聚氰胺甲醛塑料	8 280 ~ 11 040		40.8 ~ 61.2	0.4 ~ 0.6	170 ~ 217.6	
玻璃填充的环氧树脂	20 700		102 ~ 408	3 ~ 4	204 ~ 476	

2.3.4 复杂应力状态下的屈服

聚合物的拉伸和压缩屈服应力是不相同的。Whitney 和 Andrews 研究了聚苯乙烯、聚甲基丙烯酸甲酯、聚碳酸酯和聚乙烯醇缩甲醛的屈服行为（图 6.8-10）。由图可知，采用适合普通金属的 Mises 或 Tresca 准则是不合适的，可选用 Mohr-Coulomb 准则。Mohr-Coulomb 准则可简单地写成

$$\frac{1}{2}(\sigma_1-\sigma_2)+\frac{1}{2}(\sigma_1+\sigma_2)\sin\varphi=\tau_c\cos\varphi \quad (6.8\text{-}8)$$

式中，φ 和 τ_c 由实验确定。设 σ_t 和 σ_c 分别为拉伸和压缩时屈服应力的绝对值，由式（6.8-8）知

$$\sin\varphi=(\sigma_c-\sigma_T)/(\sigma_c+\sigma_T),\tau_c=\sqrt{\sigma_T\sigma_c/2}$$

另一可选用的准则是修正的 Mises 准则，它是 Mohr-Coulomb 不对称六角形的外接椭圆，即

$$\sigma_1^2+\sigma_3^2-\sigma_1\sigma_3+(\sigma_1-\sigma_3)(\sigma_c-\sigma_T)=\sigma_c\sigma_T \quad (6.8\text{-}9)$$

实验表明，修正的 Mises 准则和实验符合得更好。

对于取向聚合物，力学性质是各向异性的，需要采用各向异性体的屈服理论。

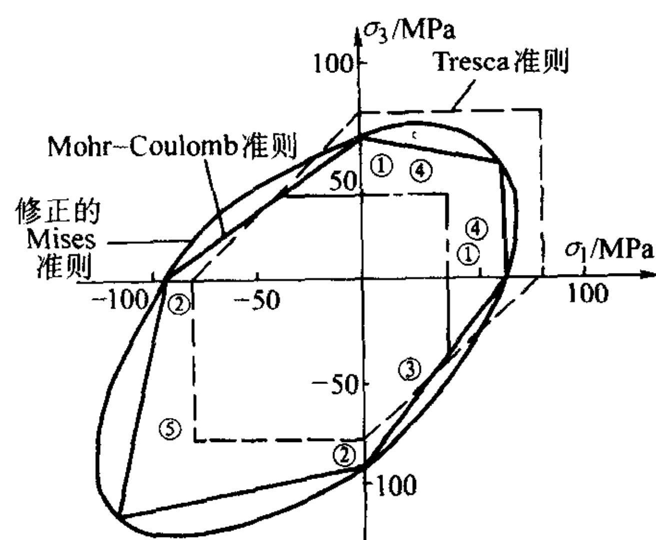


图 6.8-10 聚苯乙烯在 $\sigma_1\sigma_3$ 平面内的屈服

① 单轴拉伸；② 单轴压缩；③ 扭转；④ 双轴拉伸；⑤ 双轴压缩

2.3.5 聚合物的银纹和断裂

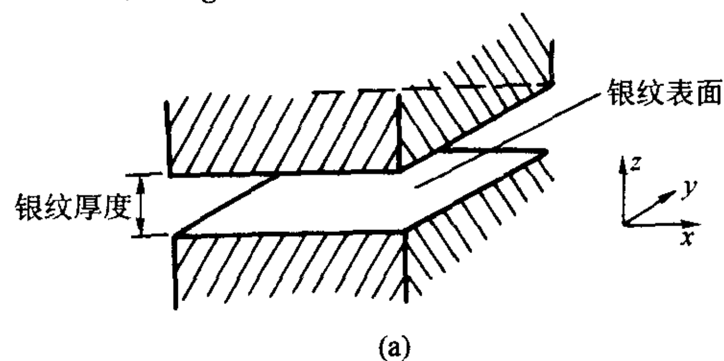
聚合物是黏弹性或黏弹塑性材料，其力学性能与时间明显相关，因而其断裂韧性不是常数，而是时间的函数，即有 $K_c(t)$、$j_c(t)$ 等。在聚合物的断裂过程中，不仅屈服引起能量耗散，而且孔洞的形成和发展也耗散能量。聚合物孔穴化的一个特有现象是产生银纹，银纹不仅在裂端产生，还以其他方式存在。

银纹（craze），原意是陶瓷表面上的细微裂纹。用于高分子材料科学中，它描述的是一种与裂纹（crack）相类似并且与之有密切联系但又不是裂纹的结构。简单地说，裂纹中不含有任何高分子材料，而银纹中却仍然有大约 30% ~ 50% 体积分数的高分子材料。取向的高分子以微纤束的形式维系着银纹的两银纹面。高分子材料中产生银纹的过程，叫银纹化（crazing）。

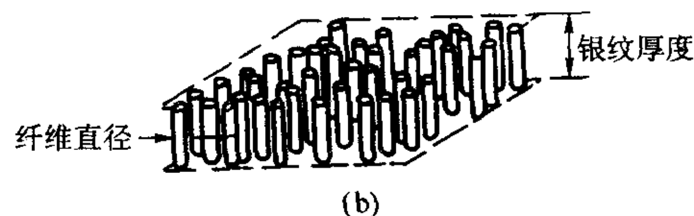
银纹现象是高分子材料特有的，它是材料宏观破坏以前微观上损伤、破坏的开始。在高分子材料的断裂（尤其是脆性断裂）、蠕变、环境应力开裂以及疲劳破坏中，银纹都具有十分重要的作用；另一方面，银纹又是高分子材料橡胶增韧的力学基础。正确认识银纹现象，有利于合成新的高分子材料的分子设计。

银纹可以在高分子材料的表面或内部形成。它是一种扁平的细长结构，外形上和裂纹类似，但不相同；银纹看起来好象在拉紧的表面上有一片细微的薄雾，其中银纹质和孔穴约各占体积的 50%，孔穴是连通相，银纹质纤维的直径约为 5 ~ 20 nm。银纹厚约为 1 ~ 2 μm，长几百微米（图 6.8-11）。由于银纹密度比基体密度小，从而因光线的反射和散射，肉眼可见；其内部的精细结构，需用电子显微镜研究。随着变形程度的增加，银纹随着生长，一定程度后发生破坏（图 6.8-12）。有机溶剂，甚至手上的汗渍都能促使银纹的引发。在有机玻璃、聚苯乙烯、聚碳酸酯之类的玻璃状高分子中，所产生的银纹由于银纹面反射光而特别容易被观察。

图 6.8-13 是 Krgner 用透射电镜测定的银纹面间的位移和



(a)



(b)

图 6.8-11 垂直于银纹面取向的圆筒状纤条示意图

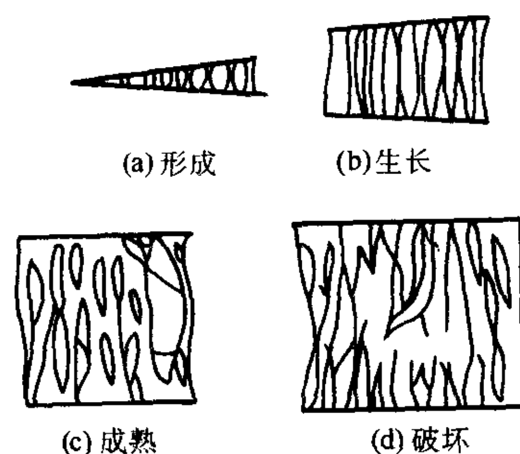


图 6.8-12 聚苯乙烯薄膜银纹在不同阶段的形状变化示意图

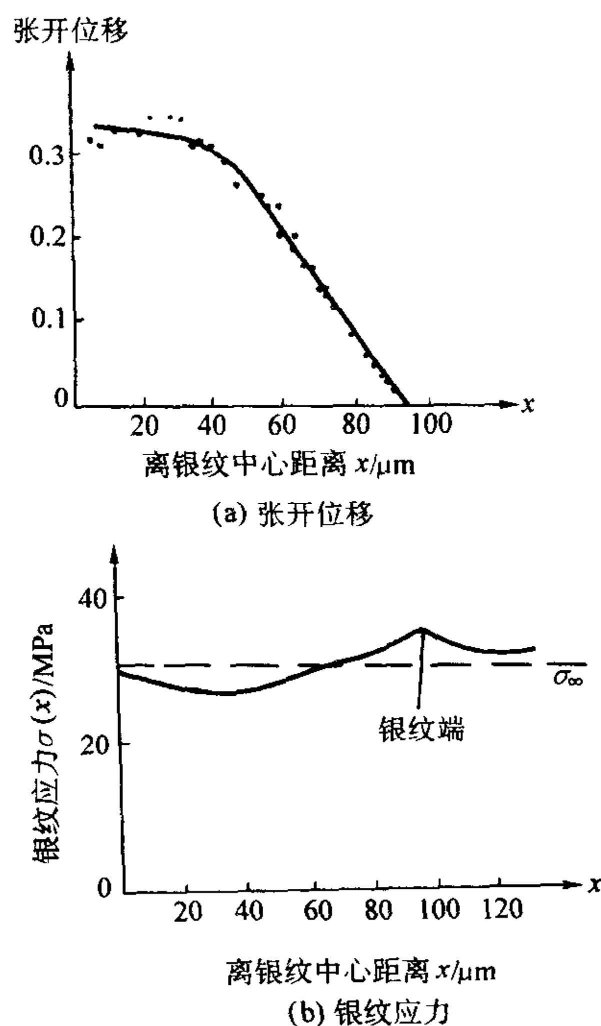


图 6.8-13 TEM 技术测定的聚苯乙烯空气银纹的张开位移和银纹应力

用弹性力学方法换算出的银纹应力。图上表明，银纹上的应力不是均匀分布的，应力集中在银纹端部。

目前认为银纹引发的最大拉应变准则较好，即

$$E\varepsilon_1 = \sigma_1 - \nu(\sigma_2 + \sigma_3) = A + B/(\sigma_1 + \sigma_2 + \sigma_3) \tag{6.8-10}$$

式中，A 和 B 是与温度和时间相关的常数。但也有实验支持，银纹引发由最大拉应力决定，而与平均压力关系不大。

对银纹的研究所涉及的内容是多方面的。大致可从以下几方面考虑。

1）银纹主要是在非晶态玻璃状高分子中产生。但某些结晶性高分子，例如在低温时变形的聚丙烯和尼龙中也能产生。热固性环氧树脂中也能产生类银纹结构，虽然这是在形式上和玻璃状高分子中的银纹相似，而微观结构尚不清楚。

2）银纹能在材料表面、内部和裂纹端部形成。在裂纹端部形成的银纹，只是裂纹端部塑性屈服区的一种形式。

3）在单一应力作用下引发的银纹，称为应力银纹（stress craze），受应力和溶剂联合作用引发的银纹，称为应力-溶剂银纹（stress-solvent craze）。溶剂能加速银纹的引发和生长。

4）银纹对高分子材料的蠕变和疲劳破坏也有重要影响。高分子在长期蠕变变形中能引发蠕变银纹。从根本上讲，疲劳裂纹的扩展，就是裂纹端部银纹的扩展过程。

2.3.6 聚合物的疲劳

由于聚合物是黏弹性材料，在循环载荷下将产生可观的能量耗散，同时因聚合物是热的不良导体，所以在裂尖处产生显著的温升；例如有试验报导，在应力强度因子变化范围 $\Delta K_1 = 3.2\ \text{MPa}\sqrt{\text{m}}$ 作用下，高抗冲尼龙裂纹尖端温度上升约100℃，这就影响到材料的变形方式。

在疲劳断口上通常出现两种条纹形态。一种是疲劳纹，纹间间距约 10 μm，对应每一周期交变应力作用的疲劳纹的扩展，在周期应力达极大值的瞬间，裂纹尖端银纹根部张开位移最大，产生断裂，随后裂纹和银纹以相同的速度向前扩展一个疲劳纹间距，这一间距通常小于银纹区长度。另一种是斑纹，纹间间距约 50 μm；在交变应力周期数达到一定值时，裂纹会产生一次突跃而形成斑纹结构，例如有实验指出，在 0.6 Hz 的低频作用下，聚氯乙烯试样大约需 370 周后才突跃一次。实验表明，低相对分子质量和低应力水平有利于形成斑纹，而高相对分子质量和高应力水平有利于形成疲劳裂纹。

聚合物疲劳裂纹扩展速率 $\mathrm{d}a/\mathrm{d}N$ 和应力强度因子幅 ΔK_1 的关系，一般服从 Paris 公式

$$\mathrm{d}a/\mathrm{d}N = A\Delta K_1^n \tag{6.8-11}$$

式中，A 和 n 是依赖于温度、加载频率和平均应力等的材料常数。图 6.8-14 给出部分聚合物 $\mathrm{d}a/\mathrm{d}N$-ΔK_1 的关系曲线。

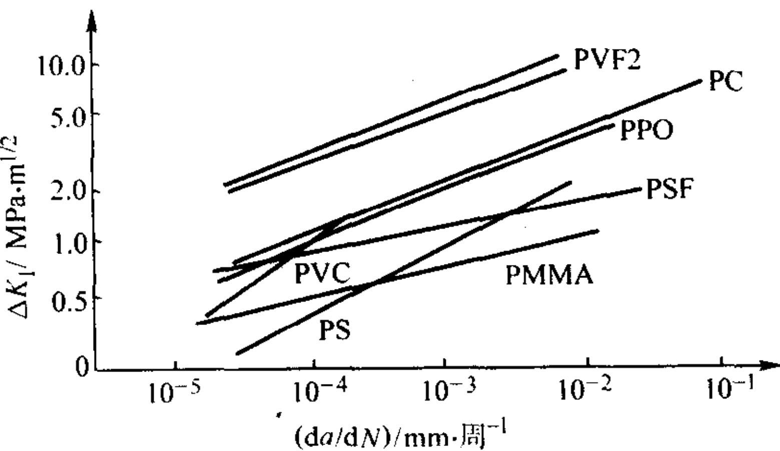


图 6.8-14 部分聚合物疲劳裂纹扩展速率 $\mathrm{d}a/\mathrm{d}N$ 与 ΔK_1 的关系

2.3.7 橡胶颗粒增韧

影响聚合物断裂韧性的因素很多，相对分子质量的大小和分布、支化、交联、取向、结晶度、晶粒尺寸、共聚和共混等均有影响；凡是能提高聚合物的强度和塑性变形能力的因素，均能提高断裂韧性。通常还用橡胶颗粒来提高脆性聚合物，如环氧树脂和聚苯乙烯等的韧性。含有恰当百分比橡胶颗粒的脆性聚合物受到外来冲击时，便会在颗粒的赤道面上引发许多银纹，这便消耗掉相当一部分能量；同时，颗粒又起到终止银纹扩展的作用，从而起到增韧作用。这是聚合物银纹—裂纹断裂机理在实际中成功应用的例子。

3 陶瓷材料

3.1 陶瓷材料的弹性性能

陶瓷材料为脆性材料，在室温下承载时几乎不能产生塑性变形。而在弹性变形范围内就产生断裂破坏，因此，其弹性性质就显得尤为重要。与其他固体材料一样，陶瓷的弹性变形可用胡克定律来描述。

陶瓷的弹性变形实际上是在外力的作用下原子间距由平衡位产生了很小位移的结果。这个原子间微小的位移所允许的临界值很小，超过此值，就会产生键的断裂（室温下的陶瓷）或产生原子面滑移塑性变形（高温下的陶瓷）。弹性模量反映的是原子间距的微小变化所需外力的大小。影响弹性模量的重要因素是原子间结合力，即化学键。表 6.8-5 给出一些陶瓷在室温下的弹性模量。

表 6.8-5 陶瓷的弹性模量数据

材料	E/GPa	材料	E/GPa	材料	E/GPa
金刚石	1 000	玻璃	35 ~ 45	SiO_2	94
WC	400 ~ 650	C_f	250 ~ 450	NaCl，LiF	15 ~ 68
TaC	310 ~ 550	AlN	310 ~ 350	$MgAl_2O_3$	240
WC-Co	400 ~ 530	$MgO\cdot SiO_2$	90	BN	84
NbC	340 ~ 520	Al_2O_3	390	MgO	250
SiC	450	BeO	380	多日石墨	10
ZrO_2	160 ~ 241	TiC	379	TiO_2	29
莫来石	145	Si_3N_4	220 ~ 320	$MgAl_2O_4$	240

弹性常数是表征材料弹性的重要物理参数，它受到晶体结构、晶格振动等微观因素所制约，是材料内部原子间结合力的一种量度，从图 6.8-15 所示的原子间结合力与原子间距的关系曲线看出，弹性模量 E 是曲线在平衡原子间距 r_0 点上切线斜率的函数。显然，原子间距的不同将反映在弹性模量的差异。提高温度将导致原子间距的膨胀，材料的 E 值亦会相应下降。压力和张力引起的弹性模量增减在陶瓷材料中并不明显，因为这里的断裂应变极限很小，原子间距变化的差别极微。

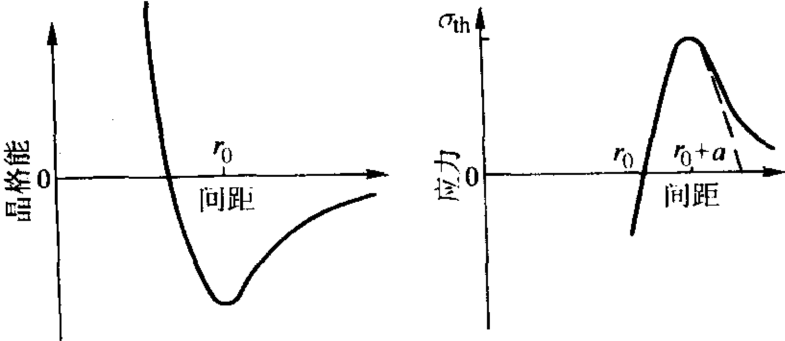


图 6.8-15 晶格能和应力随原子间距的变化

3.1.1 各向同性陶瓷的弹性模量

多晶陶瓷材料多数由各向异性的晶粒构成，微观上是各向异性，但宏观上可看作各向同性体。其弹性行为遵循着胡克定律，可用弹性模量 E 和切变模量 G 分别描述正应变 ε

和切应变γ与应力之间的关系：

$$\left.\begin{aligned}\varepsilon_x &= \frac{1}{E}[\sigma_x - \nu(\sigma_y + \sigma_z)]\\ \varepsilon_y &= \frac{1}{E}[\sigma_y - \nu(\sigma_z + \sigma_x)]\\ \varepsilon_z &= \frac{1}{E}[\sigma_z - \nu(\sigma_x + \sigma_y)]\end{aligned}\right\} \tag{6.8-12}$$

$$\left.\begin{aligned}\gamma_{xy} &= \frac{\tau_{xy}}{G}\\ \gamma_{yz} &= \frac{\tau_{yz}}{G}\\ \gamma_{zx} &= \frac{\tau_{zx}}{G}\end{aligned}\right\} \tag{6.8-13}$$

$$G = \frac{E}{2(1+\nu)} \tag{6.8-14}$$

一般来讲，在其他性能允许的情况下，可以通过在一定范围内调整两相比例来获得所需的弹性模量值。

3.1.2 单晶体陶瓷弹性模量的各向异性

单晶体陶瓷在不同的晶向上往往具有不同的弹性模量。表6.8-6给出M20及石墨的弹性模量各向异性的例子。表中S为弹性柔度系数，C为刚度系数。表6.8-7给出各向同性材料各弹性模量及泊松比之间的关系。一般的陶瓷材料都是由很小的晶粒组成的多晶体，因此，整体材料表现出各向同性。但像MgO这种对称性高的晶体，其不同晶体学方向上的弹性模量相差很大，这在微观断裂力学分析时是要特别注意的。

表6.8-6 MgO与石墨的弹性常数

(1) 25℃时MgO的弹性常数

C_{11}	C_{12}	C_{44}	S_{11}	$-S_{12}$	S_{44}
GPa			MPa		
289.2	88.0	154.9	4.03	0.94	6.47

(2) 石墨单晶的弹性常数

刚度系数/GPa	柔度系数/$10^{-3}\cdot GPa^{-1}$
$C_{11} = 1\,060 \pm 20$	$S_{11} = 0.98 \pm 0.03$
$C_{12} = 180 \pm 20$	$S_{12} = -0.16 \pm 0.06$
$C_{13} = 15 \pm 5$	$S_{13} = -0.33 \pm 0.08$
$C_{33} = 36.5 \pm 1.0$	$S_{33} = 27.5 \pm 1.0$
$C_{44} = 4.5$	$S_{44} = 240$

(3) 25℃时MgO弹性常数的各向异性

结晶方位	E/GPa	G/GPa
⟨100⟩	248.2	154.6
⟨110⟩	316.4	121.9
⟨111⟩	348.9	113.8

表6.8-7 各向同性物质弹性常数之间的关系

弹性常数	E，ν	G，ν	E，G	G，K
拉伸弹性模量E	E	$2(1+\nu)G$	E	$\frac{9KG}{3K+G}$
切变模量G	$\frac{E}{2(1+\nu)}$	G	G	G
体积弹性模量K	$\frac{E}{3(1-2\nu)}$	$\frac{2G(1+\nu)}{3(1-2\nu)}$	$\frac{G\cdot E}{3(3G-E)}$	K
泊松比ν	ν	ν	$\frac{E}{2G}-1$	$\frac{3K-2G}{6K+2G}$

3.1.3 弹性模量的影响因素

(1) 温度对弹性模量的影响

由于原子间距及结合力随温度的变化而变化，所以弹性模量对温度变化很敏感。当温度升高时，原子间距增大，由d_0变为d_1（如图6.8-16），而d_1处曲线的斜率变缓，即弹性模量降低。因此，固体的弹性模量一般均随温度的升高而降低。图6.8-17给出一些陶瓷的弹性模量随温度的变化情况，一般来说，线胀系数小的物质，往往具有较高的弹性模量。

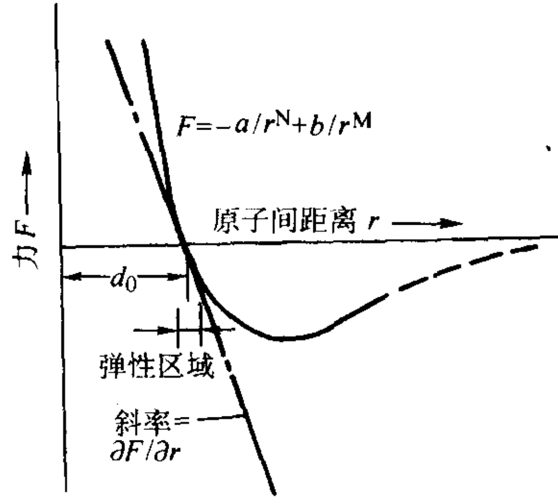


图6.8-16 原子结合力示意图

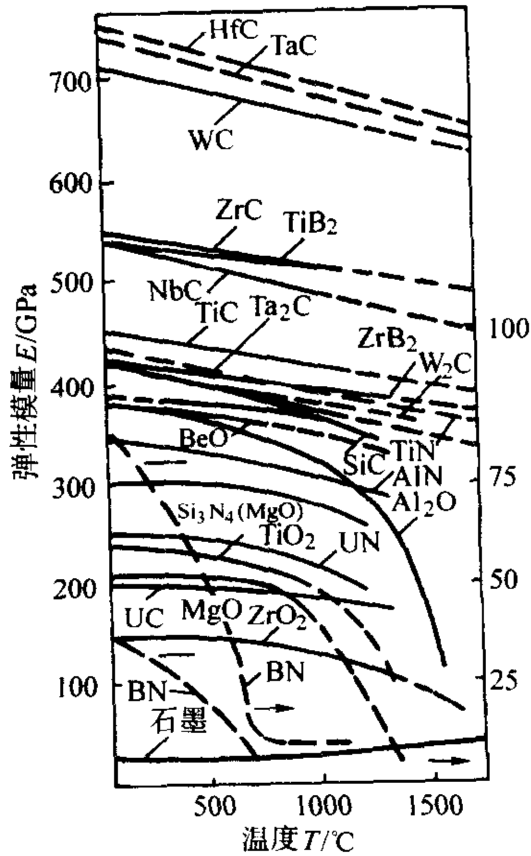


图6.8-17 温度对弹性模量的影响

(2) 弹性模量与熔点的关系

物质熔点的高低反映其原子间结合力的大小，一般来说，弹性模量与熔点成正比例关系。

例如，在300 K下，弹性模量正与熔点T_m之间满足如下关系

$$E = \frac{100kT_m}{V_a} \tag{6.8-15}$$

式中，V_a为原子体积或分子体积。图6.8-18为由Frost与Ashby总结出的E与kT_m/V_a之间的关系图，可以看出，它们符合良好的线性关系。不同种类的陶瓷材料弹性模量之间大体上有如下关系：氧化物<氮化物≈硼化物<碳化物。

(3) 弹性模量与材料致密度的关系

陶瓷材料的致密度对其弹性模量影响很大，弹性模量E与气孔率p之间满足下面关系式

$$E = E_0(1 - f_1 p + f_2 p^2) \tag{6.8-16}$$

式中，E_0为气孔率为0时的弹性模量；f_1及f_2为由气孔形状决定的常数。Mackenzie求出当气孔为球形时，$f_1 = 1.9$，

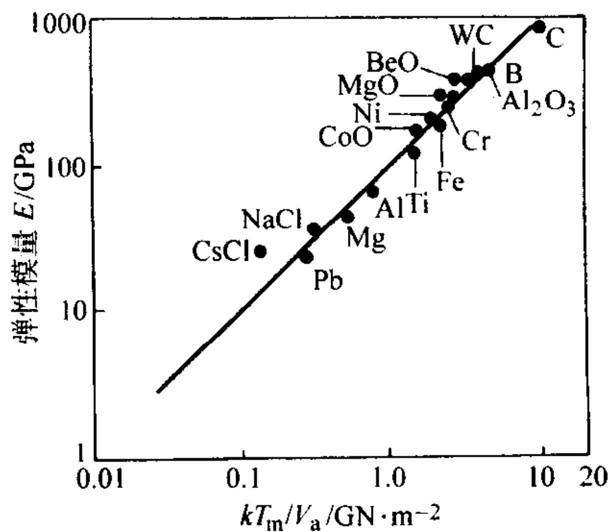


图6.8-18 弹性模量与 kT_m/V_a 之间的关系

$f_2=0.9$。图6.8-19给出 Al_2O_3 陶瓷的弹性模量随气孔率的变化及某些理论计算值的比较。Frost指出，弹性模量与气孔率之间符合指数关系

$$E = E_0\exp(-Bp) \tag{6.8-17}$$

式中，B 为常数。

总之，随着气孔率的增加，陶瓷的弹性模量急剧下降。

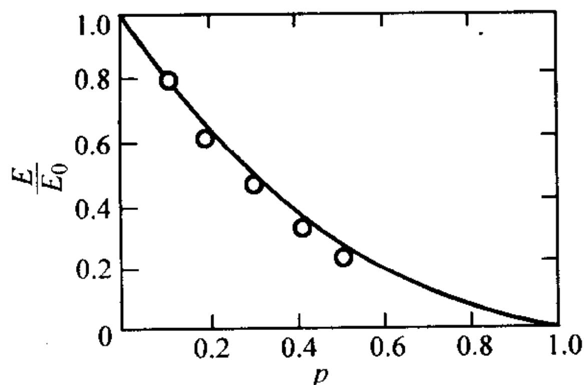


图6.8-19 气孔率对弹性模量的影响

3.2 陶瓷材料的强度及其影响因素

一般陶瓷材料在室温下的应力－应变曲线如图6.8-20中曲线1所示，即在断裂前几乎没有塑性变形，其破坏方式为脆性断裂。陶瓷材料本身的脆性来自于其化学键的种类，实际陶瓷晶体中大都以方向性较强的离子键和共价键为主，多数晶体的结构复杂，平均原子间距大，因而表面能小。因此，同金属材料相比，在室温下开动的滑移系几乎没有，位错的滑移、增殖很难发生。因此很容易由表面或内部存在的缺陷引起应力集中而产生脆性破坏。这是陶瓷材料脆性的原因所在，也是其强度值分散性较大的原因所在。因此陶瓷材料室温强度测定只能获得一个断裂强度 σ_f 值。而金属材料则可获得屈服强度 σ_s 或 $\sigma_{0.2}$ 和极限强度 σ_b（图6.8-20中曲线2，3）。

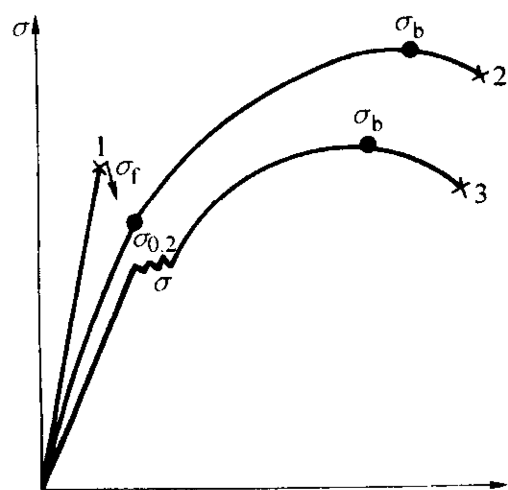


图6.8-20 陶瓷与金属的应力－应变曲线类型

由此可知，陶瓷材料的室温强度是弹性变形抗力，即当弹性变形达到极限程度而发生断裂时的应力，可采用金属材料的断裂强度计算公式进行计算。强度与弹性模量及硬度一样，是材料本身的物理参数，它决定于材料的成分及组织结构，同时也随外界条件（如温度、应力状态等）的变化而变化。

3.2.1 气孔率对强度的影响

气孔是绝大多数陶瓷的主要组织缺陷之一，气孔明显地降低了载荷作用横截面积，同时气孔也是引起应力集中的地方（对于孤立的球形气孔，应力增加一倍）。有关气孔率与强度的关系式有多种提案，其中最常用的是Ryskewitsch提出的经验公式

$$\sigma = \sigma_0\exp(-\alpha p) \tag{6.8-18}$$

式中，p 为气孔率；σ_0 为 $p=0$ 时的强度；α 为常数；其值在4～7之间，许多试验数据与此式接近。根据此关系式可推断出当 $p=10\%$ 时，陶瓷的强度就下降到无气孔时的一半。硬瓷的气孔率约为3%，陶器的气孔率约为10%～15%。当材料成分相同，气孔率的不同将引起强度的显著差异。图6.8-21示出 Al_2O_3 陶瓷的弯曲强度与气孔率之间的关系。可以看出，试验与理论值符合较好。由上述可知，为了获得高强度，应制备接近理论密度的无气孔陶瓷材料。

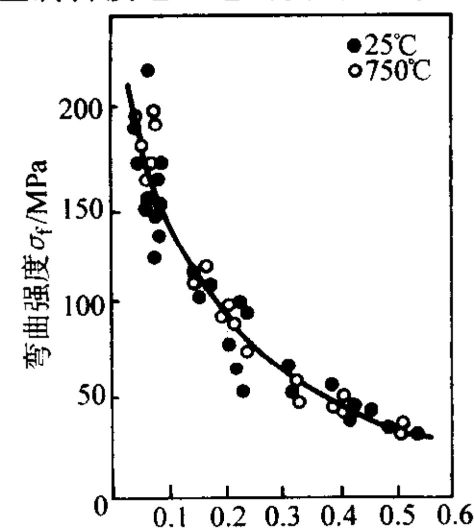


图6.8-21 Al_2O_3 的强度与气孔率的关系

3.2.2 晶粒尺寸对强度的影响

陶瓷材料的强度与晶粒尺寸的关系与金属有类似的规律，也符合Hall-Petch关系式

$$\sigma_f = \sigma_0 + kd^{\frac{1}{2}} \tag{6.8-19}$$

式中，σ_0 为无限大单晶的强度；k 为系数；d 为晶粒直径。图6.8-22给出多晶 Al_2O_3 陶瓷强度与晶粒关系；可以看出随 d 的减小强度有显著提高。图中，σ_f 与 $d^{-\frac{1}{2}}$ 的关系曲线分为两个区域，但在两区域内都成直线关系。在Ⅰ区，符合如下关系

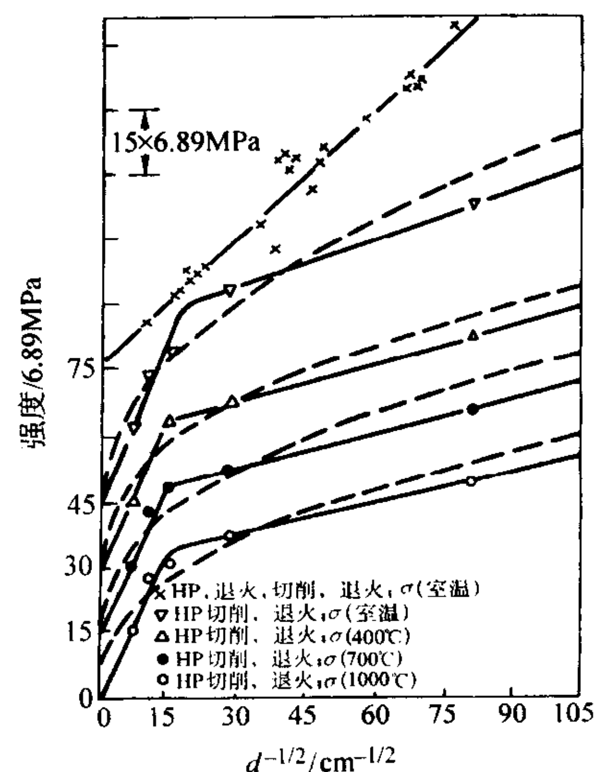


图6.8-22 多晶 Al_2O_3 强度与晶粒尺寸之间的关系

$$\sigma_f = \frac{1}{Y}\sqrt{2E\gamma^*/c} \qquad (6.8\text{-}20)$$

此时 $c \approx d$，故有 $\sigma_f \propto d^{-\frac{1}{2}}$ 的关系。在Ⅱ区，符合由金属中位错塞积（pile-up）模型推导出的滑移面剪切应力 τ_i 与位错塞积群长度 L（与晶粒 d 大小有关）之间的关系式

$$\tau_i = \tau_0 + k_s L^{-\frac{1}{2}} \qquad (6.8\text{-}21)$$

式中，τ_0 为位错运动摩擦力；k_s 为比例常数，它与裂纹形成时的表面能有关，对多晶体来说，近似地有 $\sigma_i = 2\tau$ 的关系。由于 $L \propto d$，所以有 $\sigma_f \propto d^{-\frac{1}{2}}$ 的比例关系。

从定性的角度上讲，实验研究已得到了与 $\sigma_f \propto d^{-\frac{1}{2}}$ 关系变化趋势相一致的结果。但对烧结体陶瓷来讲，要做出只有晶粒尺寸大小不同而其他组织参量都相同的试样是非常困难的，因此往往其他因素与晶粒尺寸同时对强度起影响作用。因此陶瓷中的 σ_f 与 $d^{-\frac{1}{2}}$ 的关系并非那么容易搞清，还有待于进一步研究。但无论如何，室温断裂强度无疑地随晶粒尺寸的减小而增高，所以对于结构陶瓷材料来说，努力获得细晶粒组织，对提高室温强度是有利而无害的。

3.2.3　晶界相的性质与厚度、晶粒形状对强度的影响

陶瓷材料的烧结大都要加入助烧剂，因此形成一定量的低熔点晶界相而促进致密化。晶界相的成分、性质及数量（厚度）对强度有显著影响。晶界相最好能起阻止裂纹过界扩展并松弛裂纹尖端应力场的作用。晶界玻璃相的存在对强度是不利的，所以应通过热处理使其晶化。对单相多晶陶瓷材料，晶粒形状最好为均匀的等轴晶粒，这样承载时变形均匀而不易引起应力集中，从而使强度得到充分发挥。

综上所述，高强度单相多晶陶瓷的显微组织应符合如下要求：①晶粒尺寸小，晶体缺陷少；②晶粒尺寸均匀、等轴，不易在晶界处引起应力集中；③晶界相含量适当，并尽量减少脆性玻璃相含量，应能阻止晶内裂纹过界扩展，并能松弛裂纹尖端应力集中；④减小气孔率，使其尽量接近理论密度。

3.2.4　温度对强度的影响

陶瓷材料的一个最大的特点就是高温强度比金属高得多。未来汽车用燃气发动机的预计温度为 1 370℃，这样的工作温度，Ni、Cr、Co 系的超耐热合金已无法接受，但 Si_3N_4、SiC 陶瓷却大有希望。

当温度 $T<0.5T_m$（T_m 为熔点）时，陶瓷材料的强度基本保持不变，当温度高于 $0.5T_m$ 时才出现明显的降低。Brown 等人提出图 6.8-23 所示的强度的变化曲线，可以看出整个曲线可分为三个区域。在低温 A 区，断裂前无塑性变形，陶瓷的断裂主要决定于试样内部既存缺陷（裂纹、气孔等）引起的裂纹扩展，为脆性断裂，其断裂应力为

$$\sigma_f = \frac{1}{Y}\sqrt{2E\gamma^*/c} \qquad (6.8\text{-}22)$$

式中，E、γ^* 及 c 等参数对温度不敏感，所以在 A 区 σ_f 随温度升高变化不大；在中间温度 B 区，由于断裂前产生塑性变形，因而强度对既存缺陷的敏感性降低，断裂受塑性变形控制，σ_f 随温度的上升而有明显的降低。此时的断裂应力受位错塞积机制控制，即 $\sigma_f = \sigma_0 + kd^{-\frac{1}{2}}$；当温度进一步升高时（$C$ 区），二维滑移系开动，位错塞积群中的一部分位错产生错的交叉滑移随温度的升高而变得活跃，由此而产生的对位错塞积群前端应力的松弛作用就越发明显。所以在此区域内，断裂应力有随温度的升高而上升的趋势。图 6.8-23 给出的是陶瓷材料的强度随温度变化关系的一般趋势。并非对所有的陶瓷材料都符合很好，也并非对所有陶瓷材料 A、B、C 三个区都出现。

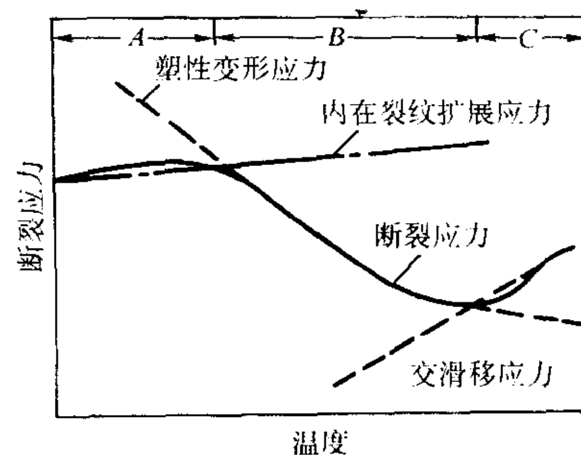


图 6.8-23　陶瓷的断裂应力与温度的依赖关系示意图

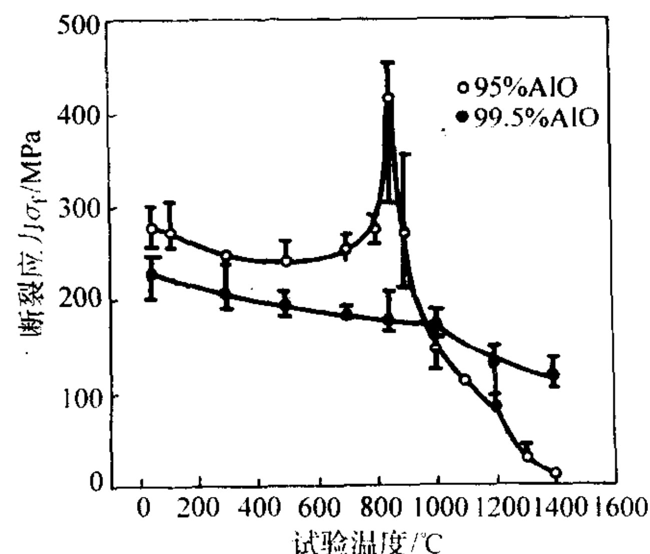


图 6.8-24　两种纯度的 Al_2O_3 的强度随温度的变化

陶瓷材料的强度随材料的纯度、微观组织结构因素及表面状态（粗糙度）的变化而变化，因此即使是同一种材料，由于制备工艺不同，其 σ_f 及其随温度的变化关系也有差异。图 6.8-24 示出两种纯度不同的 Al_2O_3 陶瓷的强度随温度的变化关系曲线。可以看出高纯 Al_2O_3 的强度变化比较简单，即随温度的升高单调下降。而低纯 Al_2O_3 陶瓷的强度在低温下高于高纯 Al_2O_3 陶瓷，且在 800℃附近出现峰值，温度在 800℃以上强度急剧下降。这是由于晶界玻璃相对致密化及愈合组织缺陷产生有利作用，因此在较低温度下玻璃相尚未软化时低纯度 Al_2O_3 的强度较高，800℃时出现的强度峰值是由于晶界玻璃相产生晶化的贡献。当温度较高时，由于玻璃相软化而使强度急剧下降。

3.3　陶瓷材料的断裂韧度

如前所述，陶瓷材料在室温下甚至在 $T/T_m \leqslant 0.5$ 的温度范围很难产生塑性变形，其断裂方式为脆性断裂，所以陶瓷材料的裂纹敏感性很强。基于陶瓷的这种特性可知，断裂力学性能是评价陶瓷材料力学性能的重要指标，同时也正是由于这种特性，其断裂行为非常适合于用线弹性断裂力学来描述。常用来评价陶瓷材料韧性的断裂力学参数就是断裂韧度（K_{IC}）。

表 6.8-8 给出一些陶瓷材料的断裂韧度值，并附常用几种金属材料的断裂韧度以作对比，可见金属材料的 K_{IC} 值比陶瓷的高一个数量级。要考虑使陶瓷材料的特长得到充分发挥，扩大在实际中的应用，就必须设法大幅度提高和改善陶瓷的韧性。

改善陶瓷韧性的方法主要有两种：一是增加陶瓷中的玻璃相，以缓冲裂纹的扩展速率；二是添加增强材料，如颗粒增强材料和纤维增强材料等，以阻止裂纹的扩展。在实际应用中，经常采用的是第二种方法，特别是晶须增强陶瓷材料的应用最为广泛。

表 6.8-8　一些陶瓷与一些金属断裂韧度值的比较

材　料	K_{IC}/MPa·m$^{1/2}$
Al_2O_3	4~4.5
Al_2O_3-ZrO_2	4~4.5
ZrO_2	1~2
ZrO_2-Y_2O_3	6~15
ZrO_2-CaO	8~10
ZrO_2-MgO	5~6
ZrO_2-CeO_2	~35
Si_3N_4	5~6
SiAlON	5~7
SiC	3.5~6
B_4C	5~7
马氏体时效钢	100
Ni-Cr-Mo 钢	45
Ti6Al4V	40
7075 铝合金	50

3.4　陶瓷材料的抗热震性

大多数结构陶瓷材料在生产和使用过程中都处于高温状态。因此，不可避免地受到温度变化的影响。材料经受温度聚变而不致破坏的能力称之为抗热震性或热稳定性。热震破坏分两大类：一类是瞬时断裂，称为热冲击断裂；另一类是在势冲击循环作用下，材料首先出现开裂、剥落，然后碎裂或变质，最后整体损坏，称为势震损伤，这里包含逐步退化过程。

陶瓷材料的抗热震能力是其力学性能和热学性能对应于各种受热条件的综合表现。材料的力学参数，如强度、断裂韧性等表征对热震破坏的抗力，而各种热环境下引起的热应力是热震破坏的动力。目前有两种观点来评价陶瓷的抗热震性：一种是基于热弹性理论，当热震温差引起的热应力超过材料的断裂应力时，导致材料瞬时断裂；另一种是基于断裂力学的概念，当热应力引起的储存于材料中的应变能足以支付裂纹扩展所需的新生表面能时，裂纹就扩展。

现以陶瓷材料的热应力和断裂强度之间的平衡为判据，分析材料在变温过程中所允许的最大温差和变温速率，简要说明在各种热震条件下，表征陶瓷材料的抗热震参数。在受热或冷却过程中，陶瓷材料出现温度梯度。在一般情况下，最大热应力 σ_{Hmax} 是多种参数的函数

$$\sigma_{Hmax} = f(m)\varphi(H)\psi(S)P(T) \tag{6.8-23}$$

式中，m 为材料特征参数，诸如力学、热学性能；H 为热处理条件，如气、液环境介质；S 为试样几何因子；T 为与温度有关的参数，如温度、变温速率、热通量、辐射温度等。对于几何形状和热处理条件相同的热震试验，φ（H）和 ψ（S）可视为常数，于是上式简化为

$$\sigma_{Hmax} = f(m)P(T) \tag{6.8-24}$$

当 σ_{Hmax} 随温度函数 P（T）变化达到材料的断裂强度 σ_f 时，相应的温度函数称为临界温度函数 P（T）。由上式得

$$P(T)_c = \sigma_f/f(m) \text{ 或 } P(T)_c = F(m)\sigma_f \tag{6.8-25}$$

临界温度函数是陶瓷材料抗热震断裂的量度。它借助于材料的力学和热学性能参数来描述，称之为抗热震参数 R。在急剧受热或冷却条件下，临界温度函数 P（T）$_c$ 就是引起临界热应力的临界温度差 ΔT_c，当一均匀试样从高温 T_1 状态下立即抛入低温 T_0 的介质中时，其表面瞬时收缩率为 α（T_0-T_1）。然而，由于保持原温度的内层并未收缩，于是表面层受到来自内层的张力为 $-E\alpha$（T_0-T_1）/（$1-\nu$），其中泊松比项（$1-\nu$）的引入是考虑多向应变导致的热应力。由此可得在急剧受热或冷却条件下，热震参数 R 为

$$R = \Delta T_c = \frac{1-\nu}{E\alpha}\sigma_f \tag{6.8-26}$$

在缓慢受热或冷却条件下，构件越接近外层则受热或冷却速率越快，而中心较小，这取决于构件的几何形状和热处理条件。在某一瞬间，构件内任一点应力取决于该点温度与构件内平均温度的差值。这时，热震参数也是临界温度差，可表示为

$$R' = \Delta T_c = \frac{Ak(1-\nu)}{E\alpha}\sigma_f = AkR \tag{6.8-27}$$

其中，A 为与构件几何形状和热处理条件相关的常数；k 为热传导率。

当构件表面以恒定速率 $\varphi = dT/dt$ 进行加热或冷却时，对于一些简单构件的表面应力可用下式表示

$$\sigma_H = \frac{E\alpha}{1-\nu} \times \frac{\varphi\gamma_m^2}{n[k/(\rho c_p)]} \tag{6.8-28}$$

式中，ρ 为材料密度；c_p 为比热容；$k/$（ρc_p）为导温系数，表征在温度变化过程中材料内部各点趋于均匀的能力；n 为构件几何形状相关参数；γ_m 为结构尺寸参数。在恒速受热或冷却条件下，热震参数是临界变温速率 φ_c，即

$$R' = \varphi_c = \frac{Ak}{\rho c_p}R \tag{6.8-29}$$

总之，根据热震条件不同，用以表征材料抗热震的参数亦不同。上述抗热震参数的共同特点是，随材料强度和热传导率的提高而增大，随着弹性模量和线胀系数的增加而降低。此外，提高材料的密度和比热容亦削弱材料的抗热震能力。表 6.8-9 列出一些陶瓷材料的力学及热学性能及其相应的抗热震参数 R 和 R'。

表 6.8-9　某些陶瓷材料的抗热震参数 R 和 R'

材　料	抗弯强度 σ_f /MPa	弹性模量 E /GPa	线胀系数 α /$10^{-6}K^{-1}$	泊松比 ν	热导率 (500℃) /kW·(m·K)$^{-1}$	R /K	R' /kW·m^{-1}
热压 Si_3N_4	850	310	3.2	0.27	17	625	11
反应烧结 Si_3N_4	240	220	3.2	0.27	15	250	3.7
反应烧结 SiC	500	410	463	0.24	84	215	18
热压 Al_2O_3	500	400	9.0	0.27	8	100	0.8
热压 BeO	200	400	8.5	0.34	63	40	2.4
烧结 WC（$w_{Co}=6\%$）	1 400	600	4.9	0.26	86	350	30

从陶瓷的生产工艺出发，人们更关心的是材料所能容忍的最大升温和冷却速率，由式（6.8-30）可得临界变温速率。

$$\varphi_c = \frac{1-\nu}{\alpha E}\left(\frac{k}{\rho c_p}\right)\left(\frac{\eta}{\gamma_m^2}\right)\sigma_f \qquad (6.8\text{-}30)$$

由此可见，当构件几何尺寸 γ_m 较大时，为保证升、降温过程中的安全，须用较小的变温速率。

3.5　陶瓷材料的疲劳

陶瓷材料疲劳比金属材料疲劳的概念含义更广。金属材料的疲劳是在循环载荷作用下的失效断裂。陶瓷材料的疲劳分为静态疲劳、动态疲劳和循环疲劳。静态疲劳是在持久载荷作用下，材料发生的失效断裂，对应于金属材料中的应力腐蚀和高温蠕变。动态疲劳是以恒定载荷速率加载，研究材料的失效断裂对加载速率的敏感性，类似于金属材料应力腐蚀研究中的慢应变速率拉伸。循环疲劳即在循环应力作用下材料的失效断裂，对应于金属中的疲劳。

对于陶瓷材料是否存在真正的循环疲劳效应，尚存在不同的看法。有些学者认为金属的疲劳效应主要是由于裂尖存在塑性区而引起的，而陶瓷材料裂尖塑性区非常有限，因此不存在循环疲劳效应。Evans 用折合算法由静态疲劳的 da/dt-K_I 曲线预测了瓷器和玻璃循环疲劳的 da/dt-K_I 曲线，预测数据和试验数据十分吻合（见图 6.8-25）。但 Ritchie 等用同样的方法，证明有明显的循环疲劳效应存在（见图 6.8-26）。看来是否存在循环疲劳效应，需视具体的材料-环境-力学体系而定。

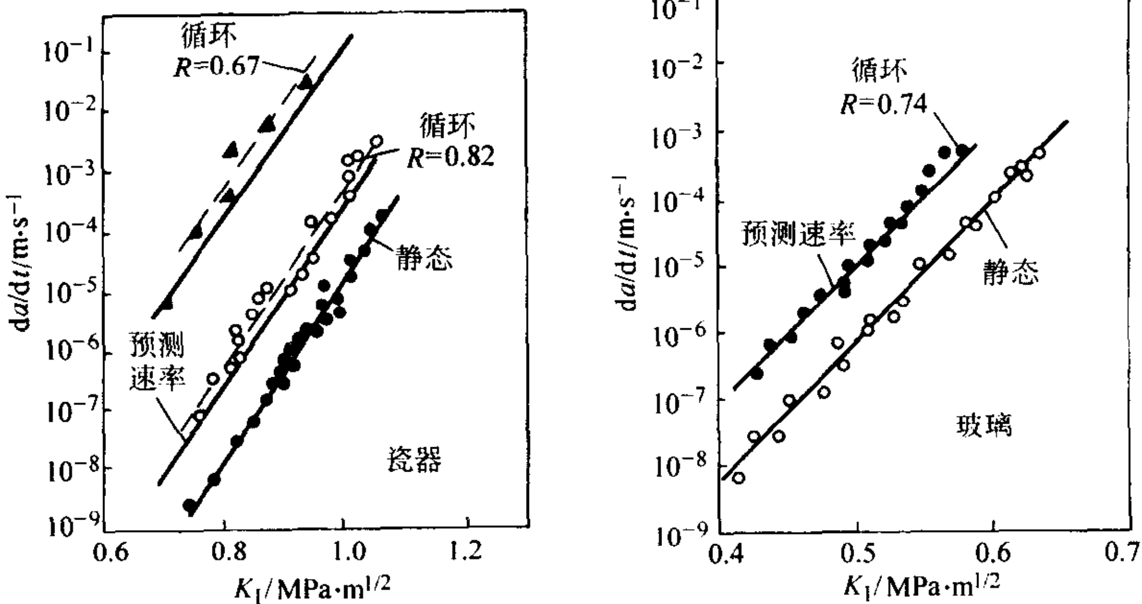


图 6.8-25　循环疲劳试验与静载预测裂纹扩展速率的比较

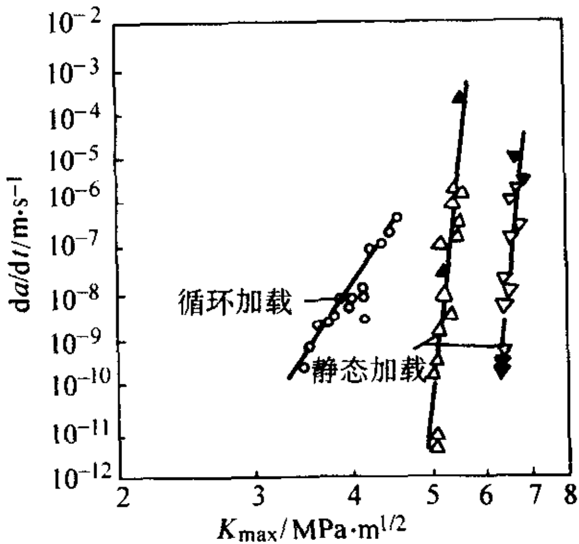


图 6.8-26　PSZ（MgO_2）材料循环疲劳与静态疲劳 da/dt-K_{max} 比较

3.6　陶瓷材料的短裂纹行为

和金属材料相比，陶瓷材料有两个特点：①由于致密性低和加工困难，陶瓷材料往往存在许多先天缺陷和微小裂纹；②由于脆性大、韧性低，陶瓷材料失稳断裂时的临界裂纹尺寸很小。因此，陶瓷材料小裂纹扩展在寿命中所占比例较大。研究表明，陶瓷材料的断裂和疲劳研究可划分为三个阶段（见图 6.8-27）。第一阶段是用光滑试样研究材料的强度和寿命，不考虑裂纹存在；第二阶段是用断裂力学方法研究含裂纹陶瓷试样的断裂，但未考虑到线弹性断裂力学当裂纹尺寸较小时求得的断裂应力将超过光滑试样的强度；第三阶段研究介乎二者之间，即小裂纹研究，研究既不能用光滑试样描述，也不能用线弹性断裂力学描述的小裂纹行为。据此，工程设计准则应为：

1）$a < a_1$，认为是光滑试样，用经典强度理论进行设计；

2）$a_1 < a < a_2$，用小裂纹断裂理论设计；

3）$a > a_2$，用线弹性断裂力学理论设计。

试验表明，陶瓷材料小裂纹现象普遍存在于短时断裂、静态疲劳、动态疲劳和循环疲劳各种断裂过程中，近年来也在陆续开展这方面的研究工作。

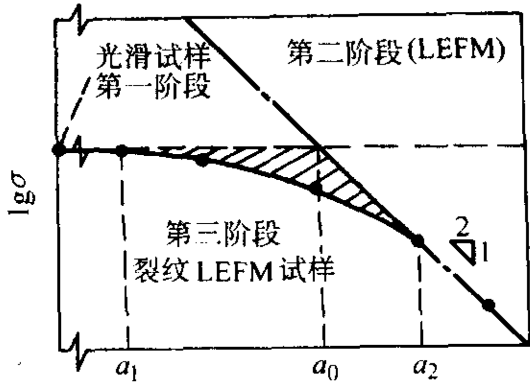


图 6.8-27　强度－裂纹尺寸效应图

编写：何雪浤（东北大学）

参考文献

1 姜伟之等编著. 工程材料的力学性能. 北京：航空航天大学出版社，2000

2 马晓青编著. 冲击动力学. 北京：北京理工大学出版社，1992

3 黄守伦主编. 实用化学热处理与表面强化新技术. 北京：机械工业出版社，2002

4 周美玲等. 材料工程基础. 北京：北京工业大学出版社，2001

5 沈莲. 机械工程材料. 北京：机械工业出版社，1999

6 李国英主编. 表面工程手册. 北京：机械工业出版社，1997

7 刘瑞堂等编. 工程材料力学性能. 哈尔滨：哈尔滨工业大学出版社，2001